Supporting Information

Optimization of an Unnatural Base Pair Towards Natural-Like Replication

Young Jun Seo, Gil Tae Hwang, Phillip Ordoukhanian, and Floyd E. Romesberg*

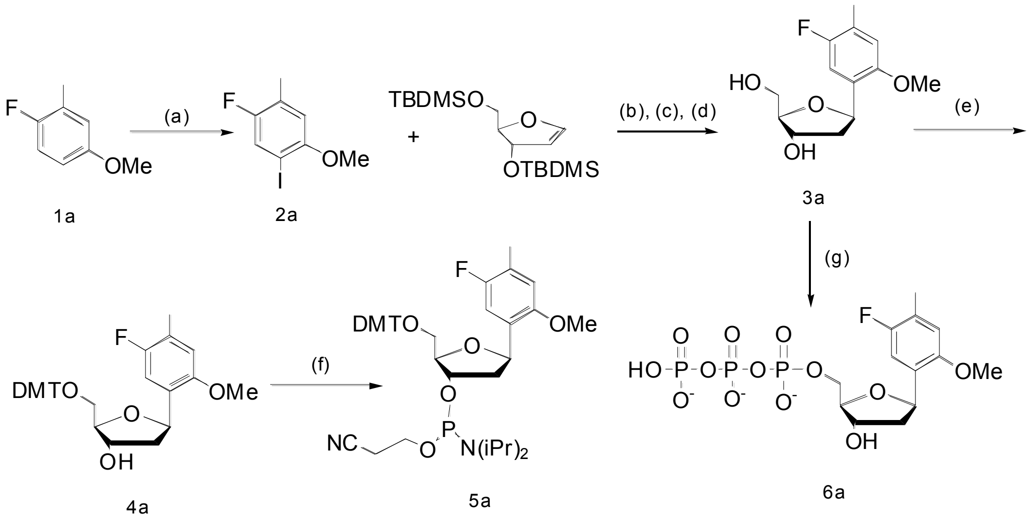
Department of Chemistry, The Scripps Research Institute, 10550 N. Torrey Pines Road, La Jolla, California, 92037

NMR and HRMS Spectrum of Novel Compounds

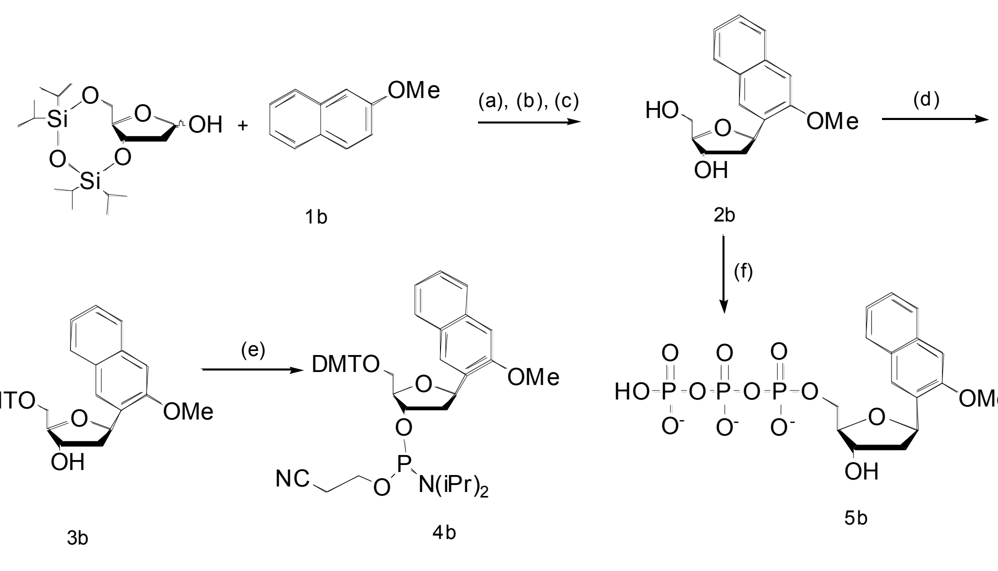
MALDI-TOF mass spectral data[a]

Sequence	MALDI-TOF signal [M+] Calc. *m*/*z*	Found *m*/*z*	Page
P1: 5'-d(TAATACGACTCACTATAGGGAGA**NaM**)	7416	7413	S44
P2: 5'-d(TAATACGACTCACTATAGGGAGA**FM**)	7398	7394	S45
P3: 5'-d(TAATACGACTCACTATAGGGAGA**5SICS**)	7433	7426	S46
P4: 5'-d(TAATACGACTCACTATAGGGAGA**MMO2**)	7380	7374	S47
T1:3'-d(ATTATGCTGAGTGATATCCCTCT**NaM**GCTAGGTTA CGGCAGGATCGC)	13890	13893	S48
T2:3'-d(ATTATGCTGAGTGATATCCCTCT**FM**GCTAGGTTA CGGCAGGATCGC)	13872	13876	S49
T3:3'-d(ATTATGCTGAGTGATATCCCTCT**5SICS**GCTAGGTTA CGGCAGGATCGC)	13907	13907	S50
T4:3'-d(ATTATGCTGAGTGATATCCCTCT**MMO2**GCTAGGTTA CGGCAGGATCGC)	13854	13857	S51
P5: 5'-d(TAATACGACTCACTATAGGGAGC**NaM**)	7393	7396	S52
P6: 5'-d(TAATACGACTCACTATAGGGAGC**FM**)	7375	7382	S53
P7: 5'-d(TAATACGACTCACTATAGGGAGC**5SICS**)	7410	7414	S54
P8: 5'-d(TAATACGACTCACTATAGGGAGC**MMO2**)	7357	7361	S55
T5:3'-:d(ATTATGCTGAGTGATATCCCTCG**NaM**TCTAGGTTA CGGCAGGATCGC)	13890	13891	S56
T6:3'-:d(ATTATGCTGAGTGATATCCCTCG**FM**TCTAGGTTA CGGCAGGATCGC)	13872	13877	S57
T7:3'-:d(ATTATGCTGAGTGATATCCCTCG**5SICS**TCTAGGTTA CGGCAGGATCGC)	13907	13888	S58
T8:3'-:d(ATTATGCTGAGTGATATCCCTCG**MMO2**TCTAGGTTA CGGCAGGATCGC)	13854	13852	S59
T9: 3'-d(ATTATGCTGAGTGATATCCCTCA**5SICS**TCTAGGTTA CGGCAGGATCGC)	13916	13915	S60
T10: 3'-d(ATTATGCTGAGTGATATCCCTCC**5SICS**TCTAGGTTA CGGCAGGATCGC)	13892	13890	S61

[a] **P1-P4**: primers for context I, **T1-T4**: templates for context I, **P5-P8**: primers for context II, **T5-T8**: templates for context II and **T9-T10:** templates for context III and IV.



Scheme S1. (a) I_2, $AgNO_3$, MeOH, yield:73%; (b) $Pd(OAc)_2$, $AsPh_3$, *n*-Bu_3N, DMF, 70 ^oC ; (c) *n*-TBAF, 0 ^oC; (d) $NaBH(OAc)_3$, AcOH, CH_3CN, overall yield:52%; (e) DMTrCl, DMAP, Et_3N, pyridine, yield:78%; (f) $NC(CH)_2OP(iPr_2N)_2$, DIPEA, CH_2Cl_2, yield:85%; (g) Proton Sponge, $POCl_3$, Bu_3N, Bu_3NPPi, $(MeO)_3P$, DMF, -20 ºC.



Scheme S2. (a) 1.6 M n-BuLi, THF, -10 ^oC; (b) n-Bu_3P, TMAD, THF, rt ; (c) TBAF, THF, rt, overall yield: 22%; (d) DMTrCl, DMAP, Et_3N, pyridine, yield: 82%; (e) $NC(CH)_2OP(iPr_2N)_2$, DIPEA, CH_2Cl_2, yield: 92%; (f) Proton Sponge, $POCl_3$, Bu_3N, Bu_3NPPi, $(MeO)_3P$, DMF, -20 ºC.

Synthetic procedures and characterizations

Compound 2b. A mixture of 2-methoxy naphatalene (700mg, 4.4 mmol) in THF (10 mL) was stirred under nitrogen atmosphere at room temperature for 30 minutes, after which 1.6 M *n*-BuLi (2.7ml, 28.5mmol) was added dropwise and the mixture was stirred for 2 h at room temperature. After cooling to -10 ^{o}C, 3,5-O-[1,1,3,3-tetrakis(1-methylethyl)-1,3-disiloxanediyl] 2-deoxy-D-erythro-pent-1-enitol (537 mg, 1.43 mmol) was added and the temperature was increased slowly to room temperature and the mixture stirred for 30 min. The reaction mixture was extracted with EtOAc (50 mL × 3) and saturated NH_4Cl aq. (50 mL) and the combined organic layers were dried over anhydrous Na_2SO_4. After filtration and evaporation, the residue was purified by silica gel column chromatography (EtOAc/hexane (1:2), R_f = 0.38). The eluted product, without further purification, was dissolved in THF (3 mL) and tri-n-butyl phosphin (87 μl, 0.35 mmol) was added. The resulting solution was cooled to 0 °C, *N*,*N*,*N′*,*N′*-tetramethylazodicarboxamide (60 mg, 0.35 mmol) was added, and the solution was stirred overnight at room temperature. Tetrabutylammoniumfluoride (0.45 mL, 0.45 mmol) was added to this mixture at room temperature and stirred for 3 h. The reaction mixture extracted with EtOAc (30 mL × 3) and saturated $NaHCO_3$ aq. (30 mL) and the combined organic layers were dried over anhydrous Na_2SO_4. After filtration and evaporation, the residue was purified by silica gel column chromatography (EtOAc R_f = 0.42) to afford **2b** (86 mg, 0.31 mmol, 22% from 3 steps) as a white solid. ^{1}H NMR (300 MHz, MeOH-*d*4) δ 7.28-8.07 (m, 6 H, ArH), 5.57 (dd, *J*=9.9 Hz, 5.7 Hz, 1 H, H-1'), 4.37-4.41 (m, 1 H, H-3'), 4.05-4.10 (m, 1 H, H-4'), 4.01 (s, 3H, OCH_3), 4.82 (dd, *J*=5.4 Hz, *J*=2.7 Hz, 2H, H-5'), 2.49-2.56 (m, 1 H, H-2'), 1.86-1.96 (m, 1 H, H-2'). ^{13}C NMR (300 MHz, MeOH-*d*4) δ 155.1,134.1, 129.0, 127.4, 126.2, 125.7, 124.5, 123.4, 104.7, 87.4, 75.7, 73.1, 62.9, 54.5, 48.6, 42.2. HRMS calcd for $C_{16}H_{18}O_4$ (MNa^+) 297.1097, found 297.1104.

Compound 3a. A mixture of palladium acetate (16 mg, 0.071 mmol) and triphenylarsine (43.6 mg, 0.14 mmol) in DMF (3 mL) was stirred under argon atmosphere at room temperature for 20 min. Then **2a** (190 mg, 0.71 mmol), 1,4-anhydro-3,5-bis-O-(tert-butyldimethylsilyl)-2-deoxy-D-erythro-pent-1-enitol (205 mg, 0.594 mmol), and tri-n-butylamine (170mg, 0.92 mmol) in DMF (5 mL) were added, and the resulting reaction mixture was stirred under nitrogen at 70 °C for 18 h. The mixture was cooled to 0 °C, and 1 M tetra-n-butylammonium fluoride (3.16mL, 3.16 mmol) in THF was added and the mixture stirred for 1 h. The reaction mixture was filtered through Celite and extracted with EtOAc (50 mL × 3) and saturated $NaHCO_3$ aq. (50 mL). The combined organic layers were dried over anhydrous Na_2SO_4. After filtration and evaporation, the residue was purified by silica gel column chromatography (5% $MeOH/CH_2Cl_2$, R_f = 0.42). The eluted product was dissolved in acetic acid (2.5 mL) and acetonitril (2.5 mL). The solution was cooled to 0 °C, and sodium triacetoxyhydride (240 mg, 1.13 mmol) was added followed by stirring for 1 h. The reaction mixture was extracted with EtOAc (50 mL × 3) and saturated $NaHCO_3$ aq. (50 mL) and the combined organic layers were dried over anhydrous Na_2SO_4. After filtration and evaporation, the residue was purified by silica gel column chromatography (EtOAc, R_f = 0.53) to afford **3a** (79 mg, 0.3 mmol, 52 % over 3 steps) as a white solid. ^{1}H NMR (300 MHz, MeOH-*d*4) δ 7.26 (d, *J*=10.5Hz, 1 H, ArH), 6.83 (d, *J*=6 Hz, 1 H ArH), 5.41 (dd, *J*=10.2 Hz, 5.7 Hz, 1 H, H-1'), 4.34 (dt, *J*=6 Hz, *J*=2.7Hz 1 H, H-3'), 3.98(td, 1 H, *J*=5.4 Hz, *J*=2.7 Hz 1 H, H-4'), 3.86(s, 3H, OCH_3), 3.74 (dd, *J*=5.4 Hz, *J*=1.2 Hz, 2H, H-5'), 2.34-2.41 (m,1H, H-2'), 2.30 (d, *J*=2.1 Hz, CH_3),1.75-1.85(m,1H, H-2'). ^{13}C NMR (300 MHz, MeOH-*d*4) δ 157.3, 130.1, 123.5, 112.9, 112.4, 87.4, 74.8, 73.2, 62.8, 55.2, 47.3, 42.2, 13.4.HRMS calcd for $C_{13}H_{17}FO_4$ (MNa^+) 279.1002, found 279.1003.

