Cap analogs containing 6-thioguanosine – reagents for the synthesis of mRNAs selectively photo-crosslinkable with cap-binding biomolecules

Monika Nowakowska,[a] Joanna Kowalska,[a] Franck Martin,[b] Arnaud d'Orchymont,[b] Joanna Zuberek,[a] Maciej Lukaszewicz,[a] Edward Darzynkiewicz[a,c] and Jacek Jemielity[a,c*]

[a]*Division of Biophysics, Institute of Experimental Physics, Faculty of Physics, University of Warsaw, Zwirki i Wigury 93, 02-089 Warsaw, Poland*

[b]*Architecture et Réactivité de l'ARN, Université de Strasbourg, CNRS, Institut de Biologie Moléculaire et Cellulaire, 15 rue René Descartes, 67084 Strasbourg CEDEX, France*

[c]*Centre of New Technologies, University of Warsaw, Zwirki i Wigury 93, 02-089 Warsaw, Poland; E-mail: jacekj@biogeo.uw.edu.pl, tel: +48 22 55 40 774, fax.: +48 22 55 40 771*

Electronic Supplementary Information (ESI)

List of content

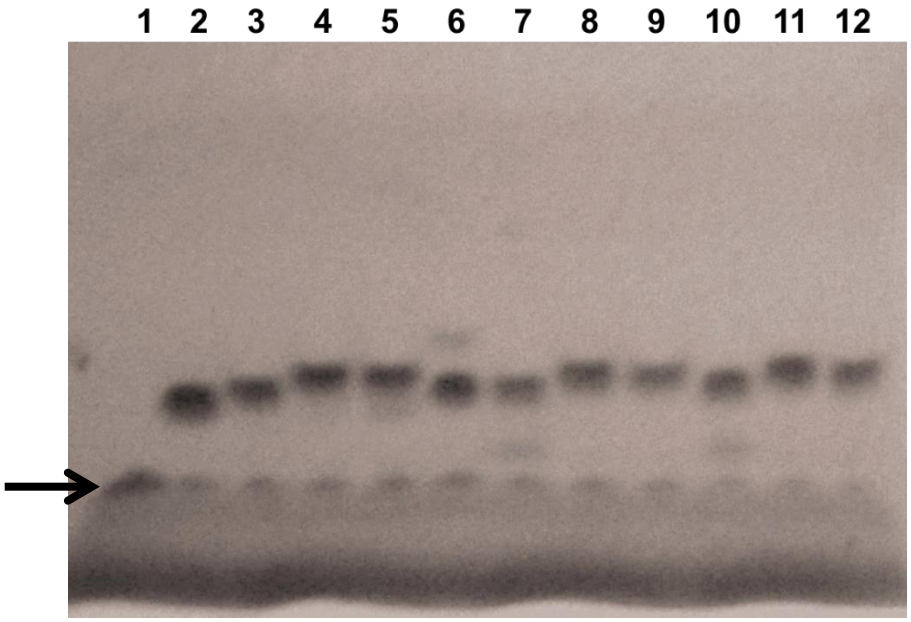


Figure S1. Electrophoretic mobility of short (6 nt) transcripts capped *in vitro* with indicated dinucleotide cap analogs. RNA transcripts were synthesized on a synthetic dsDNA template (28 bp), where the SP6 promoter sequence is directly followed by a GCCCC sequence, which undergoes transcription to RNA. A typical *in vitro* transcription reaction mixture (25 µl) contained: SP6 transcription buffer (Frementas) 3.4 µM DNA template, 1.2 U/µl SP6 RNA polymerase, 2 U/µl RiboLock Ribonuclease Inhibitor (Fermentas), 2 mM CTP, 1 mM dinucleotide cap analog and either 0.2 mM GTP (5:1 molar ratio of cap analog to GTP; lanes 1-9) or 0.1 mM GTP (10:1 cap to GTP; lanes 10, 11 and 12). The reaction was incubated at 37 °C for 60 minutes. Resultant short RNA transcripts (5 and 6 nucleotides in length) were separated on 20% polyacrylamide gel (19:1) with 7 M urea and directly after electrophoresis RNA transcripts were visualised by UV shadowing. The black arrow indicates position of a non-capped (GTP initiated) transcript.

Lane 1: pppGCCCC
Lane 2: m^7GpppGCCCC
Lane 3: iPr-m^7pppGCCCC (*Warminski et.al. 2013*)
Lane 4: iPr-m^7pp$_S$pGCCCC (D1) (*Warminski et.al. 2013*)
Lane 5: iPr-m^7pp$_S$pGCCCC (D2) (*Warminski et.al. 2013*)
Lane 6: $m_2^{7,2'\text{-O}}$Gpp$_S$pGCCCC (D2)
Lane 7: $m_2^{7,2'\text{-O}}$Gppp^{6S}GCCCC
Lane 8: $m_2^{7,2'\text{-O}}$Gpp$_S$p^{6S}GCCCC (D1)
Lane 9: $m_2^{7,2'\text{-O}}$Gpp$_S$p^{6S}GCCCC (D2)
Lane 10: $m_2^{7,2'\text{-O}}$Gppp^{6S}GCCCC (10:1)
Lane 11: $m_2^{7,2'\text{-O}}$Gpp$_S$p^{6S}GCCCC (D1) (10:1)
Lane 12: $m_2^{7,2'\text{-O}}$Gpp$_S$p^{6S}GCCCC (D2) (10:1)

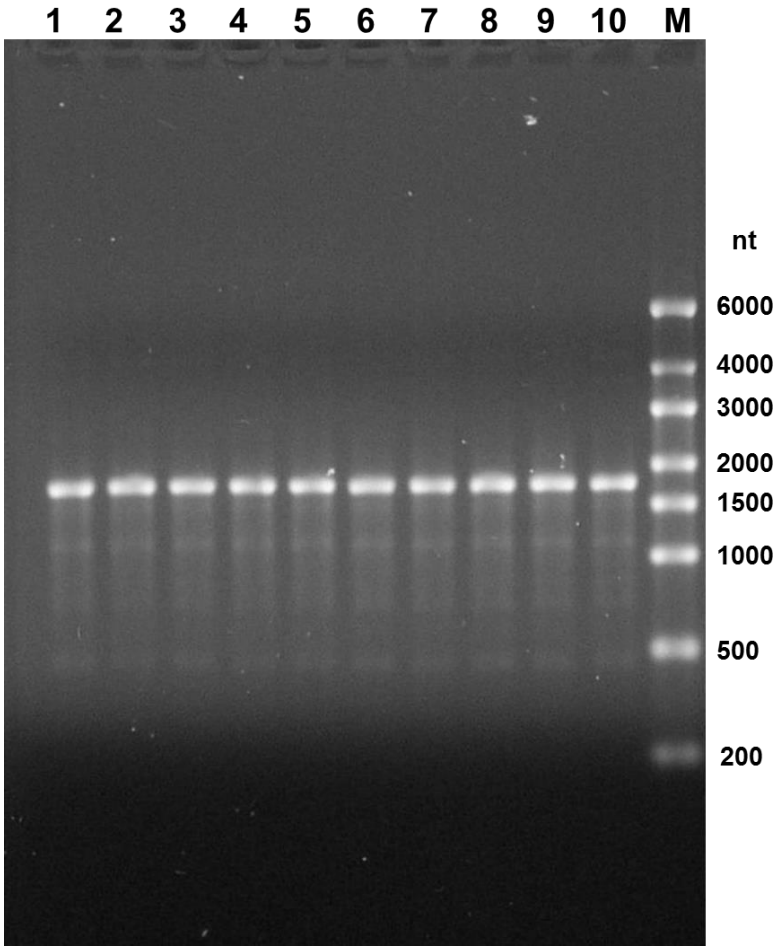


Figure S2. Analysis of *in vitro* synthesized capped luciferase mRNA transcripts.