General DMTr-protection Procedure. To a solution of free nucleoside (~100 mg, 1eq.) and *N,N*-dimethylaminopyridine (0.2 eq.) in pyridine (1 mL), 4,4'- dimethoxytrityl chloride (1.3 eq.) and triethylamine (1.3 eq.) were added. The solution was stirred for 18 h at ambient temperature. After quenching the reaction by adding MeOH, the mixture was evaporated to dryness *in vacuo*. The residue was extracted with EtOAc (50 mL x 3) and water (50 mL), and the organic layer was collected, dried over anhydrous Na_2SO_4, filtered, and evaporated to dryness under reduced pressure. The crude product was purified by silica gel column chromatography (hexane:EtOAc = 3:1 to 2:1) to afford the DMTr-protected nucleoside as a white foam.

Compound 4a. ^{1}H NMR (300 MHz, CDCl3) δ 7.12-7.41 (m, 11H, Ar-H), 6.81-6.89 (m, 4 H, Ar-H), 6.61(bs, 1 H, OH), 5.32 (dd, *J*=9.9 Hz, 5.7 Hz, 1 H, H-1'), 4.29-4.33 (m, 1 H, H-3'), 4.10-4.18(m, 1 H, H-4'), 3.83(bs, 6H, Ar-OCH_3), 3.71 (bs, 3H, Ar-OCH_3), 3.35(dd, 1H, *J*=9.6 Hz, *J*=4.8 Hz, H-5'), 3.20(dd, 1H, *J*=9.6 Hz, *J*=5.4 Hz, H-5'), 2.28-2.35 (m,1H, H-2'), 2.21(s, 3H, CH3) 1.79-1.88(m,1H, H-2'). ^{13}C NMR (300 MHz, $CDCl_3$) δ 158.6, 155.1, 151.1, 145.2, 135.4, 130.2, 129.5,128.2, 127.5, 125.3, 112.3, 111.2, 87.4, 78.7, 76.9, 76.3, 73.1, 62.8, 55.8, 42.1, 14.5. HRMS calcd for $C_{34}H_{35}FO_6$ (MNa^+) 581.2310, found 581.2312.

Compound 3b. ^{1}H NMR (300 MHz, CDCl3) δ 7.09-7.64 (m, 11 H, Ar-H), 6.72-6.75 (m, 4 H, Ar-H), 5.44 (dd, *J*=9.0 Hz, 6.3 Hz, 1 H, H-1'), 4.30-4.34 (m, 1 H, H-3'), 3.97-4.07(m, 1 H, H-4'), 3.83(s, 6H, Ar-OCH_3), 3.68 (s, 3H, Ar-OCH_3), 3.36(dd, 1H, *J*=9.9 Hz, *J*=4.5Hz, H-5'), 3.26(dd, 1H, *J*=9.6 Hz, *J*=5.4 Hz, H-5'), 2.36-2.43 (m,1H, H-2'), 1.83-1.96(m,1H, H-2'). ^{13}C NMR (300 MHz, $CDCl_3$) δ 158.7, 155.0, 145.0, 136.2, 133.7, 132.7, 130.7, 128.3, 127.9, 126.9, 124.8, 123.7, 113.2, 104.9, 86.3, 85.6, 77.5, 75.5, 74.7, 64.5, 55.4, 42.4. HRMS calcd for $C_{37}H_{36}O_6$ (MNa^+) 599.2404, found 599.2402.

General Phosphoramidite Synthesis Procedure. To a solution of DMTr-protected nucleoside (100 mg, 1 eq.) and diisopropylethylamine (4 eq.) in dichloromethane (1mL) was added 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite (1.5 eq.) at 0 °C. After stirring for 30 min at ambient temperature, the reaction mixture was extracted with EtOAc (30 mL × 3) and saturated $NaHCO_3$ aq. (20 mL). The organic layers were collected, dried over anhydrous Na_2SO_4, filtered, and evaporated to dryness under reduced pressure. The crude product was purified by silica gel column chromatography (hexane:EtOAc = 3:1 to 2:1) to afford compound as a pale yellow foam.

Compound 5a. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.18-7.47 (m, 11 H, Ar-H), 6.80-6.84 (m, 4 H, Ar-H), 6.5-6.6 (m, 1 H, Ar-H), 5.32(ddd, *J*=7.8 Hz, 5.7 Hz, 3.6Hz, 1 H, H-1'), 4.36-4.49 (m, 1 H, H-3'), 4.17-4.20(m, 1 H, H-4'), 3.76(t, 6H, Ar-OCH_3), 3.74 (bs, 3H, Ar-OCH_3), 3.64-3.83(m, 2H, CH2), 3.50-3.62(m, 2H, CH2), 3.22-3.27 (m, 2H, H-5'), 2.58-2.78 (m, 2H, CH), 2.41-2.52 (m,1H, H-2'), 2.23(s, 3H, CH3), 1.73-1.82(m,1H, H-2'), 1.04-1.27 (m, 12 H, N-CH3). ^{31}P NMR (400 MHz, $CDCl_3$): δ 148.5, 148.1. HRMS calcd for $C_{43}H_{52}FN_2O_7P(MH^+)$ 759.3569, found 759.3559.

Compound 4b. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.10-8.08 (m, 11 H, Ar-H), 6.80-6.83 (m, 4 H, Ar-H), 5.51-5.56 (m, 1 H, H-1'), 4.40-4.57 (m, 1 H, H-3'), 4.20-4.28(m, 1 H, H-4'), 3.93(d, 3H, Ar-OCH_3), 3.77 (bs, 6H, Ar-OCH_3), 3.65-3.73(m, 2H, CH2), 3.54-3.66(m, 2H, CH2), 3.30-3.43 (m, 2H, H-5'), 2.61-2.67(m, 2H, CH), 2.47-2.51(m, 1H, H-2'), 1.99(s, 3H, CH3), 1.91-1.98(m,1H, H-2'), 1.06-1.31 (m, 12 H, N-CH3). ^{31}P NMR (400MHz, $CDCl_3$): δ 148.7, 148.1. HRMS calcd for $C_{46}H_{53}N_2O_7P$ (MH^+) 777.3663, found 777.3659.

General Procedure for Triphosphate Synthesis. Proton sponge (1.5 equiv) and the free nucleoside (1 equiv) were dissolved in trimethyl phosphate (0.3 M) and cooled to -20 ºC. $POCl_3$ (1.5 equiv) was added dropwise, and the purple slurry was stirred at -20 ºC for 2 h. Tributylamine (6.2 equiv) was added, followed by a solution of tributylammonium pyrophosphate (5.0 equiv) in DMF (0.5 M). After 5 min, the reaction was quenched by addition of 0.5 M aqueous $Et_3NH_2CO_3$ (20 vol-equiv) and the resulting solution was lyophilized. Purification by reverse-phase (C18) HPLC (4-35% CH_3CN in 0.1 M $Et_3NH_2CO_3$, pH 7.5) followed by lyophilization afforded the triphosphate as a white solid.

Compound 6a. ^{31}P NMR (400 MHz, D_2O): δ -9.00 – -9.11(m; γ-P), -10.64 (d, *J* = 37.2 Hz; α-P), -22.5 (t, *J* = 37.5 Hz; β-P).

Compound 5b. ^{31}P NMR (400 MHz, D_2O): δ -10.02 – -10.83 (m; γ-P, α-P), -22.65 - -22.98 (m; β-P).

Table S1. Kinetic data for insertion of d**Y**TP d**5SICS** in sequence context I[a]

```
5'-d(TAATACGACTCACTATAGGGAGA)
3'-d(ATTATGCTGAGTGATATCCCTCTXGCTAGGTTACGGCAGGATCGC)
```

X(Template)	d**Y**TP	k_{cat} (min^{-1})	K_M (μM)	k_{cat}/K_M (M^{-1} min^{-1})
5SICS	**MMO2**	12 ± 1	33 ± 6	3.6×10^5
	5FM	8.2 ± 1.8	2.3 ± 0.3	3.6×10^6
	NaM	25 ± 5	4.9 ± 1.2	5.0×10^6
	5SICS	1.7 ± 0.5	63 ± 4	2.7×10^4
	A	1.2 ± 0.3	54 ± 6	2.2×10^4
	C	nd[b]	nd[b]	$<1.0 \times 10^3$
	G	4.6 ± 0.6	36 ± 4	1.3×10^5
	T	1.7 ± 0.2	130 ± 19	1.3×10^4

[a] See Experimental Section for experimental details. [b] Rates too slow to independently determine k_{cat} and K_M.

Table S2. Kinetic data for insertion of d**Y**TP opposite d**MMO2**, d**5FM**, or d**NaM** in sequence context I[a]

5'-d(TAATACGACTCACTATAGGGAGA)
3'-d(ATTATGCTGAGTGATATCCCTCT**X**GCTAGGTTACGGCAGGATCGC)

X(Template)	dYTP	k_{cat} (min^{-1})	K_M (μM)	k_{cat}/K_M (M^{-1} min^{-1})
MMO2	**5SICS**	8.5 ± 0.1	0.18 ± 0.01	4.7×10^7
	MMO2	5.1 ± 0.7	44 ± 4	1.2×10^5
	A	3.3 ± 0.3	32 ± 4	1.0×10^5
	G	nd[b]	nd[b]	$<1.0 \times 10^3$
	C	nd[b]	nd[b]	$<1.0 \times 10^3$
	T	nd[b]	nd[b]	$<1.0 \times 10^3$
5FM	**5SICS**	3.3 ± 1.0	0.23 ± 0.03	1.4×10^7
	FM	10.8 ± 2.4	12.2 ± 3.9	8.9×10^5
	A	6.1 ± 1.0	20 ± 2	3.0×10^5
	G	nd[b]	nd[b]	$<1.0 \times 10^3$
	C	nd[b]	nd[b]	$<1.0 \times 10^3$
	T	0.31 ± 0.1	131 ± 15	2.4×10^3
NaM	**5SICS**	3.7 ± 0.8	0.10 ± 0.03	3.7×10^7
	NaM	6.9 ± 1.6	2.0 ± 0.4	3.4×10^6
	A	8.1 ± 1.7	11.4 ± 4.3	7.1×10^5
	G	nd[b]	nd[b]	$<1.0 \times 10^3$
	C	nd[b]	nd[b]	$<1.0 \times 10^3$
	T	0.77 ± 0.12	136 ± 4.3	5.6×10^3

[a] See Experimental Section for experimental details. [b] Rates too slow to independently determine k_{cat} and K_M.

Table S3. Kinetic data for extension of dN:d**5SICS** by dCTP insertion in sequence context I[a]

```
5'-d(TAATACGACTCACTATAGGGAGAY)
3'-d(ATTATGCTGAGTGATATCCCTCTXGCTAGGTTACGGCAGGATCGC)
```

X(Template)	**Y**	k_{cat} (min^{-1})	K_M (μM)	k_{cat}/K_M (M^{-1} min^{-1})
5SICS	**MMO2**	6.4 ± 1.1	3.4 ± 0.3	1.9×10^6
	5FM	3.49 ± 0.11	0.63 ± 0.10	5.5×10^6
	NaM	3.1 ± 0.9	2.47 ± 0.40	1.2×10^6
	5SICS	nd[b]	nd[b]	$<1.0 \times 10^3$
	A	0.79 ± 0.01	76 ± 11	1.0×10^4
	G	0.62 ± 0.02	127 ± 8	4.9×10^3
	C	0.18 ± 0.01	44 ± 9	4.2 ± 10^3
	T	2.0 ± 0.1	4.9 ± 0.3	4.0 ± 10^5

[a] See Experimental Section for experimental details. [b] Rates too slow to independently determine k_{cat} and K_M.

Table S4. Kinetic data for extension of dN:d**MMO2**, dN:d**5FM**, and dN:d**NaM** by dCTP insertion in sequence context I[a]

```
5'-d(TAATACGACTCACTATAGGGAGAY)
3'-d(ATTATGCTGAGTGATATCCCTCTXGCTAGGTTACGGCAGGATCGC)
```

X(Template)	**Y**	k_{cat} (min^{-1})	K_M (μM)	k_{cat}/K_M (M^{-1} min^{-1})
MMO2	**5SICS**	3.8 ± 0.3	5.7 ± 1.2	6.7×10^5
	MMO2	0.87 ± 0.28	165 ± 47	5.3×10^3
	A	8.7 ± 0.9	187 ± 17	4.6×10^4
	G	nd[b]	nd[b]	$<1.0 \times 10^3$
	C	3.90 ± 1.32	3.17 ± 0.23	1.2×10^6
	T	6.23 ± 1.01	9.45 ± 1.23	6.6×10^5
FM	**5SICS**	3.2 ± 0.7	1.4 ± 0.3	2.3×10^6
	5FM	2.57 ± 0.39	99.5 ± 17.9	2.6×10^4
	A	5.49 ± 0.08	17.3 ± 1.6	3.2×10^5
	G	nd[b]	nd[b]	$<1.0 \times 10^3$
	C	2.4 ± 0.23	1.19 ± 0.10	2.0×10^6
	T	4.14 ± 0.19	2.06 ± 0.17	2.0×10^6
NaM	**5SICS**	5.73 ± 0.51	2.15 ± 0.50	2.7×10^6
	NaM	0.78 ± 0.21	69 ± 25	1.1×10^4
	A	3.99 ± 1.46	102 ± 35	3.9×10^4
	G	nd[b]	nd[b]	$<1.0 \times 10^3$
	C	7.29 ± 0.69	4.78 ± 0.24	1.5×10^6
	T	4.73 ± 0.42	8.33 ± 0.27	5.7×10^5

[a] See Experimental Section for experimental details. [b] Rates too slow to independently determine k_{cat} and K_M.

Table S5. Kinetic data for insertion of d**Y**TP opposite d**5SICS** in sequence context II[a]

```
5'-d(TAATACGACTCACTATAGGGAGC)
3'-d(ATTATGCTGAGTGATATCCCTCGXTCTAGGTTACGGCAGGATCGC)
```

X(Template)	d**Y**TP	k_{cat} (min^{-1})	K_M (μM)	k_{cat}/K_M (M^{-1} min^{-1})
5SICS	**MMO2**	2.91 ± 0.76	6.04 ± 1.9	4.8×10^5
	5FM	9.52 ± 1.14	2.29 ± 0.21	4.2×10^6
	NaM	15.0 ± 1.10	0.37 ± 0.07	4.1×10^7
	5SICS	1.41 ± 0.04	5.84 ±0.20	2.4×10^5
	A	1.23 ± 0.24	20.8 ± 2.86	5.9×10^4
	C	nd[b]	nd[b]	$<1.0 \times 10^3$
	G	3.11 ± 0.53	15.9 ± 2.34	1.9×10^5
	T	0.62 ± 0.20	74.3 ± 42.4	8.5×10^3

[a] See Experimental Section for experimental details. [b] Rates too slow to independently determine k_{cat} and K_M.

Table S6. Kinetic data for insertion of d**Y**TP opposite d**MMO2**, d**5FM**, or d**NaM** in sequence context II[a]

```
5'-d(TAATACGACTCACTATAGGGAGC)
3'-d(ATTATGCTGAGTGATATCCCTCGXTCTAGGTTACGGCAGGATCGC)
```

X(Template)	d**Y**TP	k_{cat} (min^{-1})	K_M (μM)	k_{cat}/K_M (M^{-1} min^{-1})
MMO2	**5SICS**	10.7 ± 2.4	0.16 ± 0.04	6.6×10^7
	MMO2	15.2 ± 3.4	8.34 ± 3.8	1.8×10^6
	A	17.4 ± 4.1	10.2 ±3.3	1.7×10^6
	G	0.16 ± 0.001	20.3 ± 5.0	7.9×10^3
	C	0.65 ± 0.04	219 ± 46	3.0×10^3
	T	1.21 ± 0.06	235 ± 123	5.2×10^3
5FM	**5SICS**	12.3 ± 2.6	0.15 ± 0.04	8.0×10^7
	5FM	5.37 ± 0.33	1.98 ± 0.39	2.7×10^6
	A	22.6 ± 4.1	7.4 ± 1.5	3.1×10^6
	G	016 ± 0.02	135 ± 52	1.2×10^3
	C	1.15 ± 0.01	152 ± 88	7.5×10^3
	T	2.95 ± 0.64	253 ± 158	1.2×10^4
NaM	**5SICS**	12.4 ± 3.6	0.22 ± 0.04	5.5×10^7
	NaM	5.36 ± 0.99	1.11 ± 0.08	4.8×10^6
	A	21.9 ± 4.2	15.3 ± 1.2	1.4×10^6
	G	0.1± 0.003	26.68 ± 3.2	3.7×10^3
	C	0.63 ± 0.19	98.3 ± 10.4	6.4×10^3
	T	1.18 ± 0.35	141 ± 17.9	8.3×10^3

[a] See Experimental Section for experimental details. [b] Rates too slow to independently determine k_{cat} and K_M.

Table S7. Kinetic data for extension of dN:d**5SICS** by dATP insertion in sequence context II[a]

```
5'-d(TAATACGACTCACTATAGGGAGCY)
3'-d(ATTATGCTGAGTGATATCCCTCGXTCTAGGTTACGGCAGGATCGC)
```

X(Template)	**Y**	k_{cat} (min^{-1})	K_M (μM)	k_{cat}/K_M (M^{-1} min^{-1})
5SICS	**MMO2**	4.33 ± 1.17	13.5 ± 3.79	3.2×10^5
	5FM	7.28 ± 1.41	5.82 ± 1.70	1.3×10^6
	NaM	2.15 ± 0.18	3.15 ± 1.29	6.8×10^5
	5SICS	nd[b]	nd[b]	$<1.0 \times 10^3$
	A	0.10 ± 0.03	29.4 ± 15.9	3.6×10^3
	G	0.212 ± 0.04	324 ± 35	6.5×10^2
	C	nd[b]	nd[b]	$<1.0 \times 10^3$
	T	4.99 ± 0.32	12.0 ± 0.3	3.3×10^5

[a] See Experimental Section for experimental details. [b] Rates too slow to independently determine k_{cat} and K_M.

Table S8. Kinetic data for extension of dN:d**MMO2**, dN:d**5FM**, and dN:d**NaM** by dATP insertion in sequence context II[a]

```
5'-d(TAATACGACTCACTATAGGGAGCY)
3'-d(ATTATGCTGAGTGATATCCCTCGXTCTAGGTTACGGCAGGATCGC)
```

X(Template)	**Y**	k_{cat} (min^{-1})	K_M (μM)	k_{cat}/K_M (M^{-1} min^{-1})
MMO2	**5SICS**	12.4 ± 1.44	7.30 ± 2.39	1.7×10^6
	MMO2	nd[b]	nd[b]	$<1.0 \times 10^3$
	A	0.42 ± 0.12	38.4 ± 12.5	1.1×10^4
	G	nd[b]	nd[b]	$<1.0 \times 10^3$
	C	4.91 ± 0.56	11.2 ± 2.41	4.4×10^5
	T	6.67 ± 0.99	3.37 ± 0.51	2.0×10^6
5FM	**5SICS**	13.2 ± 0.31	1.63 ± 0.02	8.2×10^6
	5FM	nd[b]	nd[b]	$<1.0 \times 10^3$
	A	0.59 ± 0.06	13.4 ± 4.0	4.4×10^4
	G	nd[b]	nd[b]	$<1.0 \times 10^3$
	C	7.17 ± 0.60	4.76 ± 0.77	1.5×10^6
	T	6.97 ± 1.46	1.72 ± 0.23	4.1×10^6
NaM	**5SICS**	17.3 ± 5.9	4.16 ±1.40	4.1×10^6
	NaM	nd[b]	nd[b]	$<1.0 \times 10^3$
	A	0.34 ± 0.17	23.8 ± 12.7	1.4×10^4
	G	nd[b]	nd[b]	$<1.0 \times 10^3$
	C	6.29 ± 0.83	21.0 ± 6.1	3.0×10^5
	T	5.66 ± 0.28	6.85 ± 0.80	8.3×10^5

[a] See Experimental Section for experimental details. [b] Rates too slow to independently determine k_{cat} and K_M.

Table S9. Kinetic data for insertion of d**NaM** opposite d**5SICS** in sequence context III[a]

```
5'-d(TAATACGACTCACTATAGGGAGT)
3'-d(ATTATGCTGAGTGATATCCCTCAXTCTAGGTTACGGCAGGATCGC)
```

X(Template)	d**Y**TP	k_{cat} (min^{-1})	K_M (μM)	k_{cat}/K_M (M^{-1} min^{-1})
5SICS	**NaM**	20.4 ± 4.41	3.44 ± 1.37	5.9×10^6

[a] See Experimental Section for experimental details.

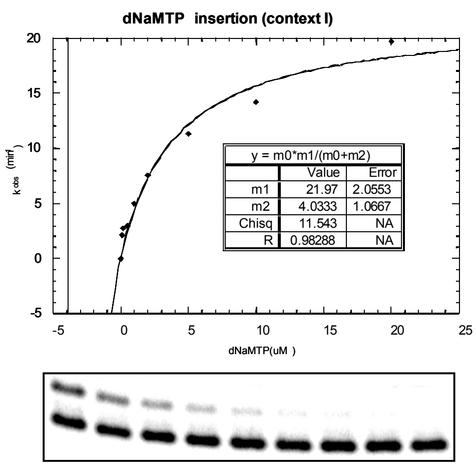
Table S10 Kinetic data for insertion of d**NaM** opposite d**5SICS** in sequence context IV[a]

```
5'-d(TAATACGACTCACTATAGGGAGG)
3'-d(ATTATGCTGAGTGATATCCCTCCXTCTAGGTTACGGCAGGATCGC)
```

X(Template)	d**Y**TP	k_{cat} (min^{-1})	K_M (μM)	k_{cat}/K_M (M^{-1} min^{-1})
5SICS	**NaM**	29.0 ± 7.7	0.45 ± 0.06	6.4×10^7

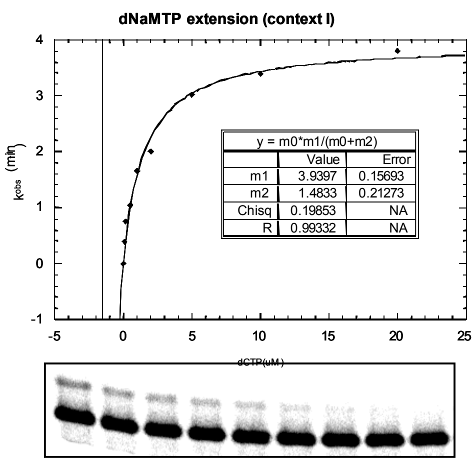
[a] See Experimental Section for experimental details.

1)
5'-d(TAATACGACTCACTATAGGGAGA) ← **dNaMTP**
3'-d(ATTATGCTGAGTGATATCCCTCT(**d5SICS**)GCTAGGTTACGGCAGGATCGC)



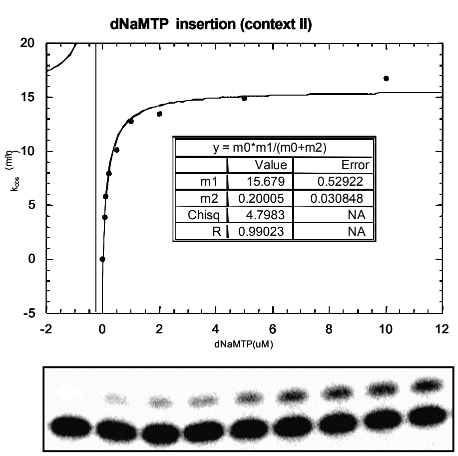
dNaMTP [μM] 20 10 5 2 1 0.5 0.2 0.1 0.05

2)
5'-d(TAATACGACTCACTATAGGGAG**dNaM**) ← **dCTP**
3'-d(ATTATGCTGAGTGATATCCCTCT(d5SICS)GCTAGGTTACGGCAGGATCGC)



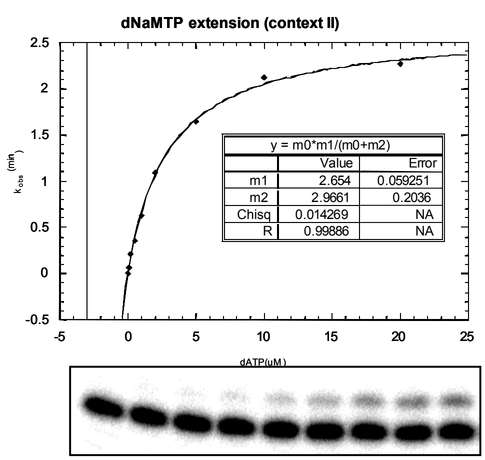
dCTP [μM] 20 10 5 2 1 0.5 0.2 0.1 0.05

3)
5'-d(TAATACGACTCACTATAGGGAGC) ← **dNaMTP**
3'-d(ATTATGCTGAGTGATATCCCTCG(**d5SICS**)TCTAGGTTACGGCAGGATCGC)