Lane 1: m^7GpppG-luciferase-polyA$_{31}$
Lane 2: $m_2^{7,3'\text{-O}}$GpppG-luciferase-polyA$_{31}$
Lane 3: $m_2^{7,2'\text{-O}}$Gppp^{6S}G-luciferase-polyA$_{31}$
Lane 4: $m_2^{7,2'\text{-O}}$Gpp$_S$p^{6S}G -luciferase-polyA$_{31}$ (D1)
Lane 5: $m_2^{7,2'\text{-O}}$Gpp$_S$p^{6S}G -luciferase-polyA$_{31}$ (D2)
Lane 6: $m_2^{7,2'\text{-O}}$Gpp$_S$pG-luciferase-polyA$_{31}$ (D2)
Lane 7: iPr-m^7pppG-luciferase-polyA$_{31}$
Lane 8: iPr-m^7pp$_S$pG-luciferase-polyA$_{31}$ (D1)
Lane 9: iPr-m^7pp$_S$pG-luciferase-polyA$_{31}$ (D2)
Lane 10: ApppG-luciferase-polyA$_{31}$
M: RNA size marker (Fermentas)

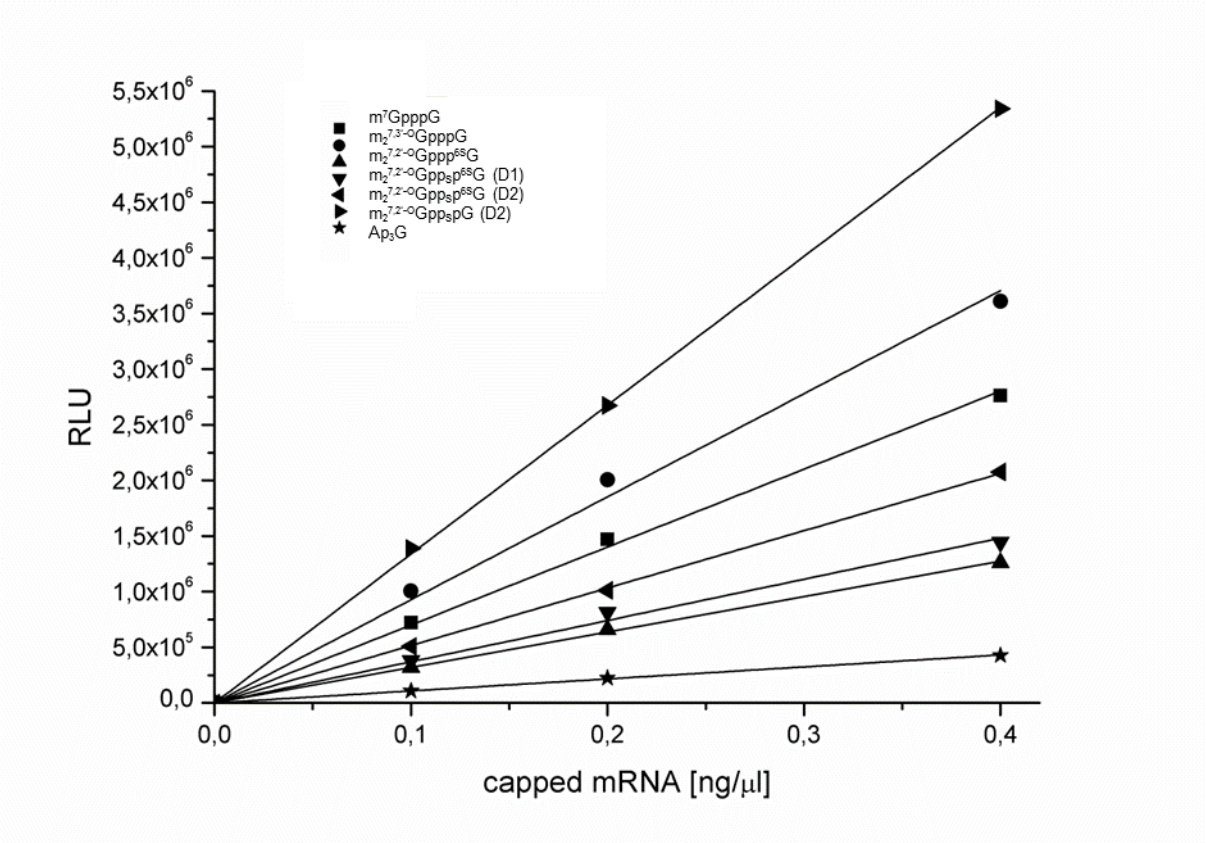
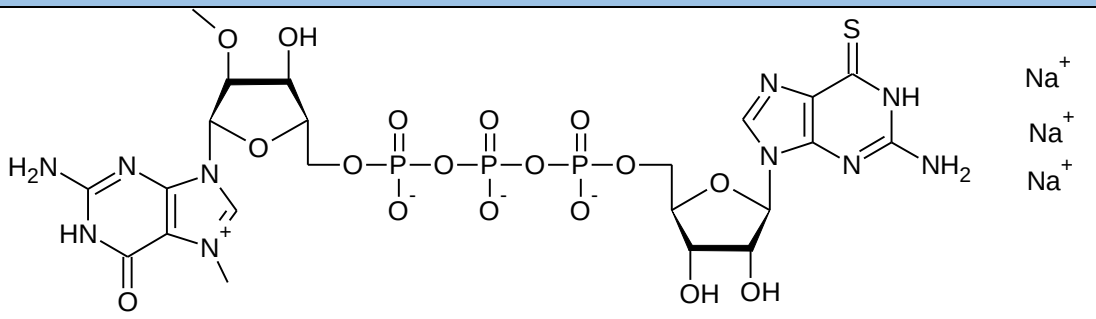
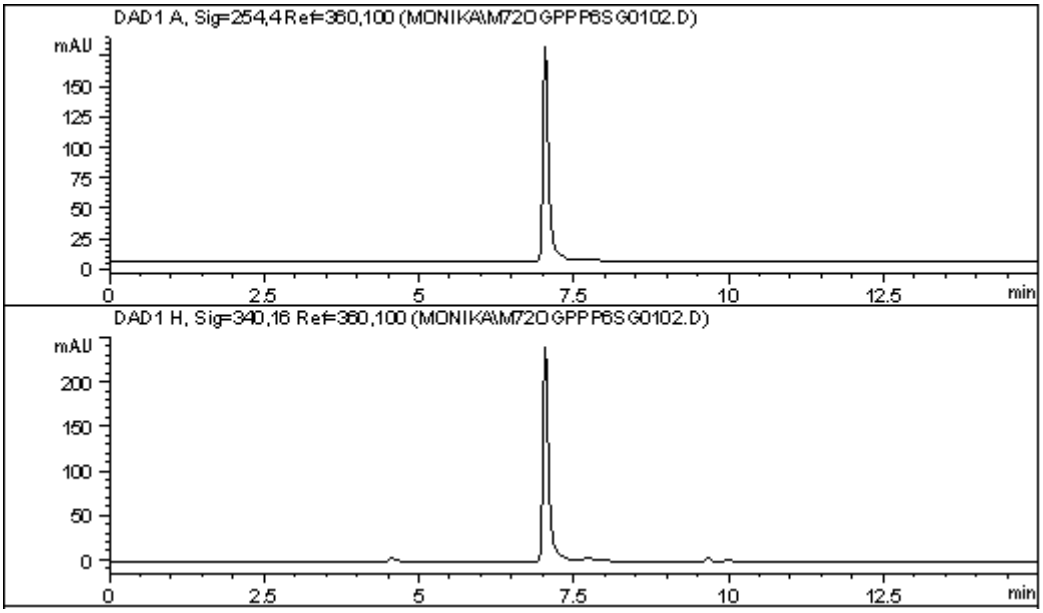
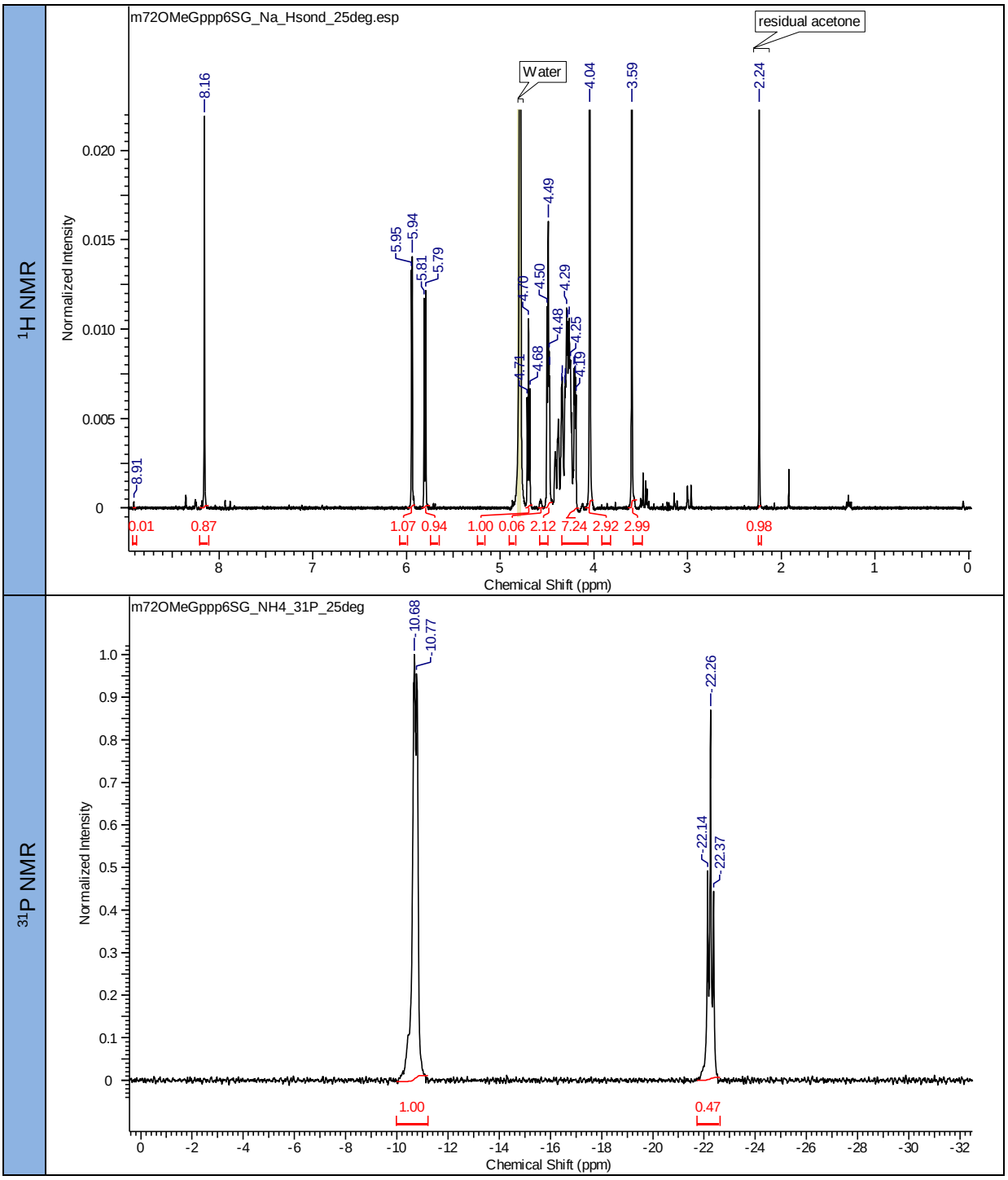
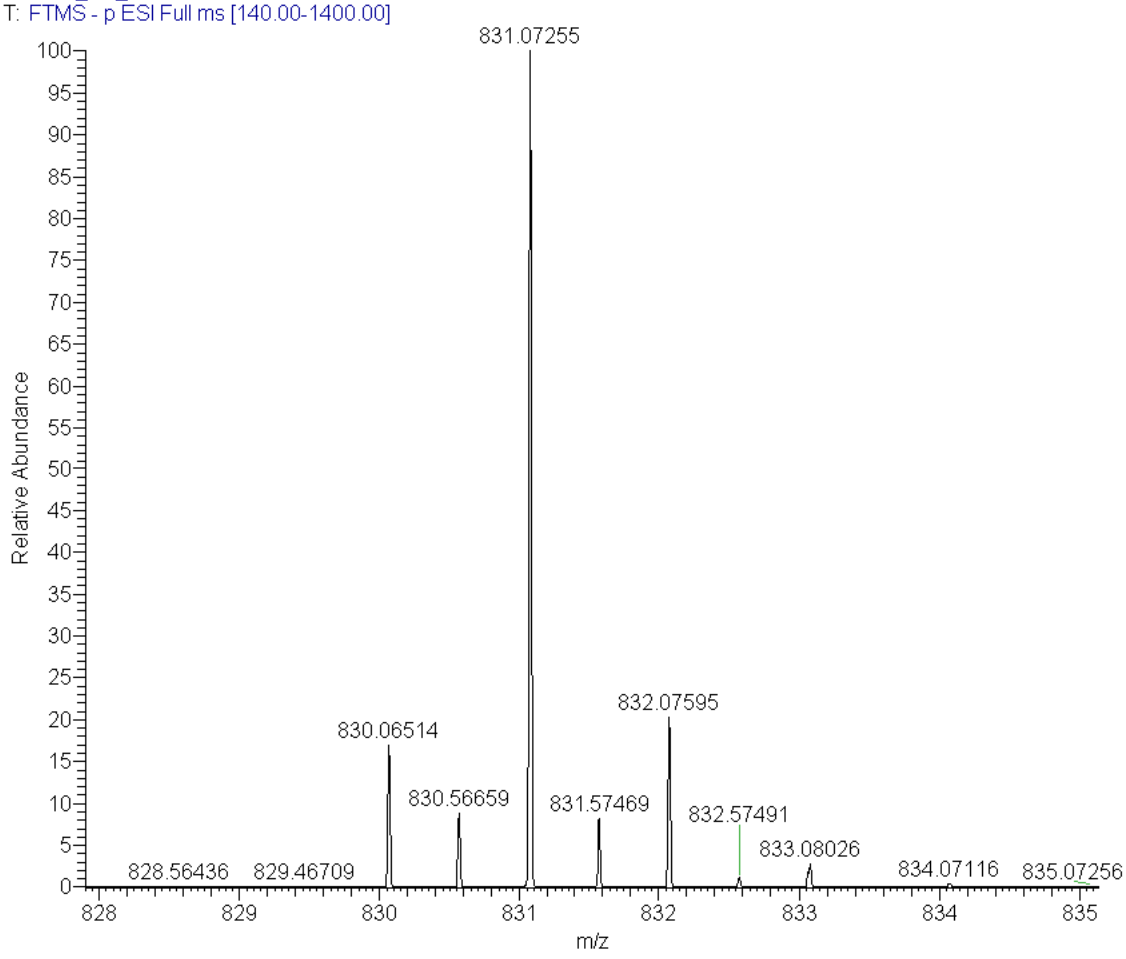


Figure S3. Representative experiment of translation efficiency in RRL lysate of luciferase mRNA transcripts capped with different cap analogs. Activity of synthesized luciferase is shown as a function of mRNA concentration.

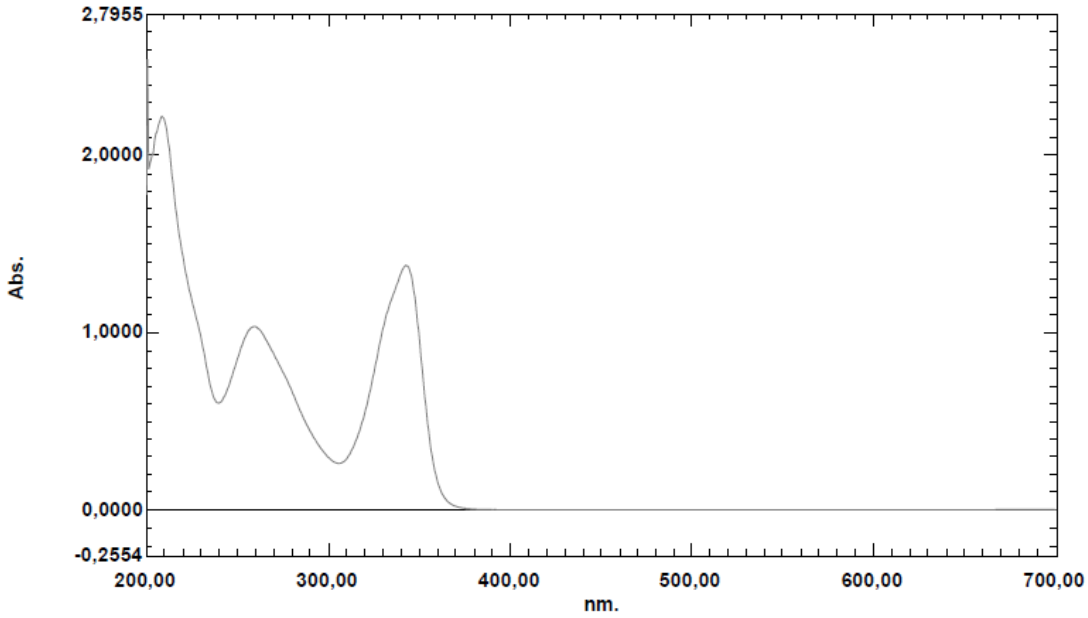
$m^{7,2\text{-}O}Gppp^{6S}G$ (**1**)	
Structure	
HPLC	


^{1}H NMR
m72OMeGppp6SG_Na_Hsond_25deg.esp
residual acetone
Water
8.16
4.04
3.59
2.24
4.49
5.95
5.94
5.81
5.79
4.70
4.50
4.29
4.48
4.25
4.71
4.68
4.19
8.91
0.01
0.87
1.07
0.94
1.00
0.06
2.12
7.24
2.92
2.99
0.98
Normalized Intensity
Chemical Shift (ppm)
^{31}P NMR
m72OMeGppp6SG_NH4_31P_25deg
-10.68
-10.77
-22.26
-22.14
-22.37
1.00
0.47
Normalized Intensity
Chemical Shift (ppm)