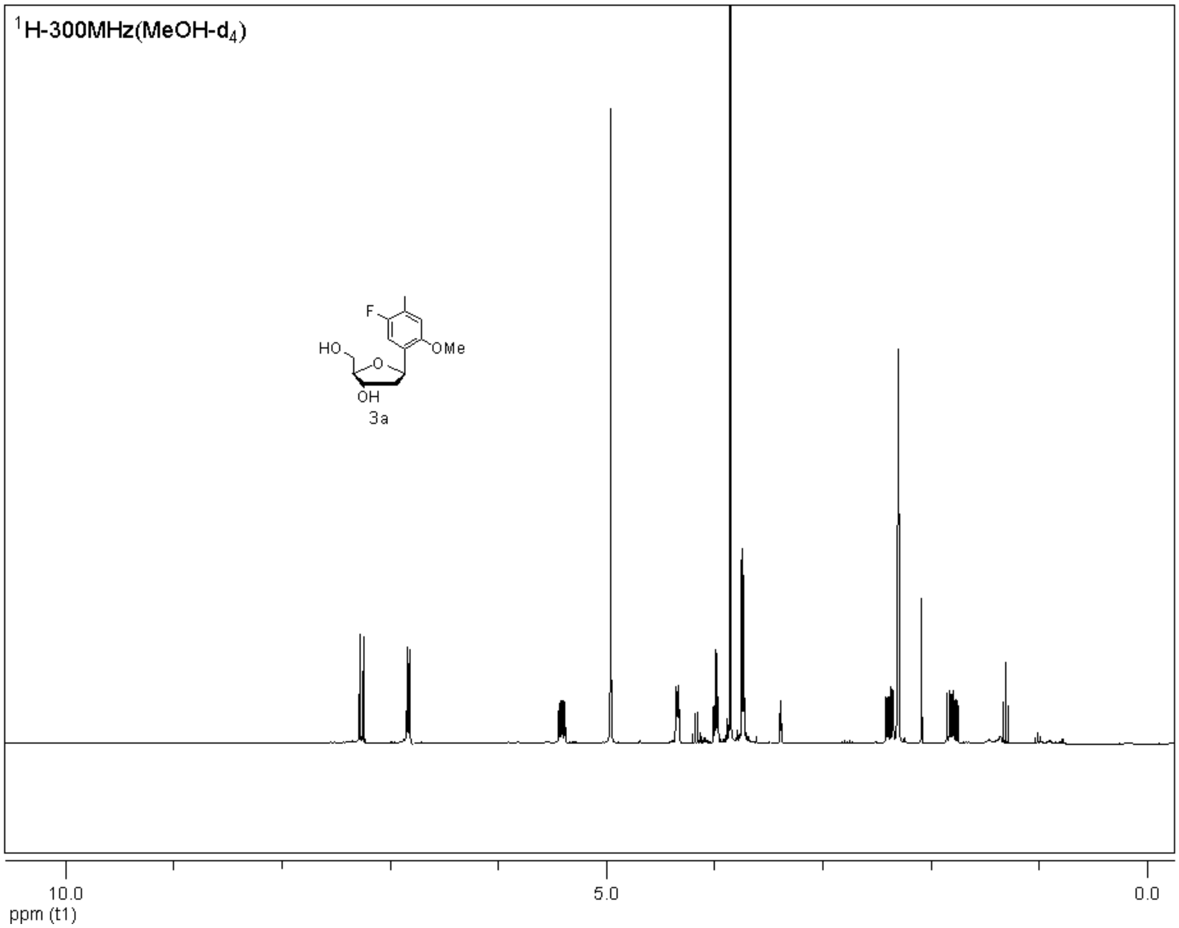
dNaMTP [μM] 0.02 0.05 0.1 0.2 0.5 1 2 5 10

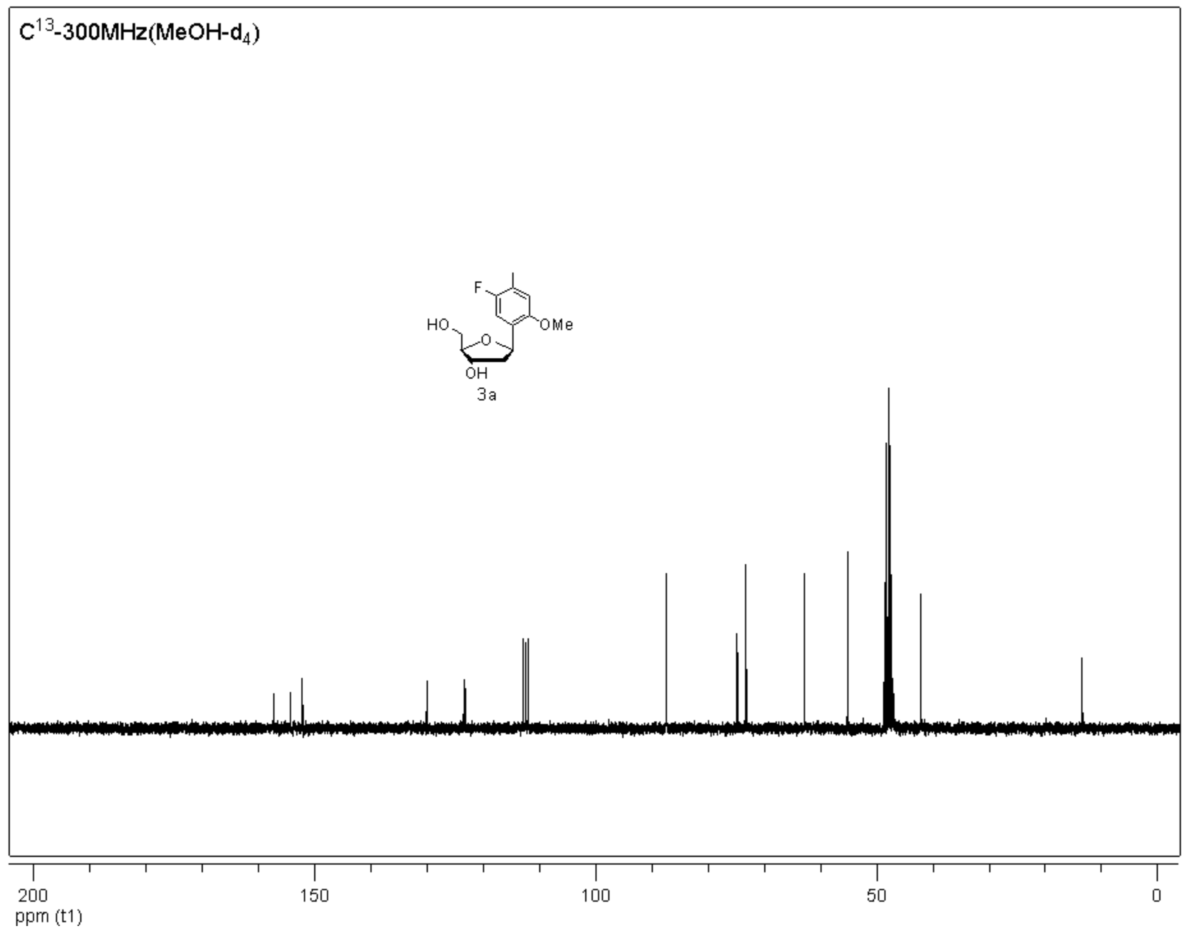
4)
5'-d(TAATACGACTCACTATAGGGAG**dNaM**) ← **dATP**
3'-d(ATTATGCTGAGTGATATCCCTCG(d5SICS)TCTAGGTTACGGCAGGATCGC)



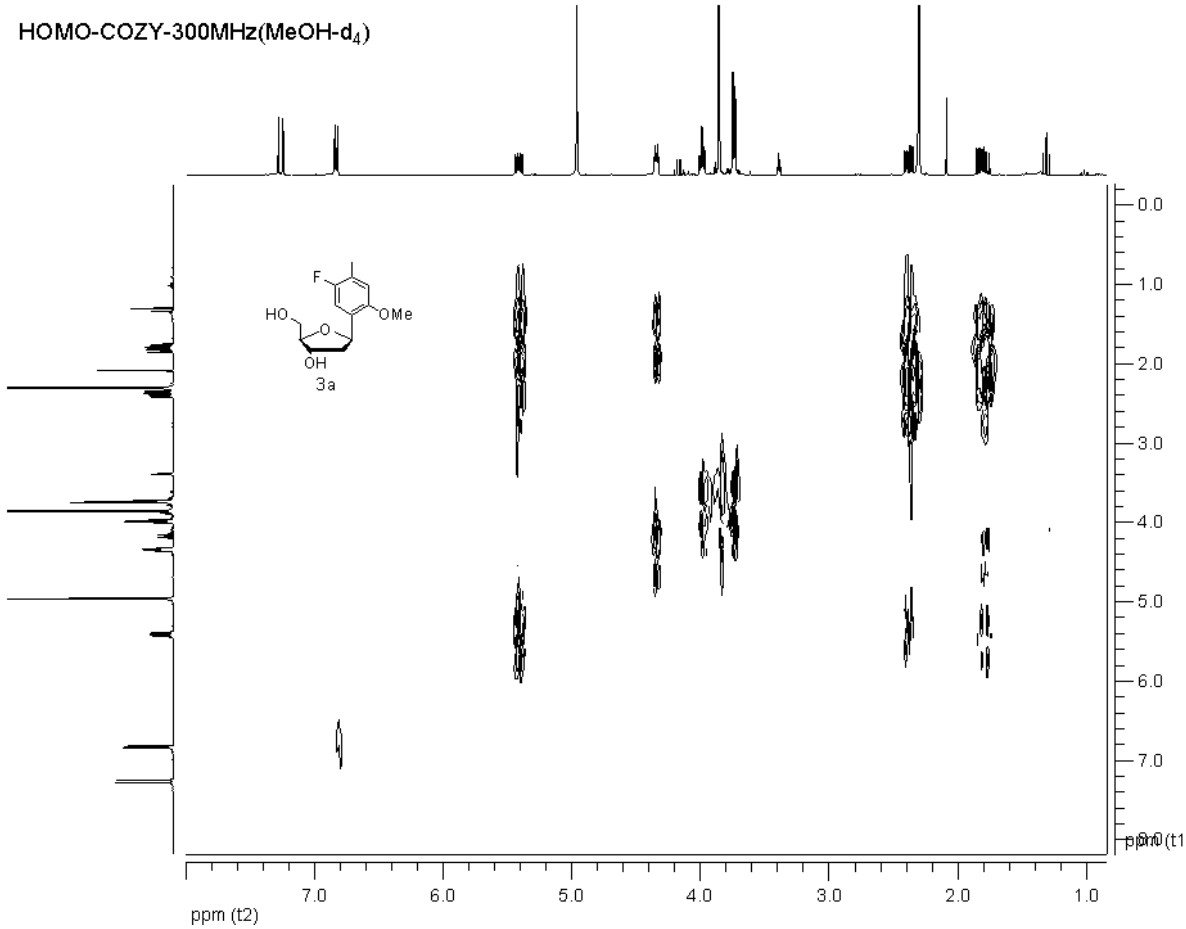
dATP [μM] 0.05 0.1 0.2 0.5 1 2 5 10 20

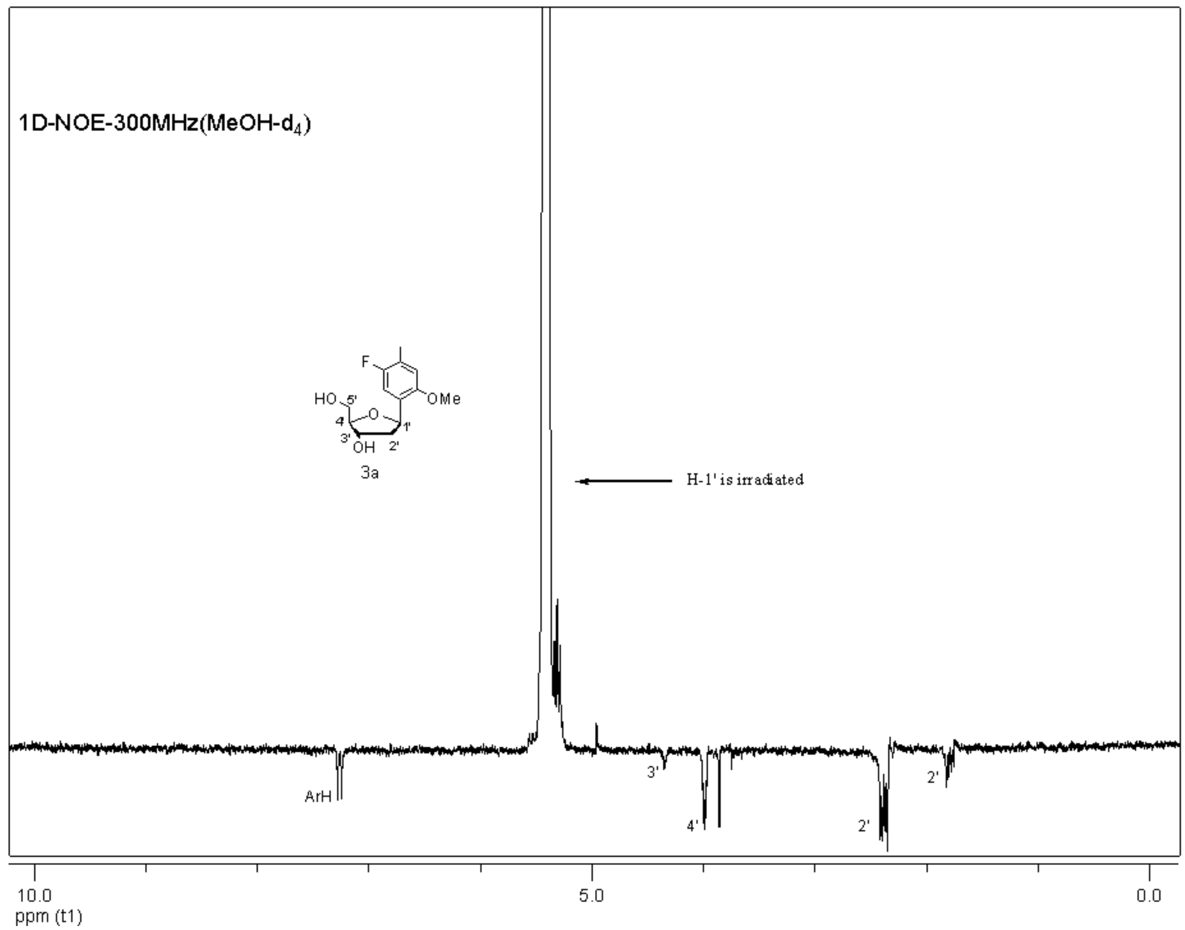
Figure S1. Raw gel data and plots of representative data (d**NaMTP** insertion opposite d**5SICS** and extension in sequence context I and II)