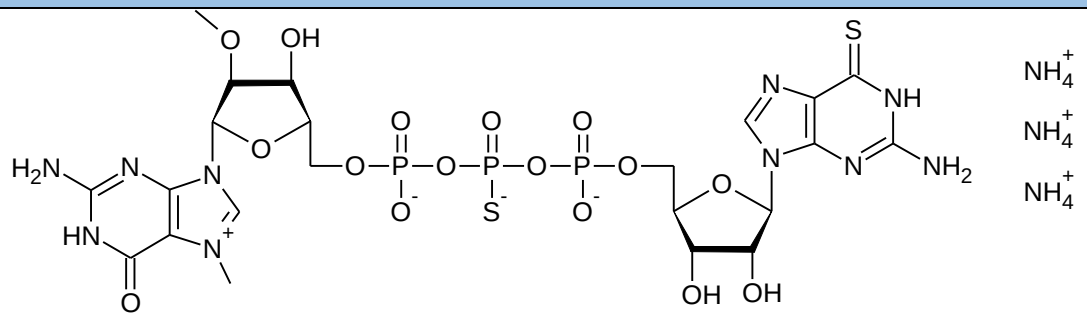
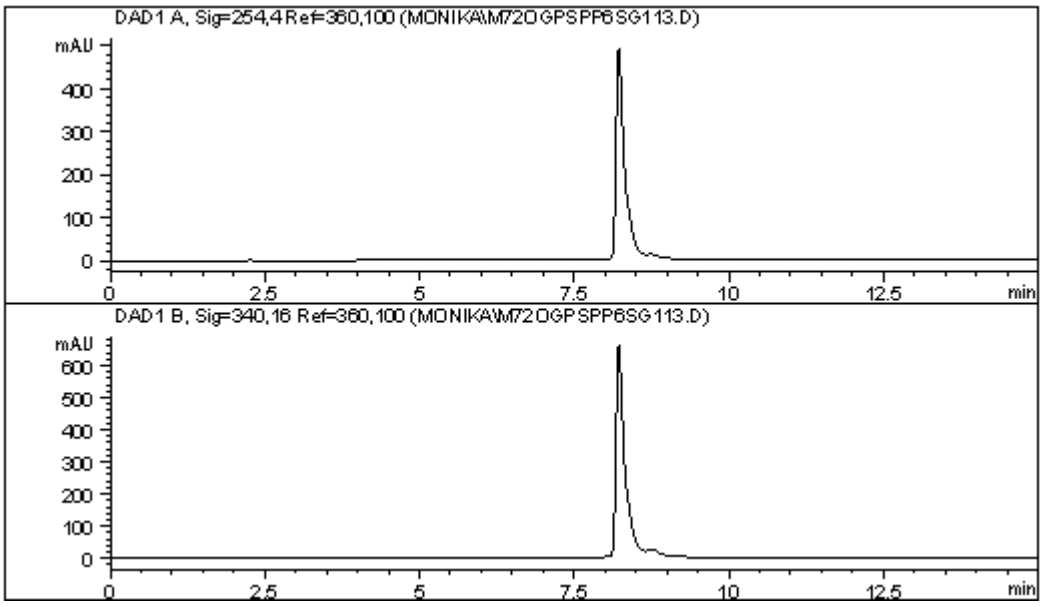
HRMS

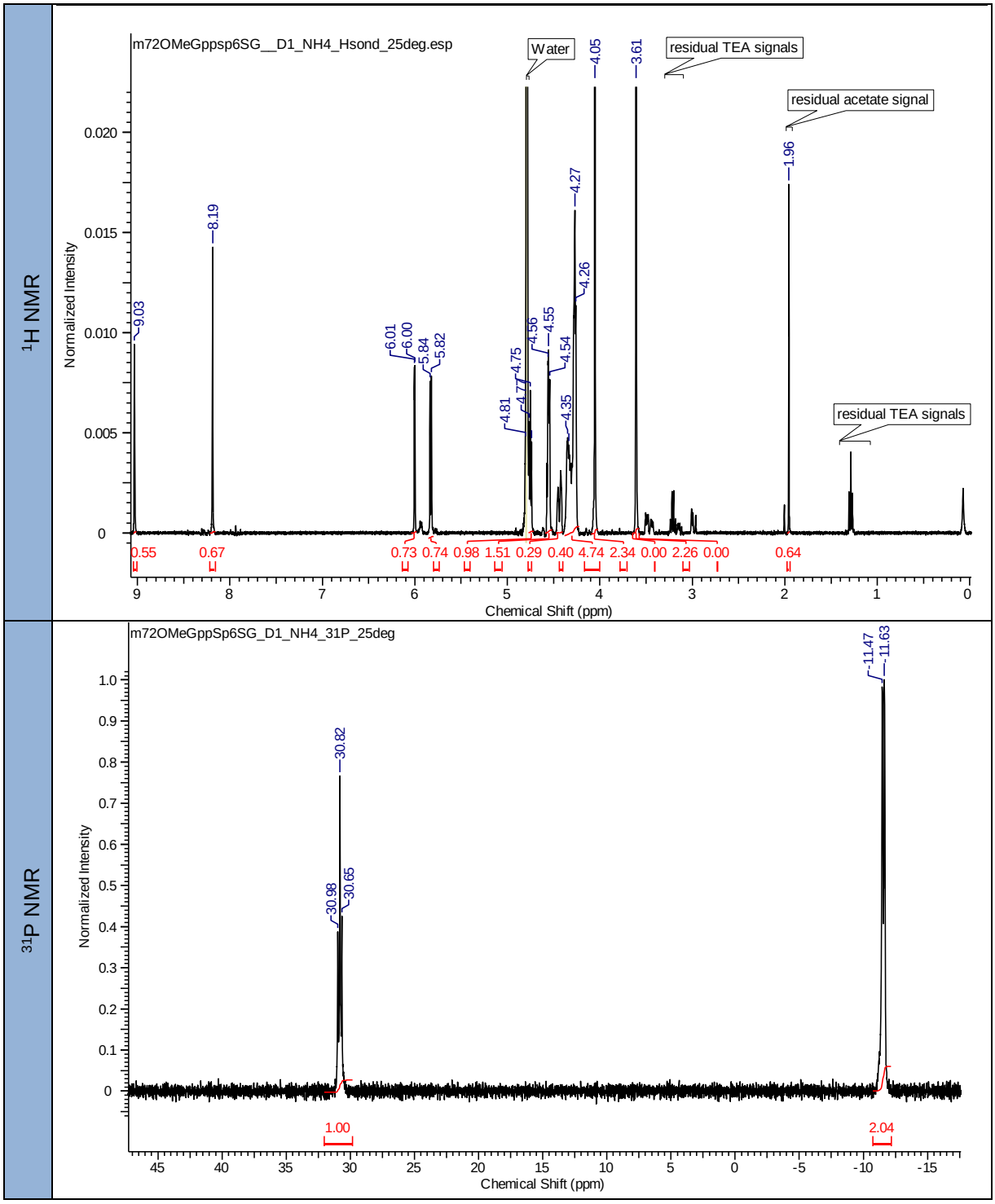


UV-VIS

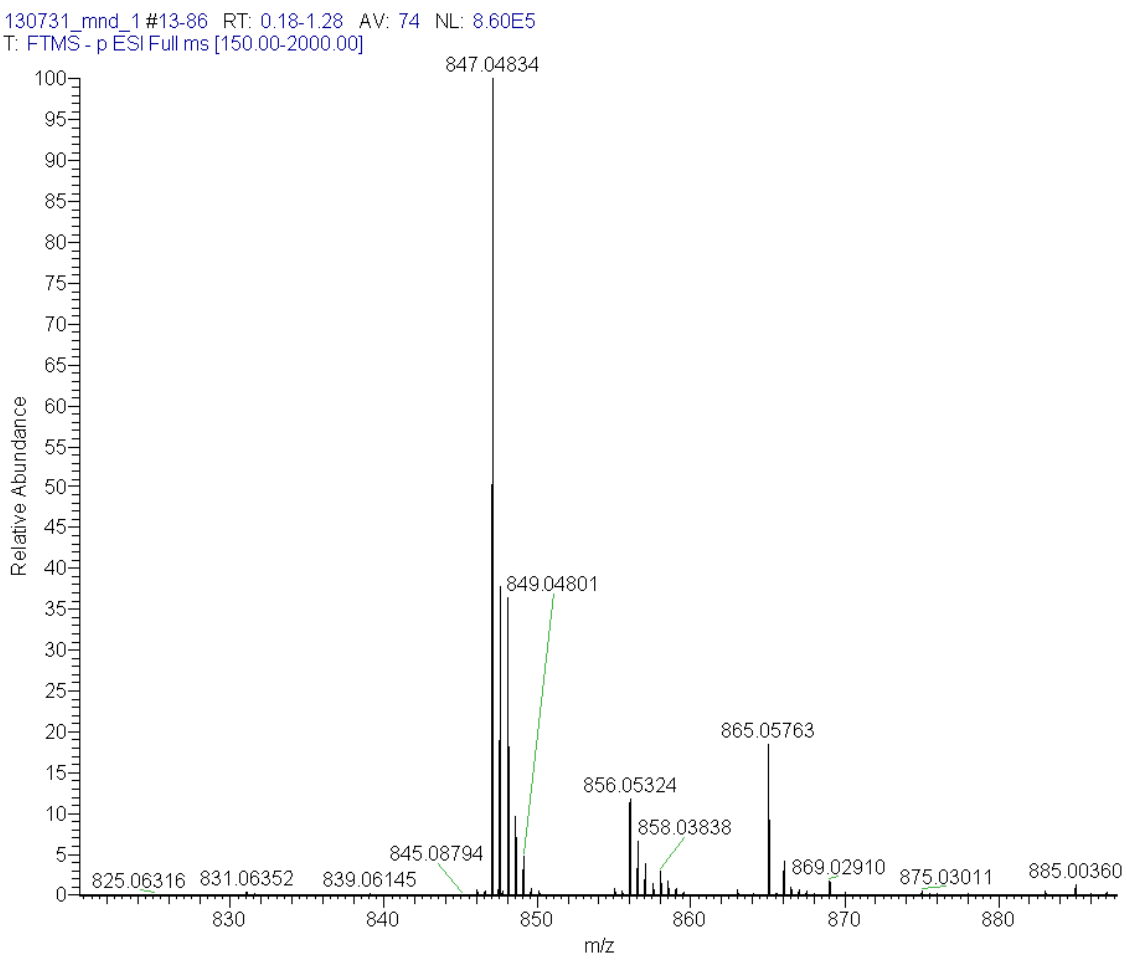


c= 0.033 mmol/ L, λ=342.50 A=0.8204, λ=259.50 A=0.5820, λ=208.00 nm A=1.2872

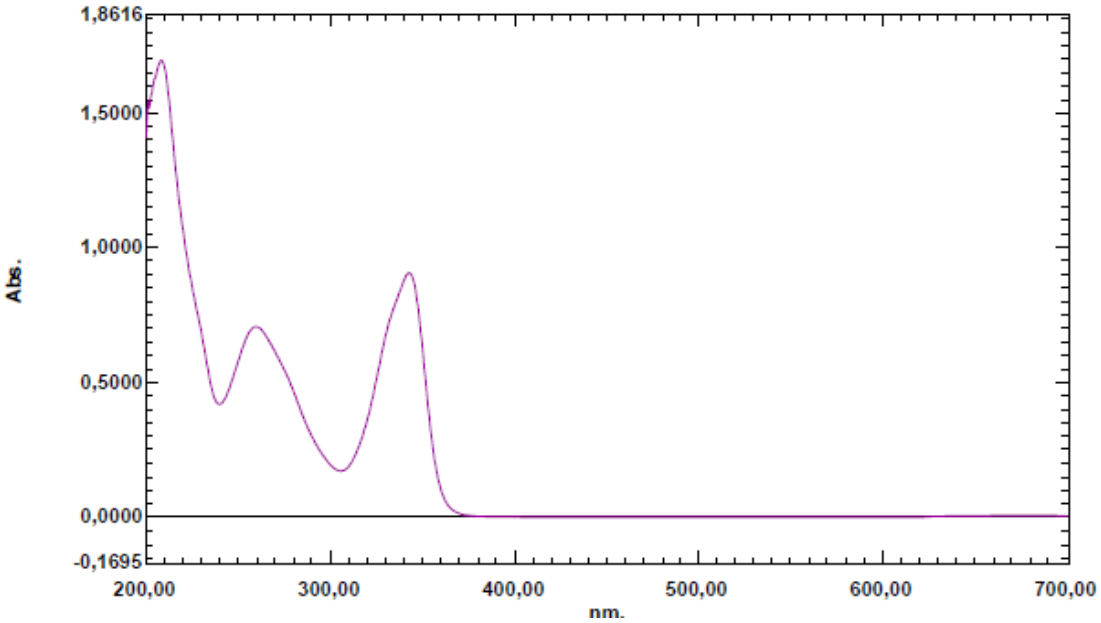
$m^{7,2\text{-}O}GppSp^{6S}G$ D1 (**2a**)	
Structure	
HPLC	


^{1}H NMR
m72OMeGppsp6SG__D1_NH4_Hsond_25deg.esp
Water
residual TEA signals
residual acetate signal
residual TEA signals
Normalized Intensity
Chemical Shift (ppm)
^{31}P NMR
m72OMeGppSp6SG_D1_NH4_31P_25deg
Normalized Intensity
Chemical Shift (ppm)

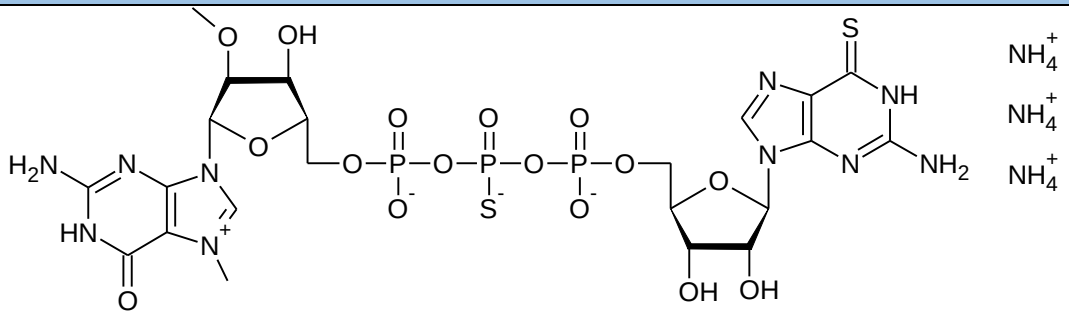
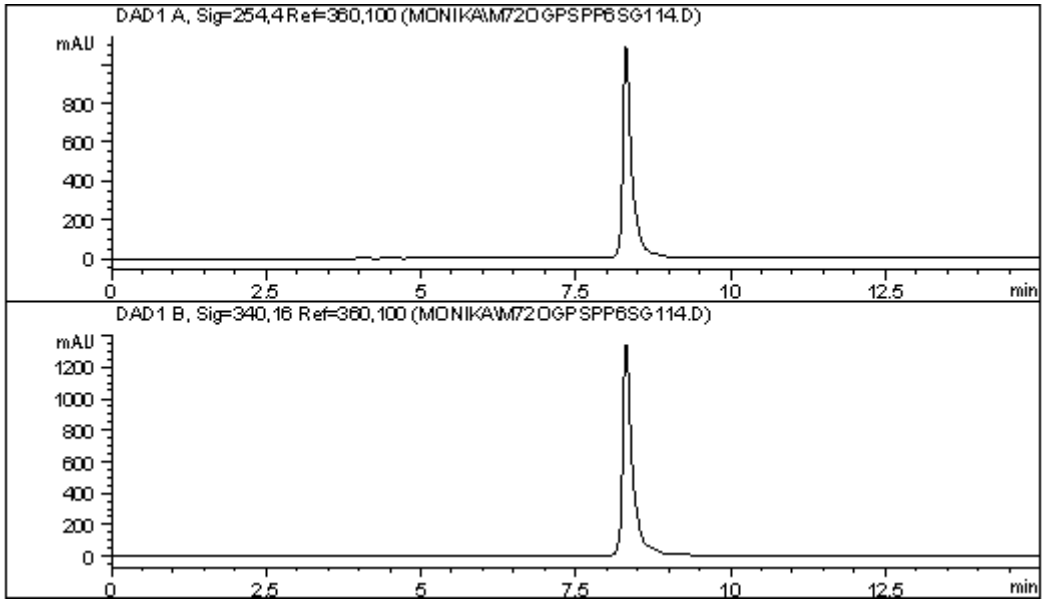
HRMS