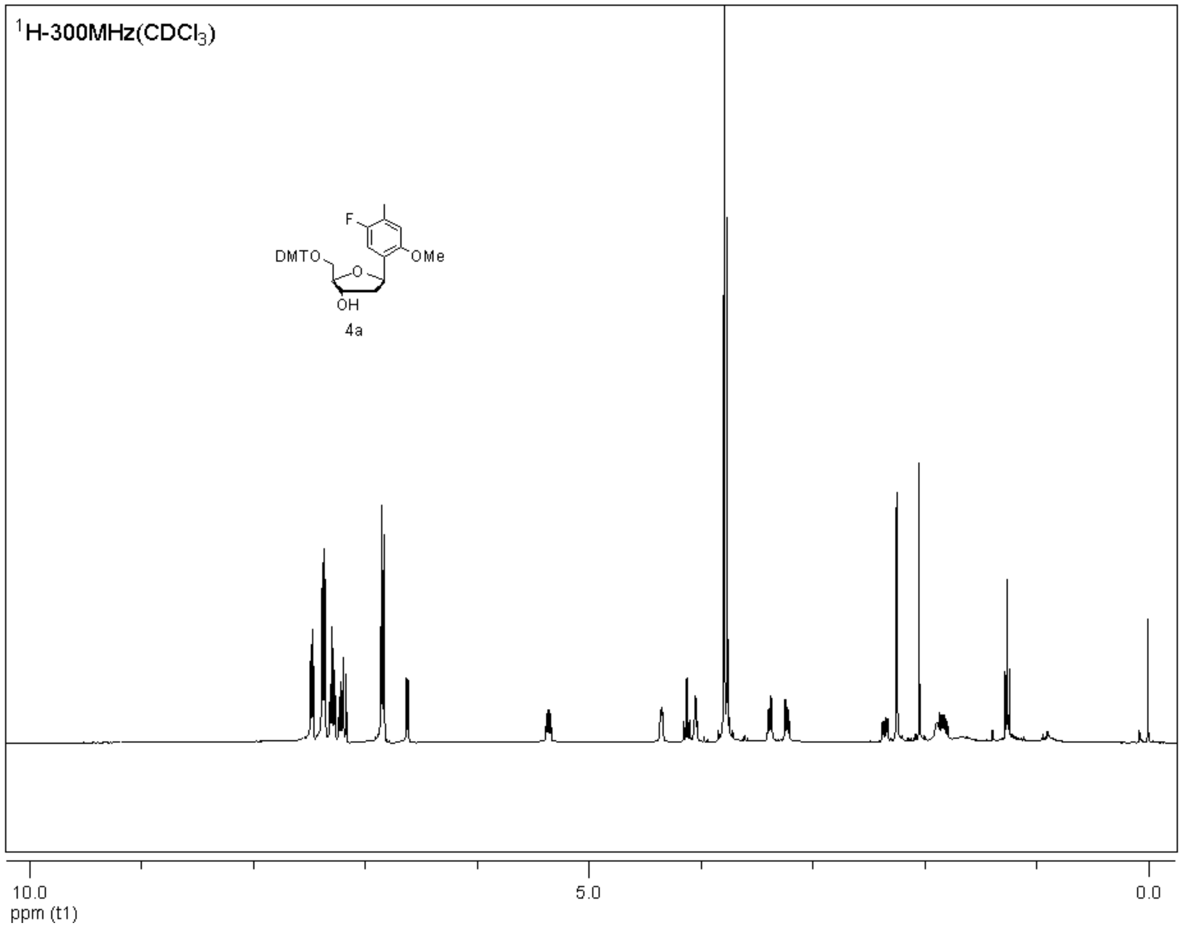

1H-300MHz(MeOH-d4)
F
HO
O
OMe
OH
3a
10.0
5.0
0.0
ppm (t1)


C13-300MHz(MeOH-d4)
F
HO
O
OMe
OH
3a
200
150
100
50
0
ppm (t1)

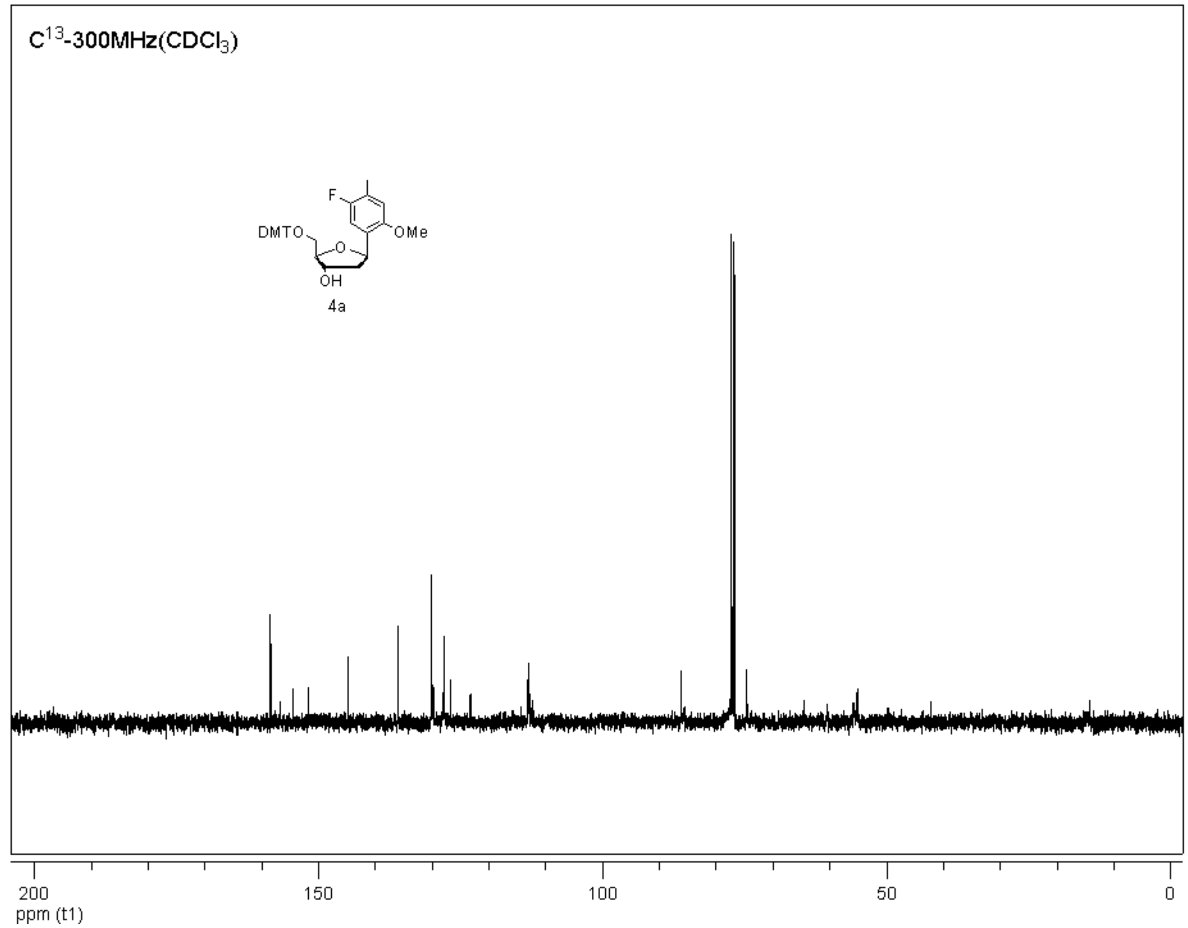
HOMO-COZY-300MHz(MeOH-d_4)



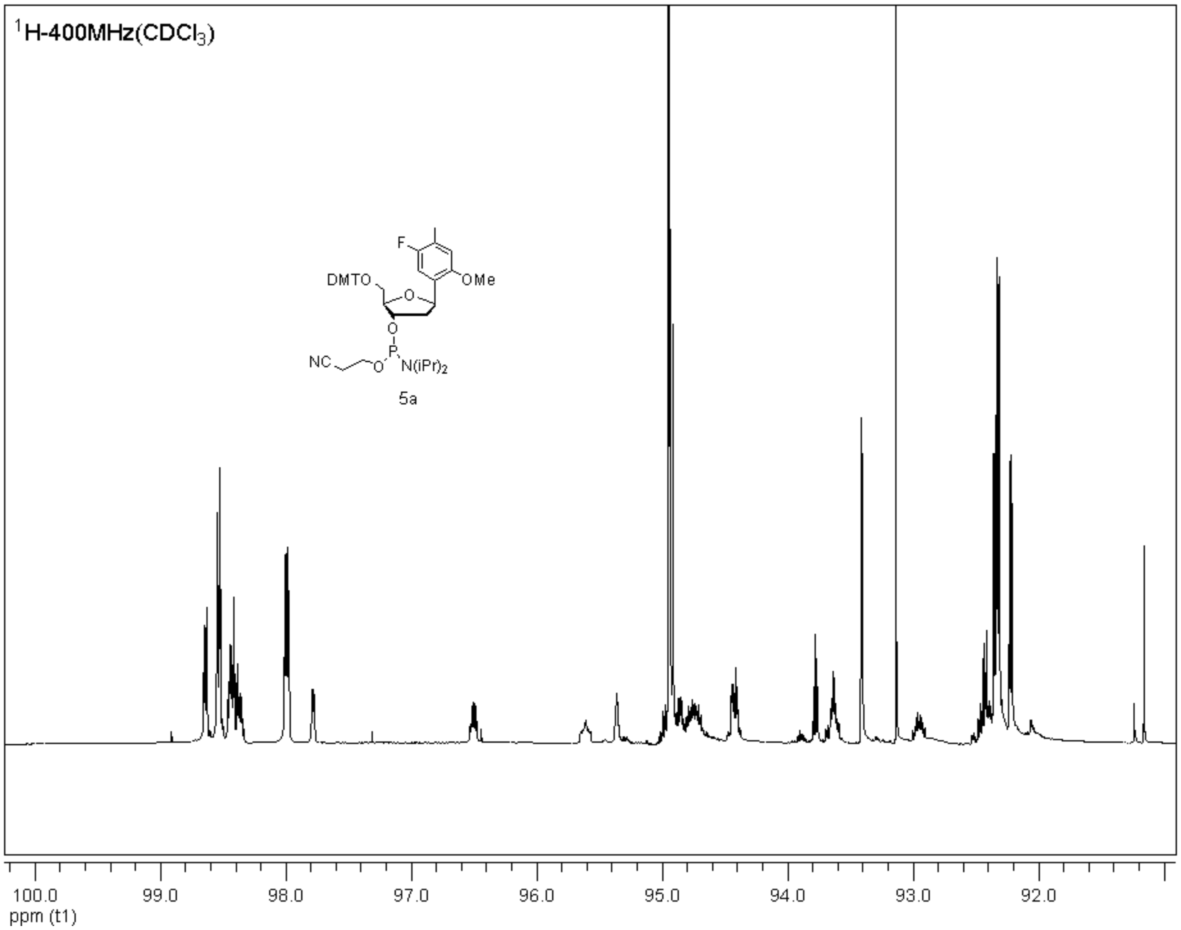

1D-NOE-300MHz(MeOH-d4)
F
HO
5'
4'
O
1'
OMe
3'
OH
2'
3a
H-1' is irradiated
ArH
3'
4'
2'
2'
10.0
ppm (t1)
5.0
0.0



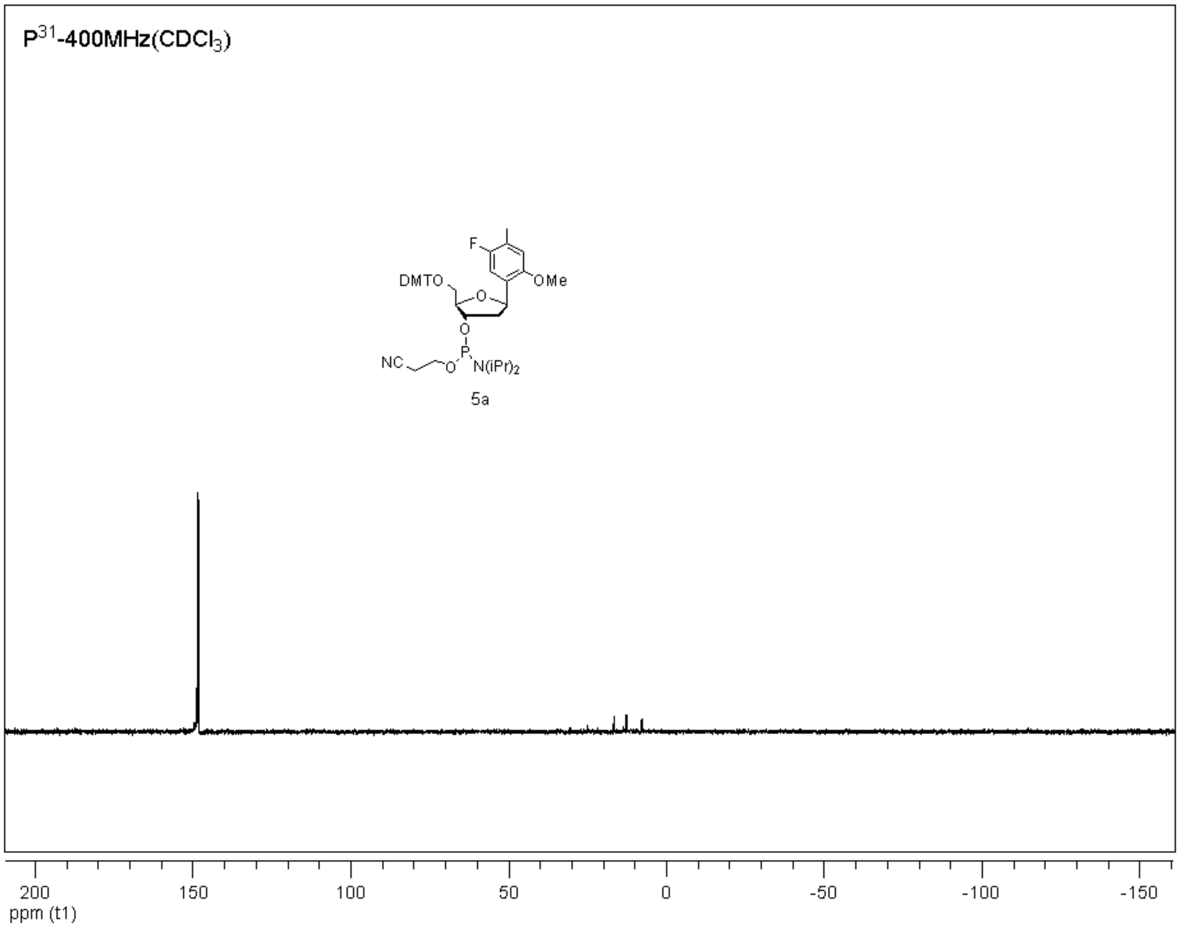
1H-300MHz(CDCl3)
F
DMTO
O
OMe
OH
4a
10.0
5.0
0.0
ppm (t1)



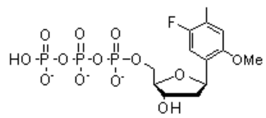
C13-300MHz(CDCl3)
DMTO
O
F
OMe
OH
4a
200
150
100
50
0
ppm (t1)



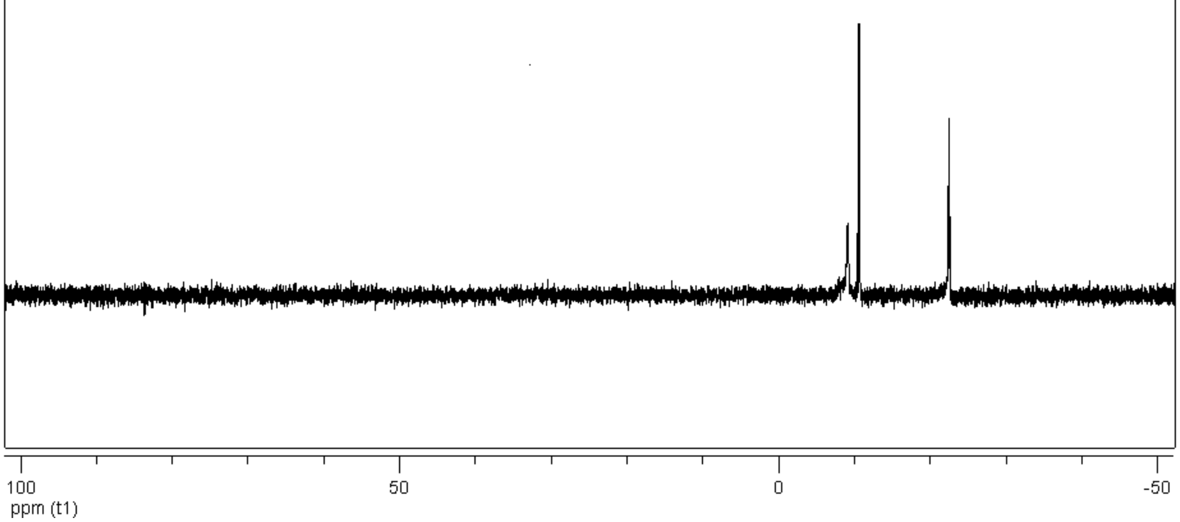
^{1}H-400MHz(CDCl$_3$)
DMTO
F
OMe
O
P
NC
N(iPr)$_2$
5a
100.0
99.0
98.0
97.0
96.0
95.0
94.0
93.0
92.0
ppm (t1)

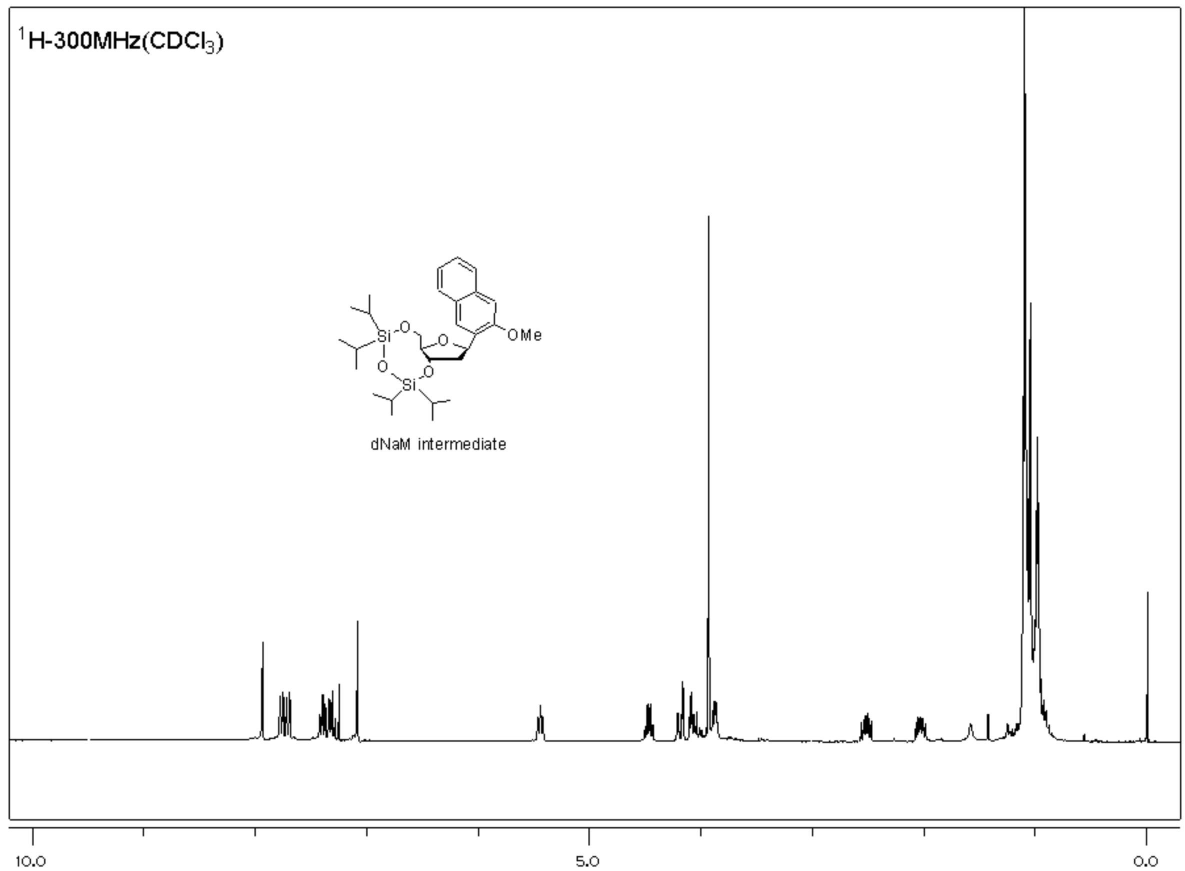

P31-400MHz(CDCl3)
DMTO
F
OMe
O
O
P
NC
O
N(iPr)2
5a
200
150
100
50
0
-50
-100
-150
ppm (t1)

P^{31}-400MHz(D_2O)