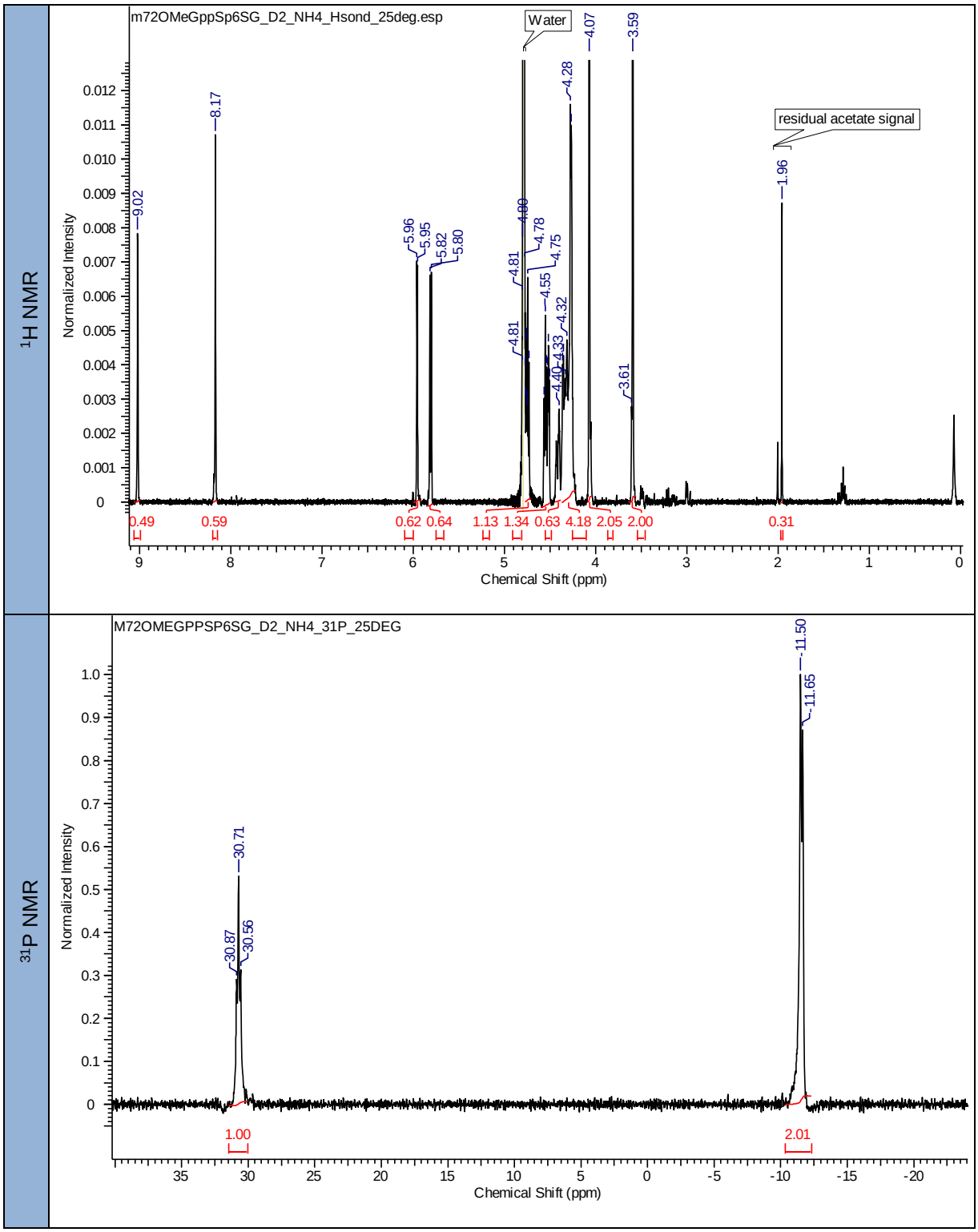


UV-VIS

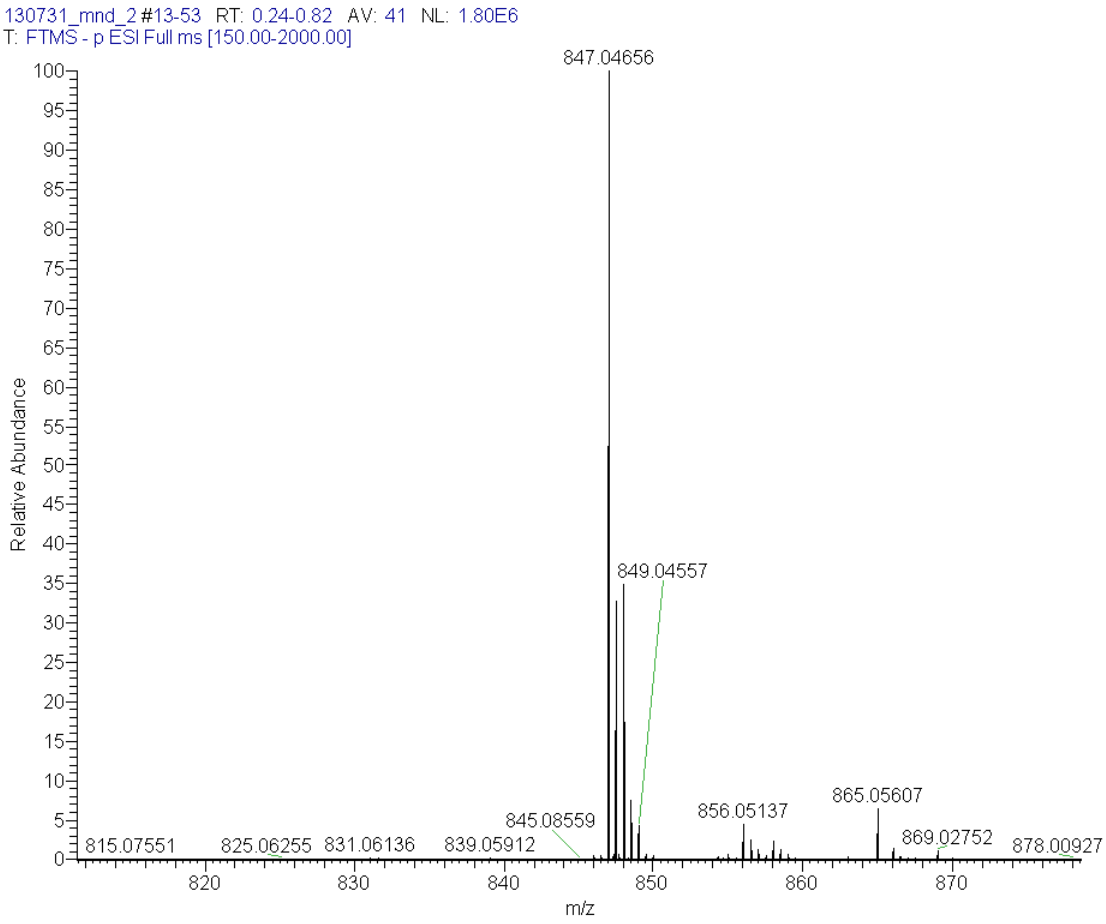


c= 0.036 mmol/L, λ=342.50 nm A=0.9047, λ=259.50 nm A=0.7054, λ=208.50 nm A=1.6924

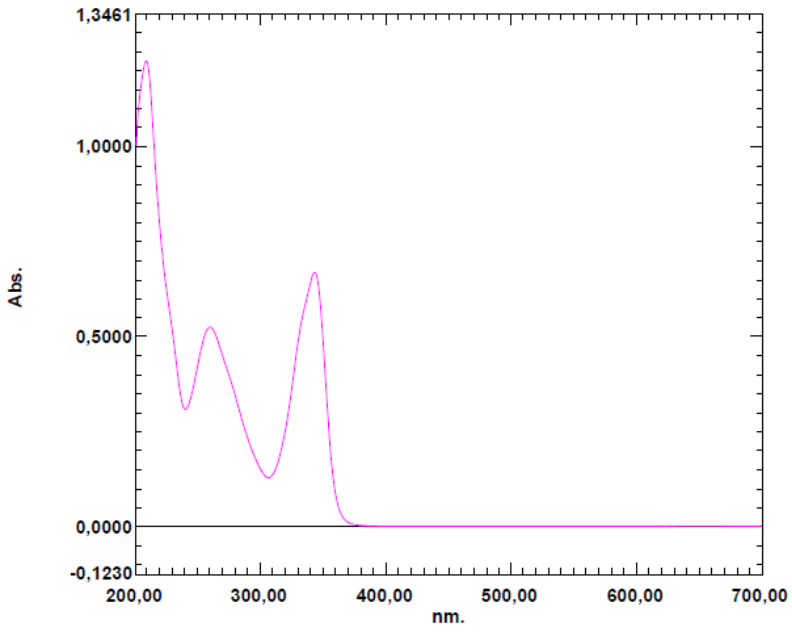
$m^{7,2\text{-}O}Gpp_Sp^{6S}G$ D2 (**2b**)	
Structure	
HPLC	


^{1}H NMR
m72OMeGppSp6SG_D2_NH4_Hsond_25deg.esp
Water
residual acetate signal
Normalized Intensity
Chemical Shift (ppm)
^{31}P NMR
M72OMEGPPSP6SG_D2_NH4_31P_25DEG
Normalized Intensity
Chemical Shift (ppm)

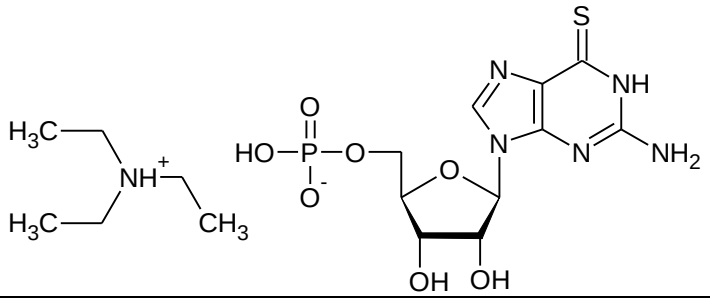
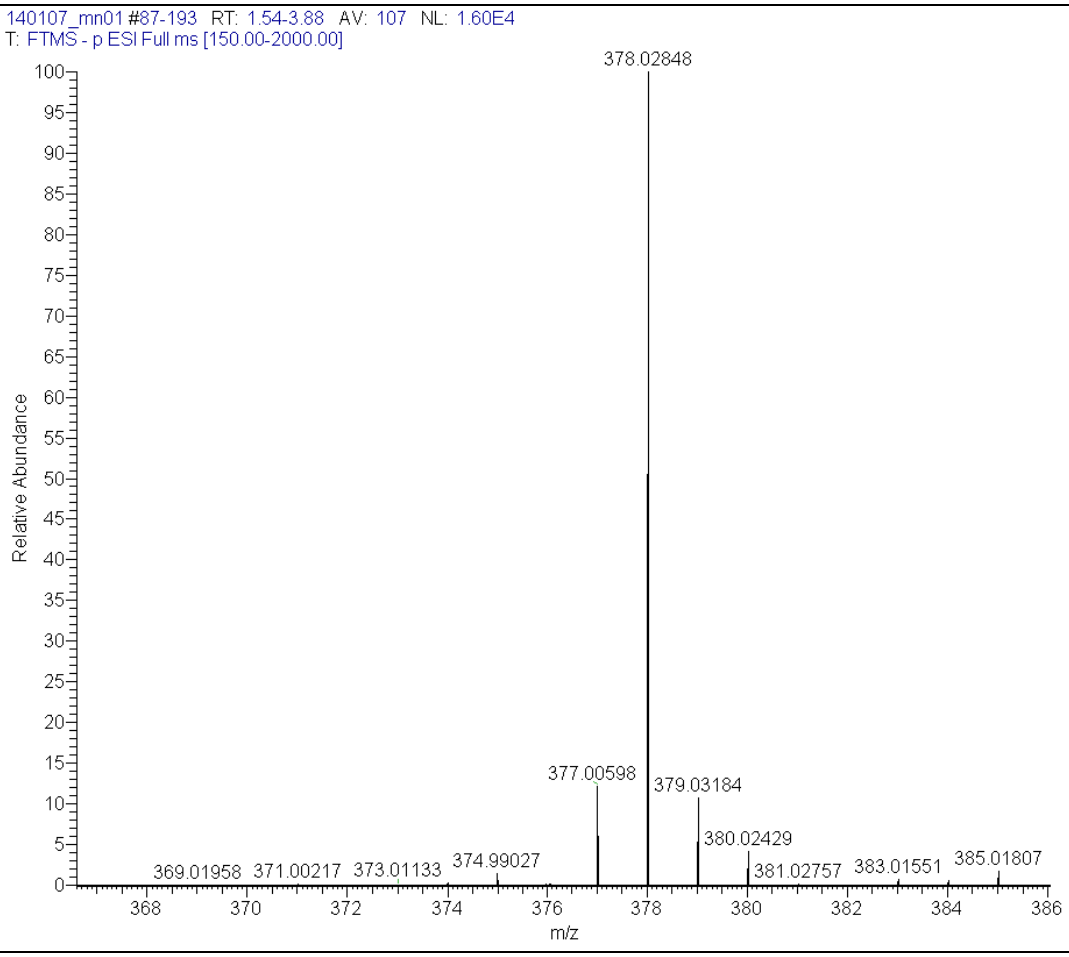
HRMS