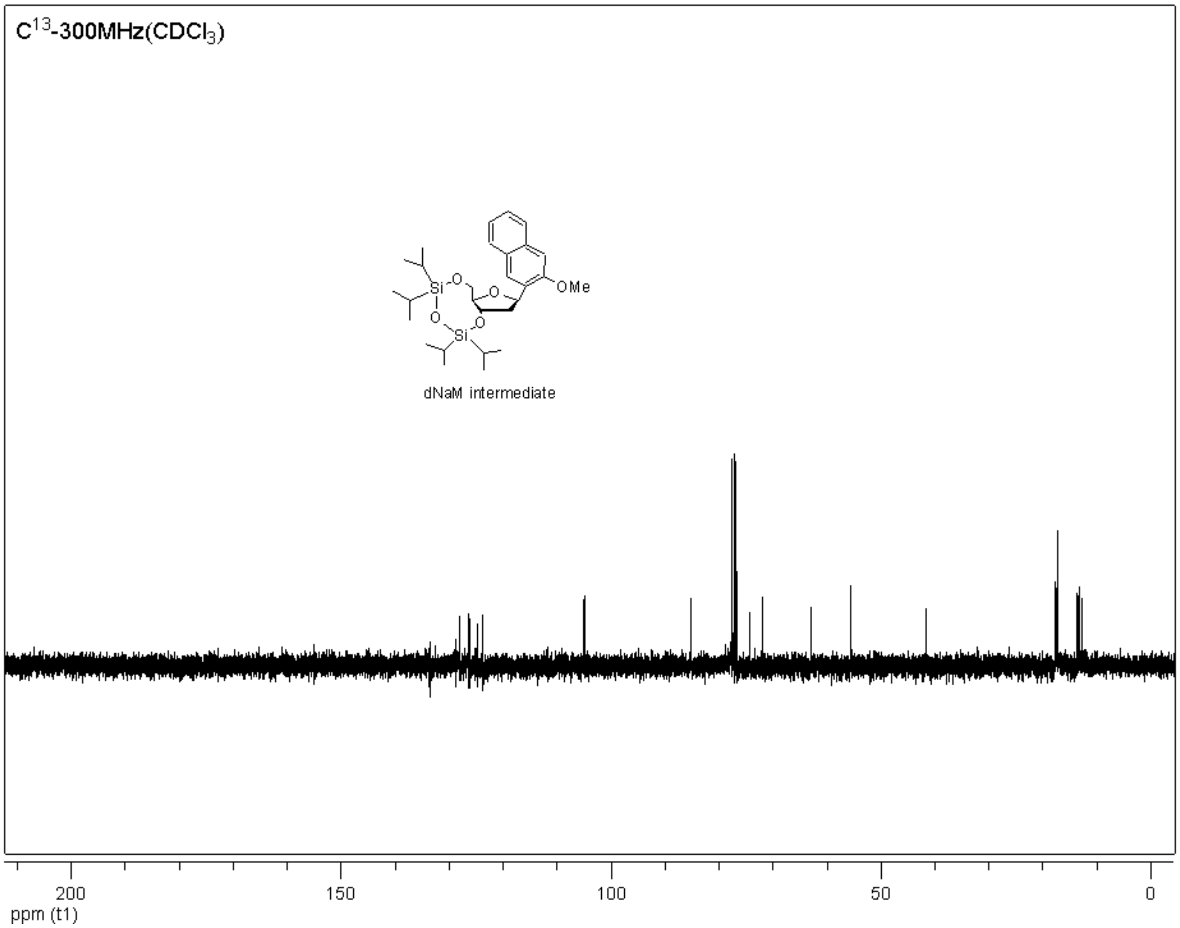


6a

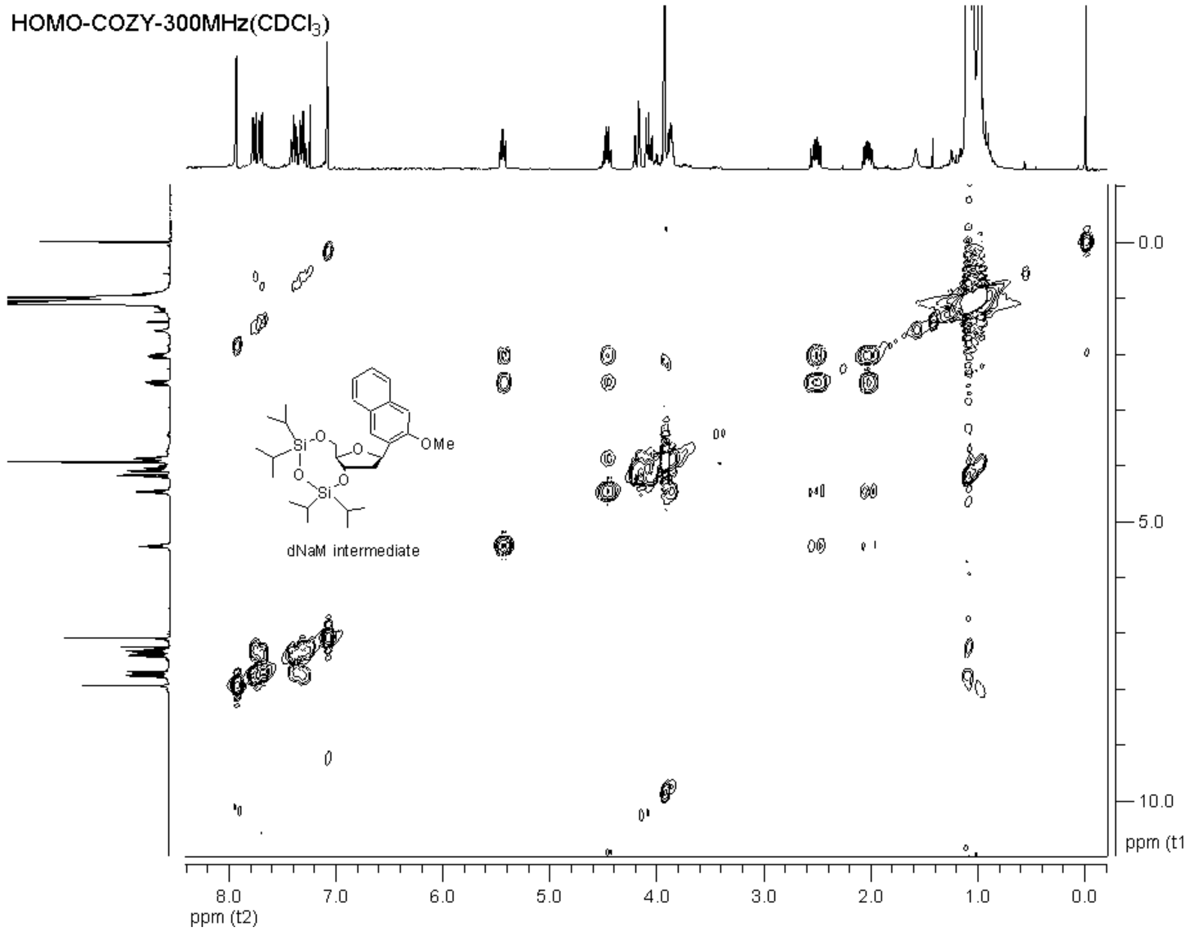


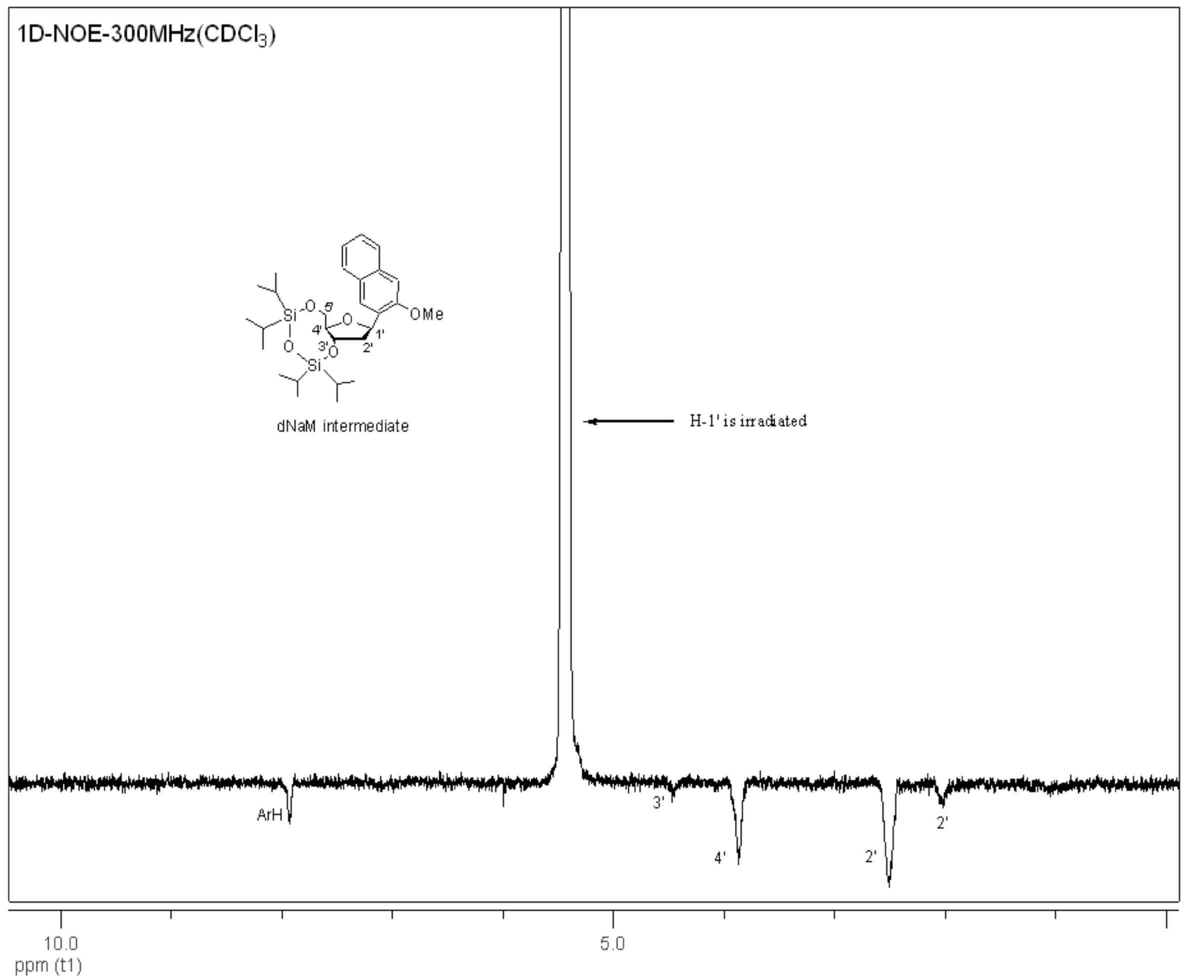


^{1}H-300MHz(CDCl$_3$)
OMe
Si
O
Si
O
O
O
dNaM intermediate
10.0
5.0
0.0

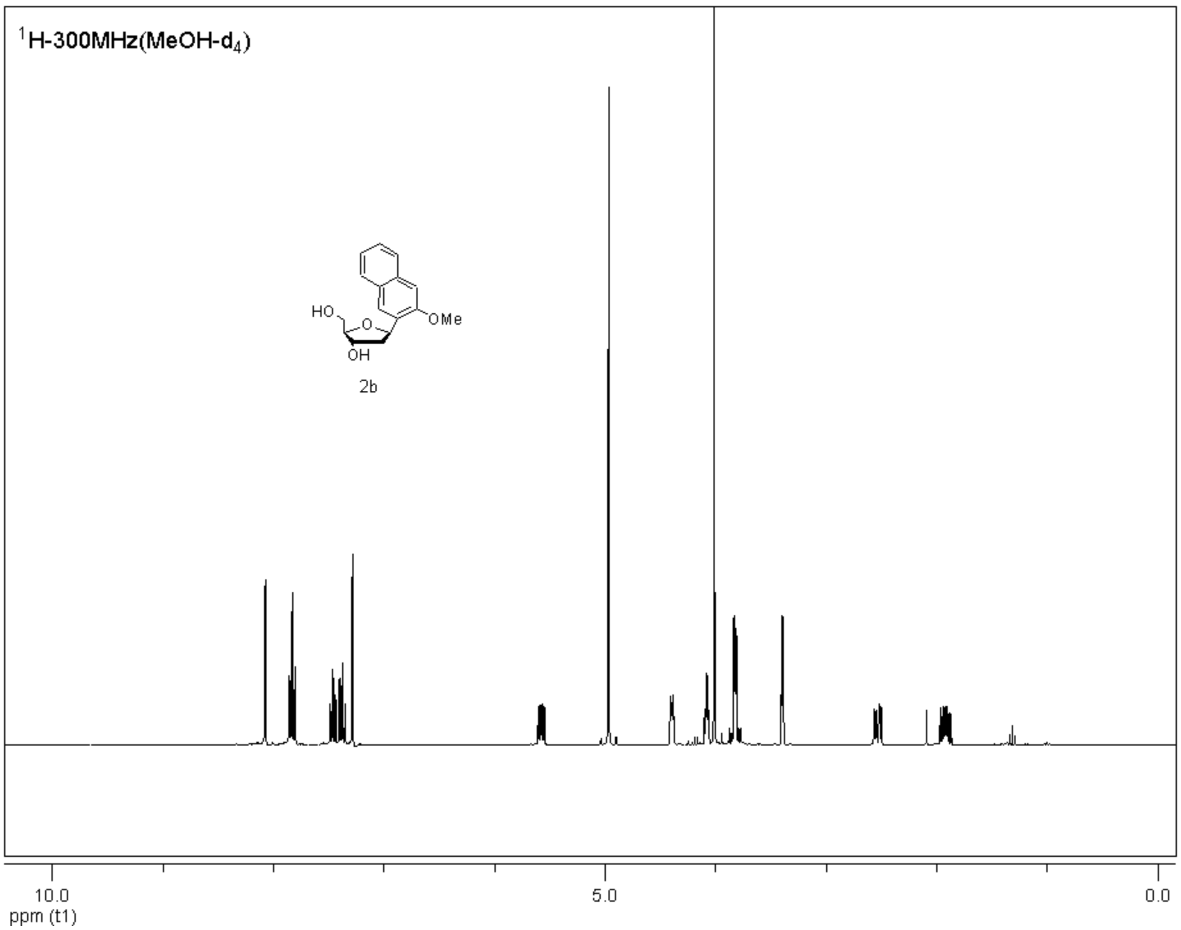

C13-300MHz(CDCl3)
OMe
dNaM intermediate
200
150
100
50
0
ppm (t1)

HOMO-COZY-300MHz($CDCl_3$)