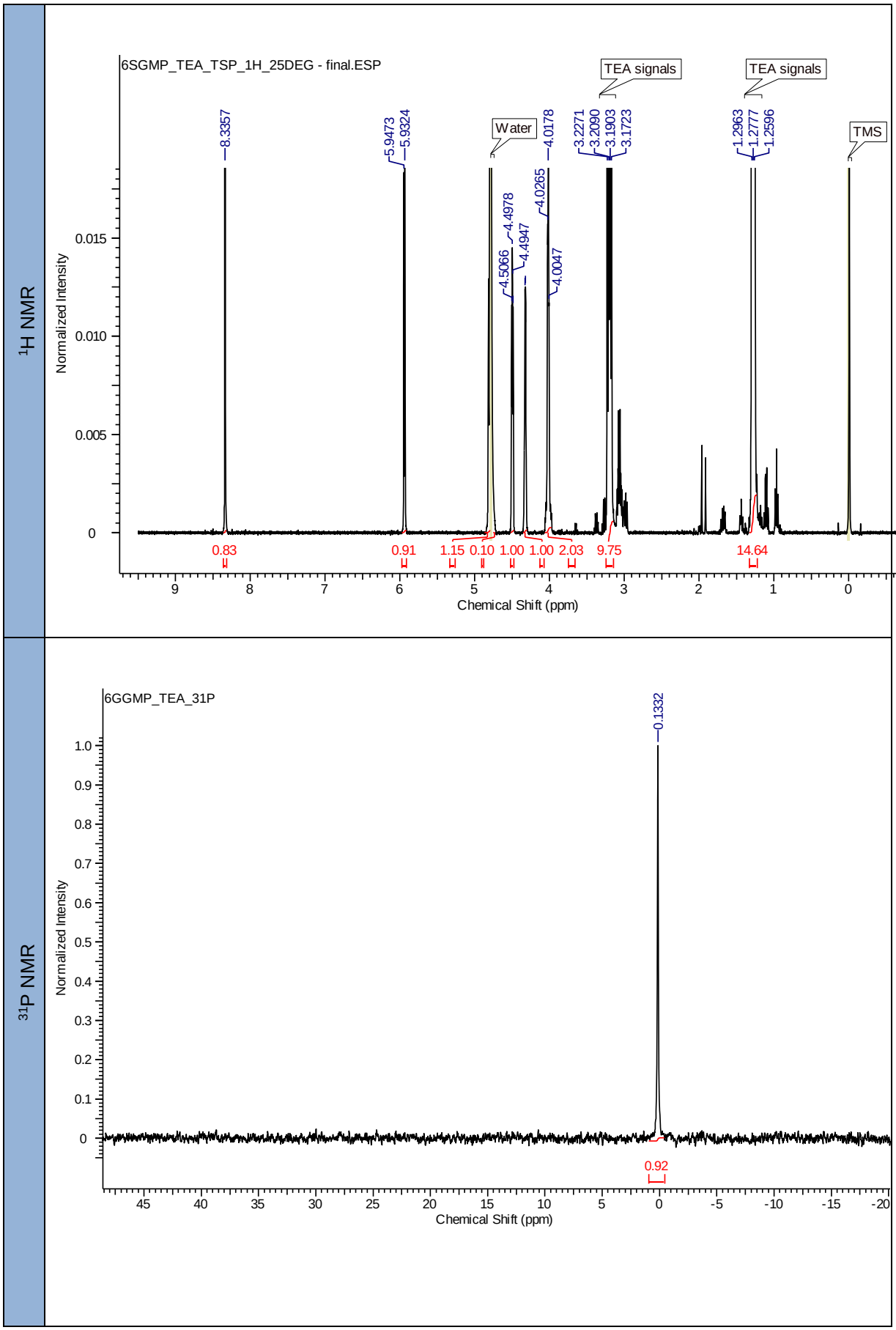


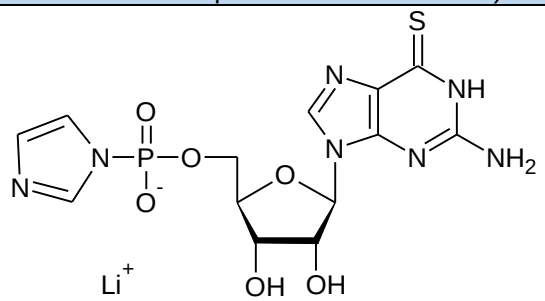
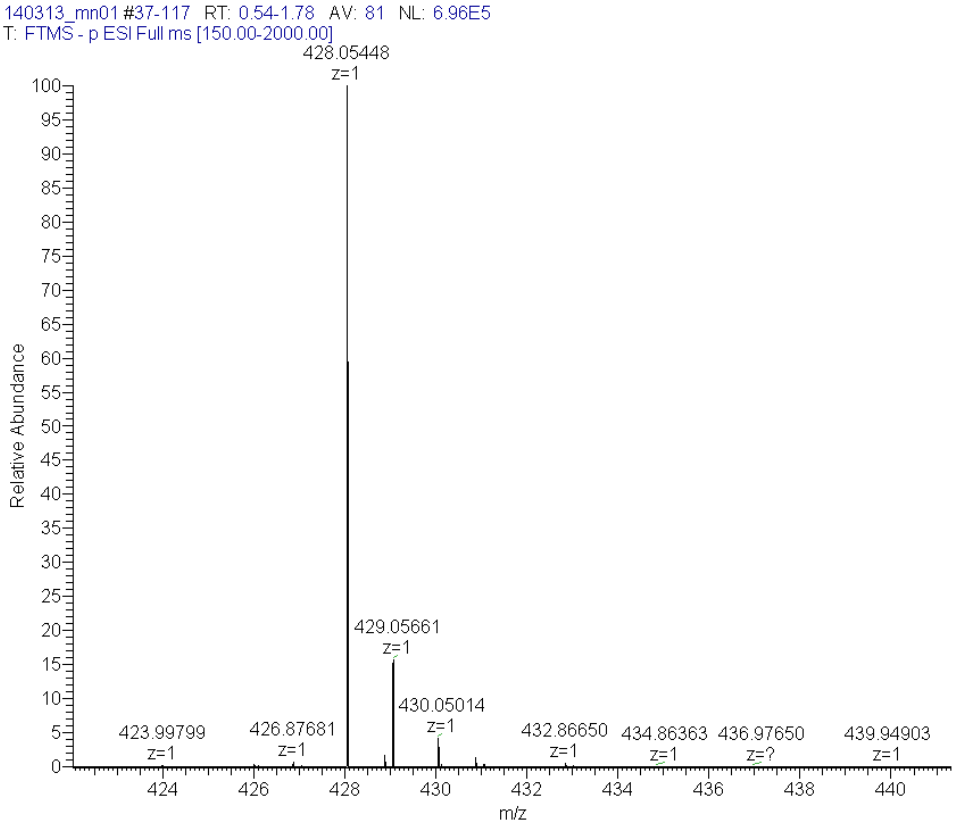
UV-VIS

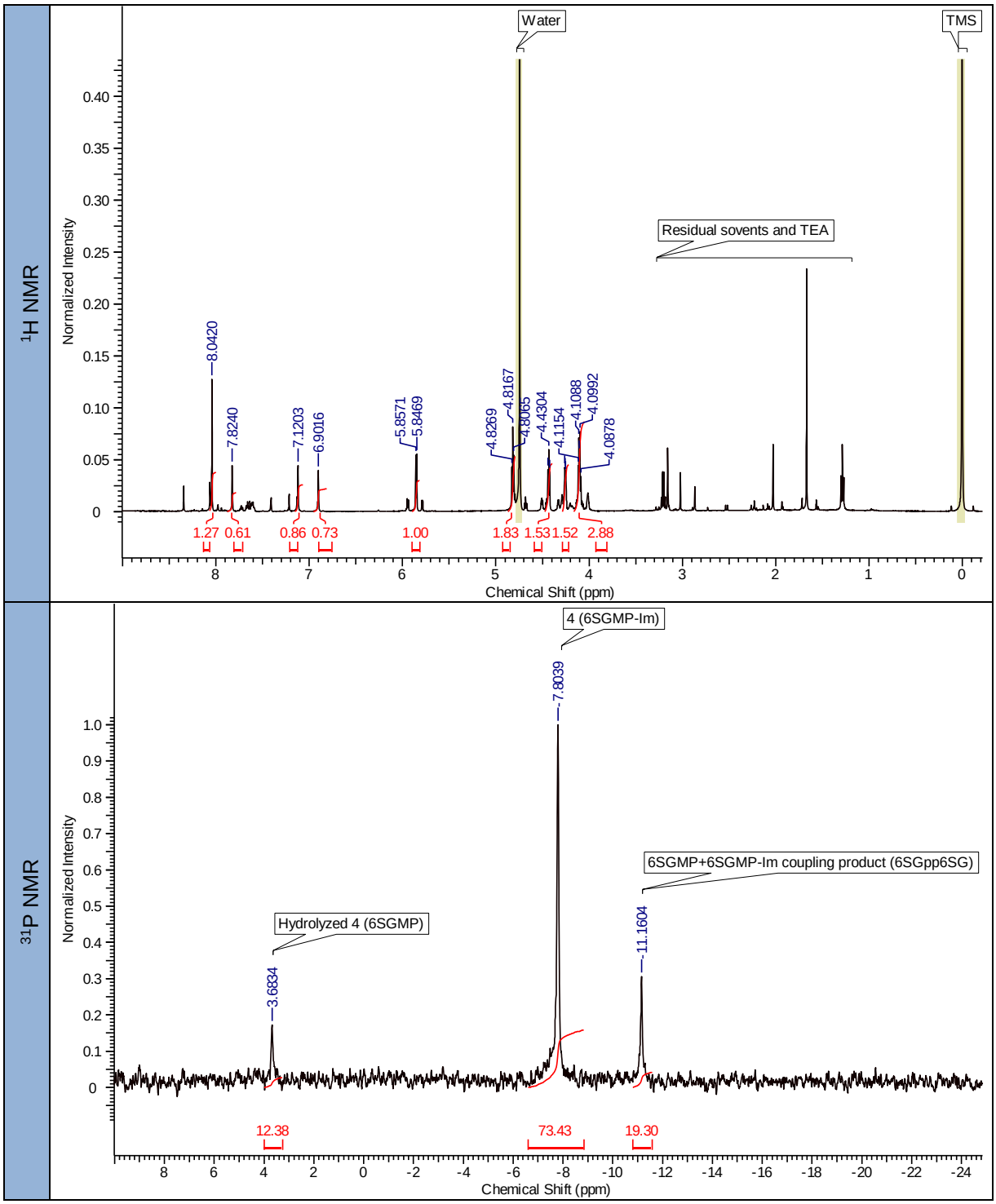


c= 0.027 mmol/L, λ=343.00 nm A=0.6676, λ=259.00 nm A=0.5238, λ=208.50 nm A=1.2236

	^{6S}GMP (**3**)
Structure	
HRMS	


^{1}H NMR
6SGMP_TEA_TSP_1H_25DEG - final.ESP
TEA signals
TEA signals
Water
TMS
8.3357
5.9473
5.9324
4.0178
3.2271
3.2090
3.1903
3.1723
1.2963
1.2777
1.2596
4.0265
4.4978
4.4947
4.5066
4.0047
Normalized Intensity
0.83
0.91
1.15
0.10
1.00
1.00
2.03
9.75
14.64
Chemical Shift (ppm)
^{31}P NMR
6GGMP_TEA_31P
0.1332
Normalized Intensity
0.92
Chemical Shift (ppm)

	^{6S}GMP-Im (**4**) (The compound is unstable in water and, therefore, domposition products were observed in ^{1}H and ^{31}P NMR spectra refcorded in D_2O)
Structure	
HRMS	


^{1}H NMR
Water
TMS
Residual sovents and TEA
Normalized Intensity
8.0420
7.8240
7.1203
6.9016
5.8571
5.8469
4.8269
4.8167
4.8065
4.4304
4.1154
4.1088
4.0992
4.0878
1.27 0.61 0.86 0.73 1.00 1.83 1.53 1.52 2.88
Chemical Shift (ppm)
^{31}P NMR
4 (6SGMP-Im)
7.8039
6SGMP+6SGMP-Im coupling product (6SGpp6SG)
Hydrolyzed 4 (6SGMP)
11.1604
3.6834
12.38 73.43 19.30
Normalized Intensity
Chemical Shift (ppm)