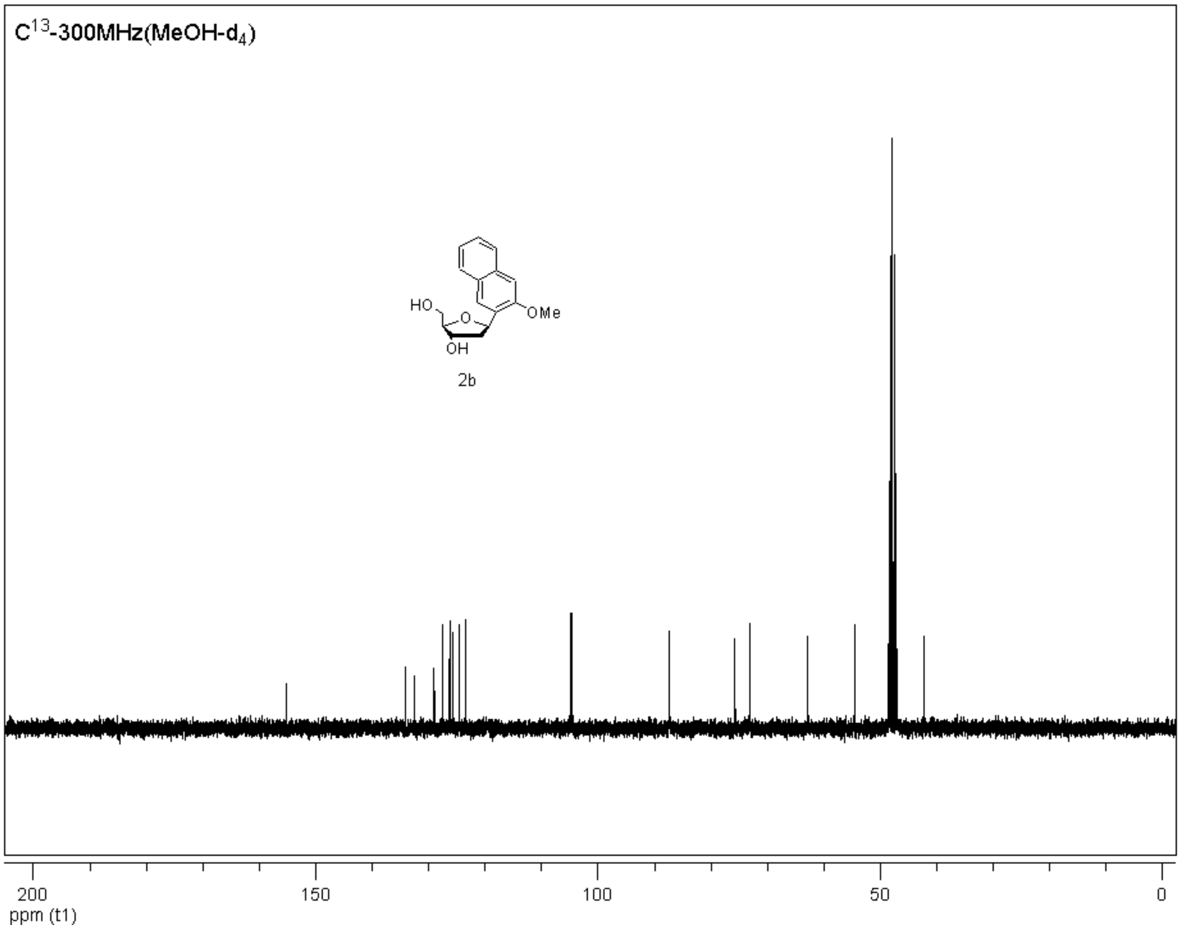




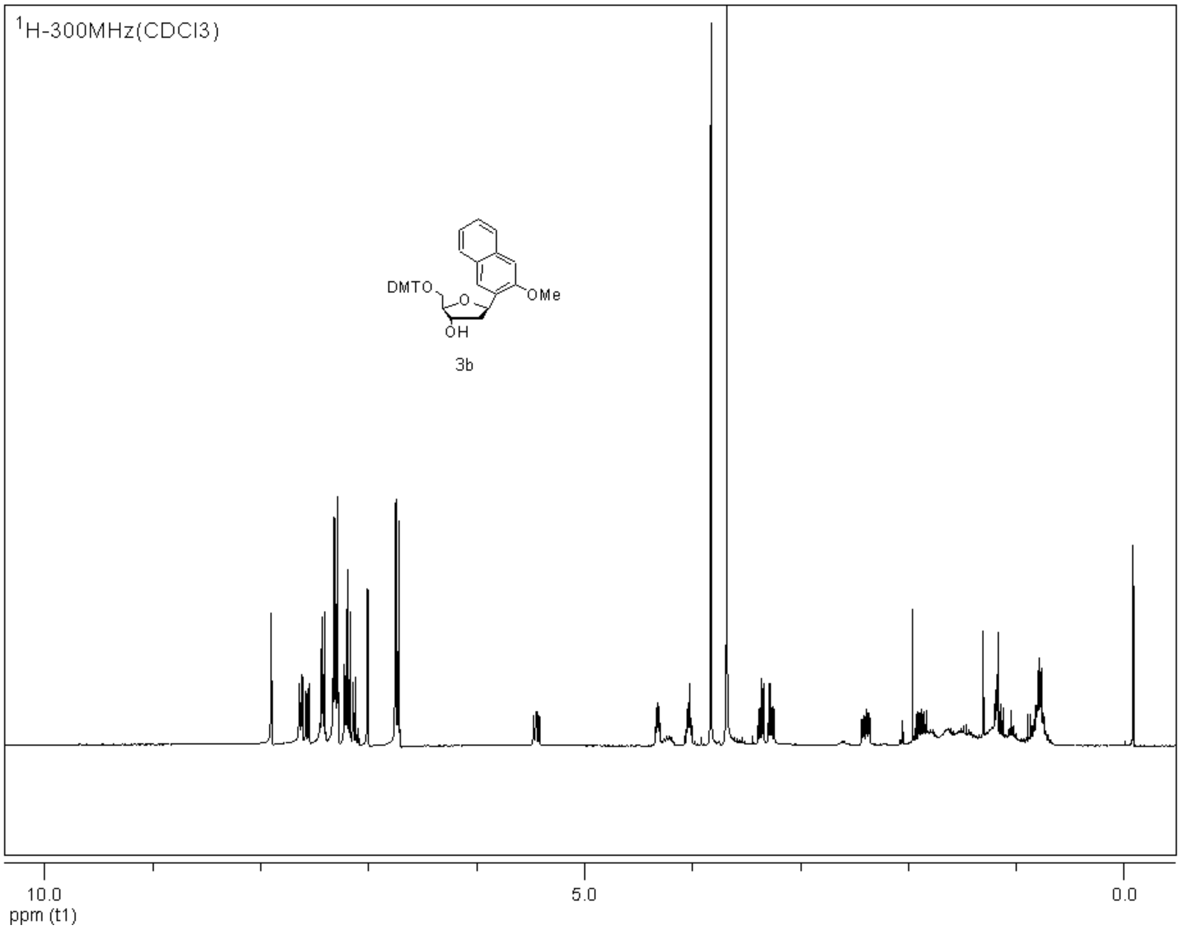
1D-NOE-300MHz(CDCl3)
OMe
5'
4'
3'
1'
2'
Si
O
dNaM intermediate
H-1' is irradiated
ArH
3'
4'
2'
2'
10.0
5.0
ppm (t1)



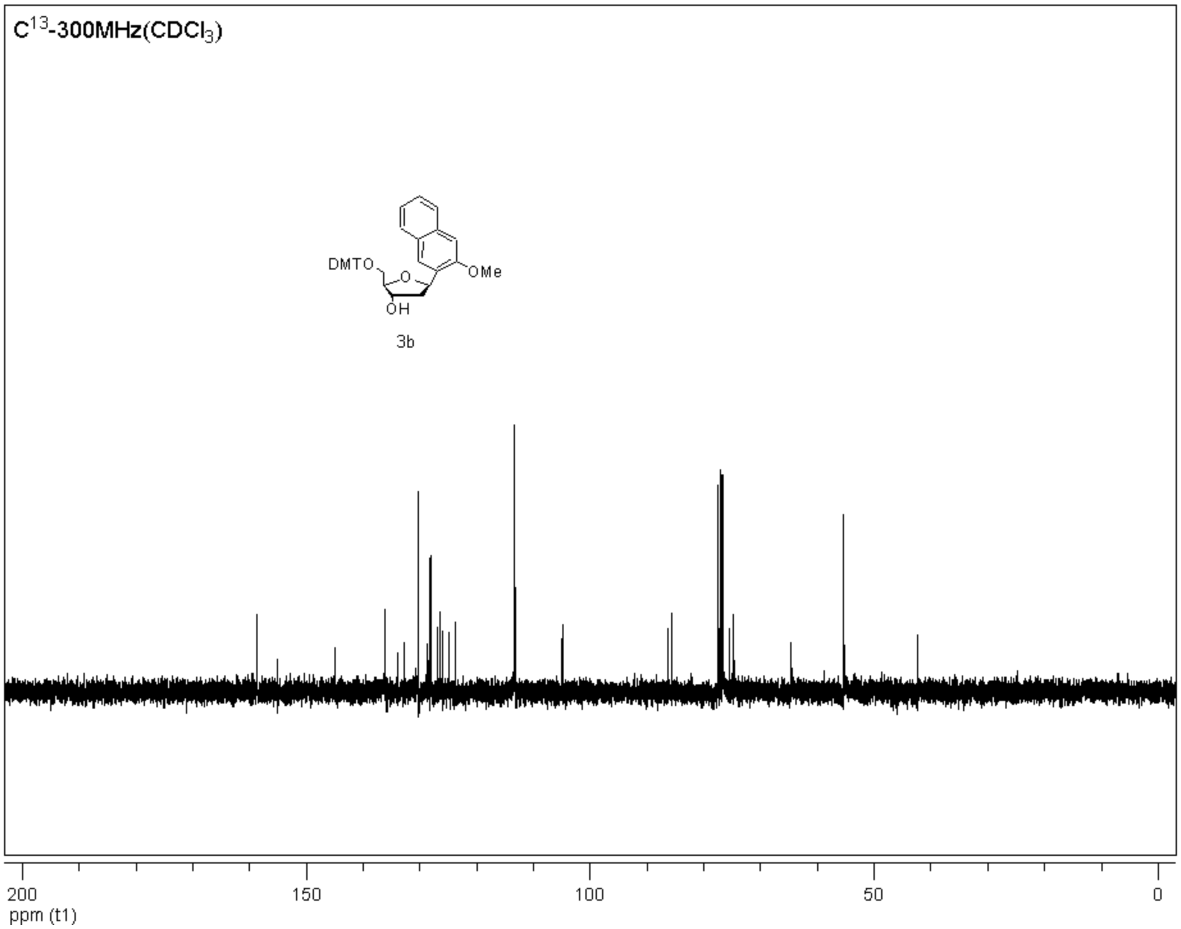
^{1}H-300MHz(MeOH-d_4)
HO
O
OH
OMe
2b
10.0
ppm (t1)
5.0
0.0

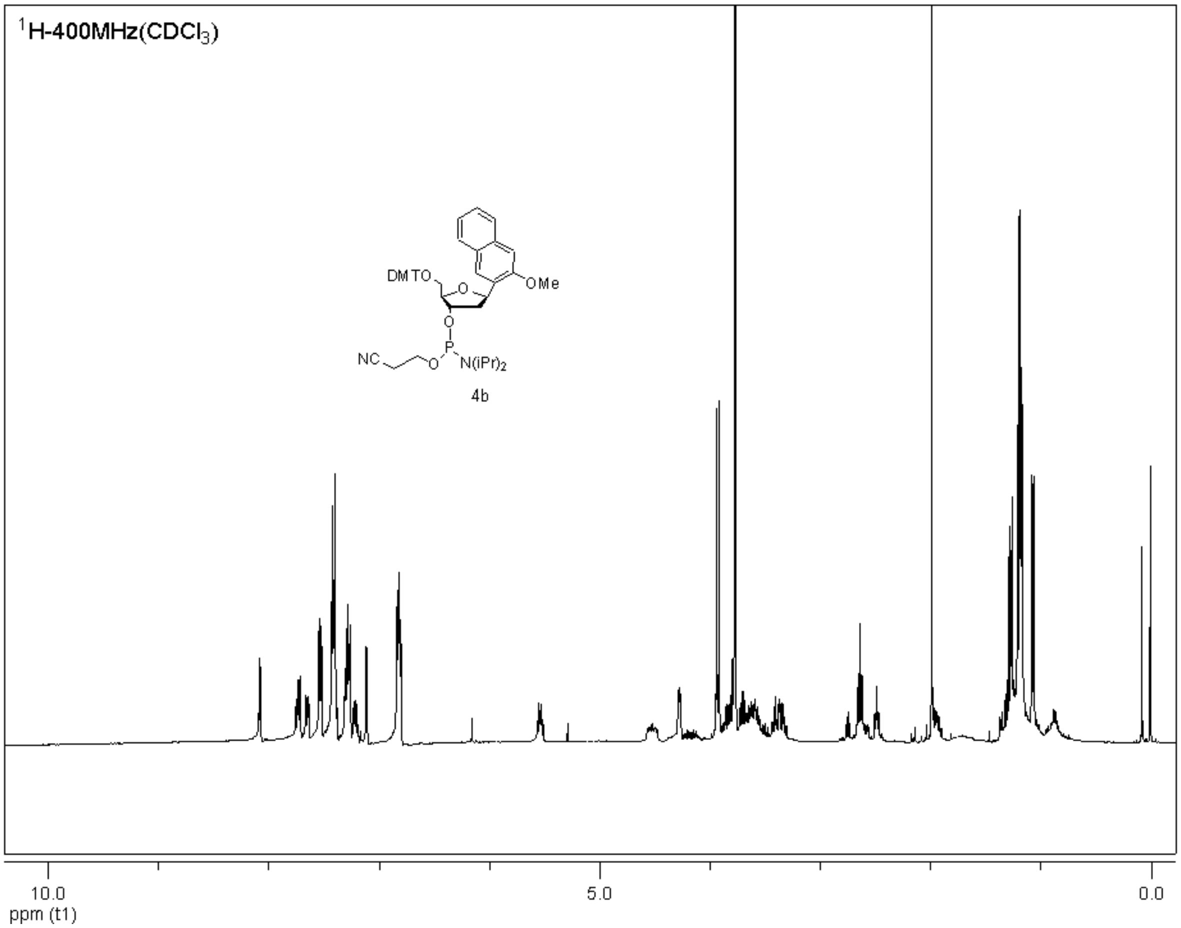


C13-300MHz(MeOH-d4)
HO
O
OH
OMe
2b
200
150
100
50
0
ppm (t1)

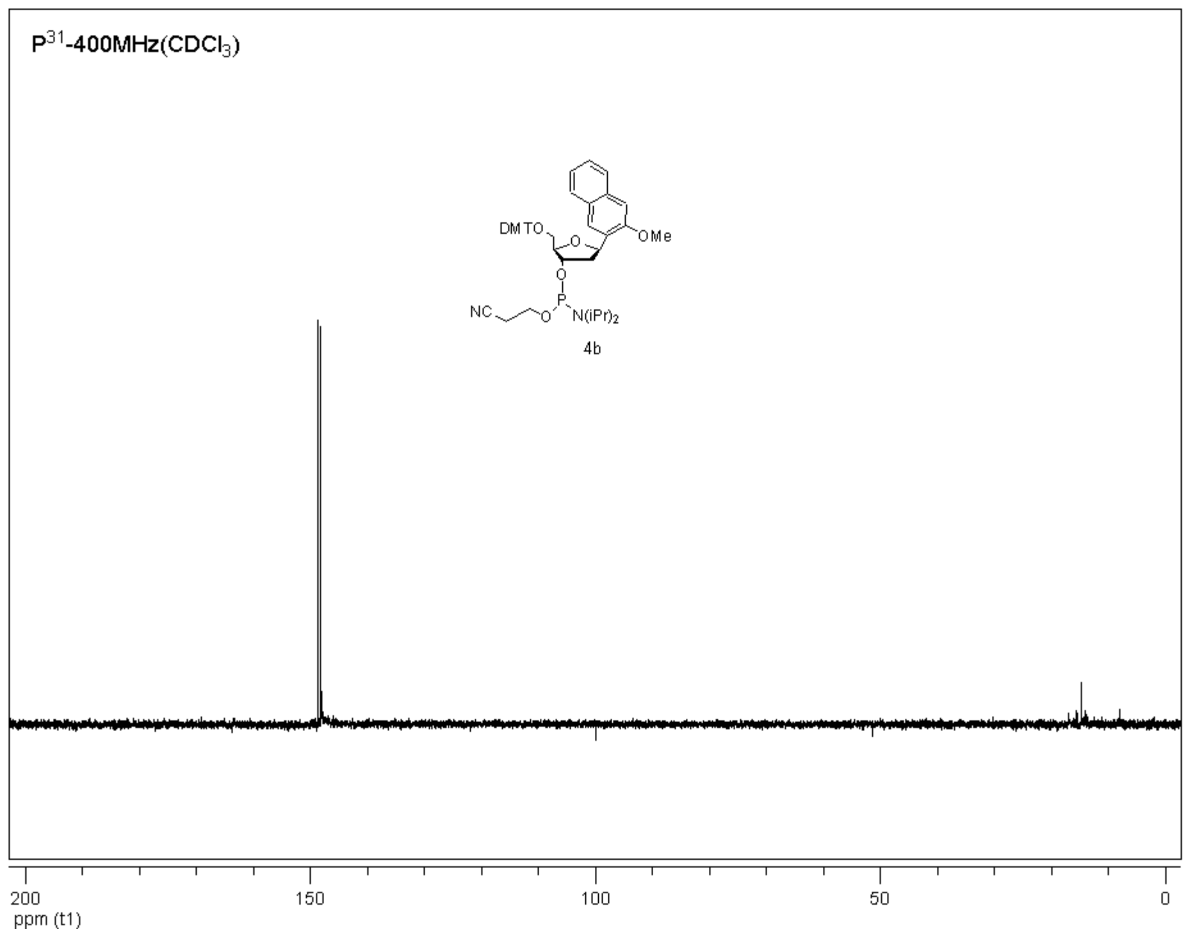


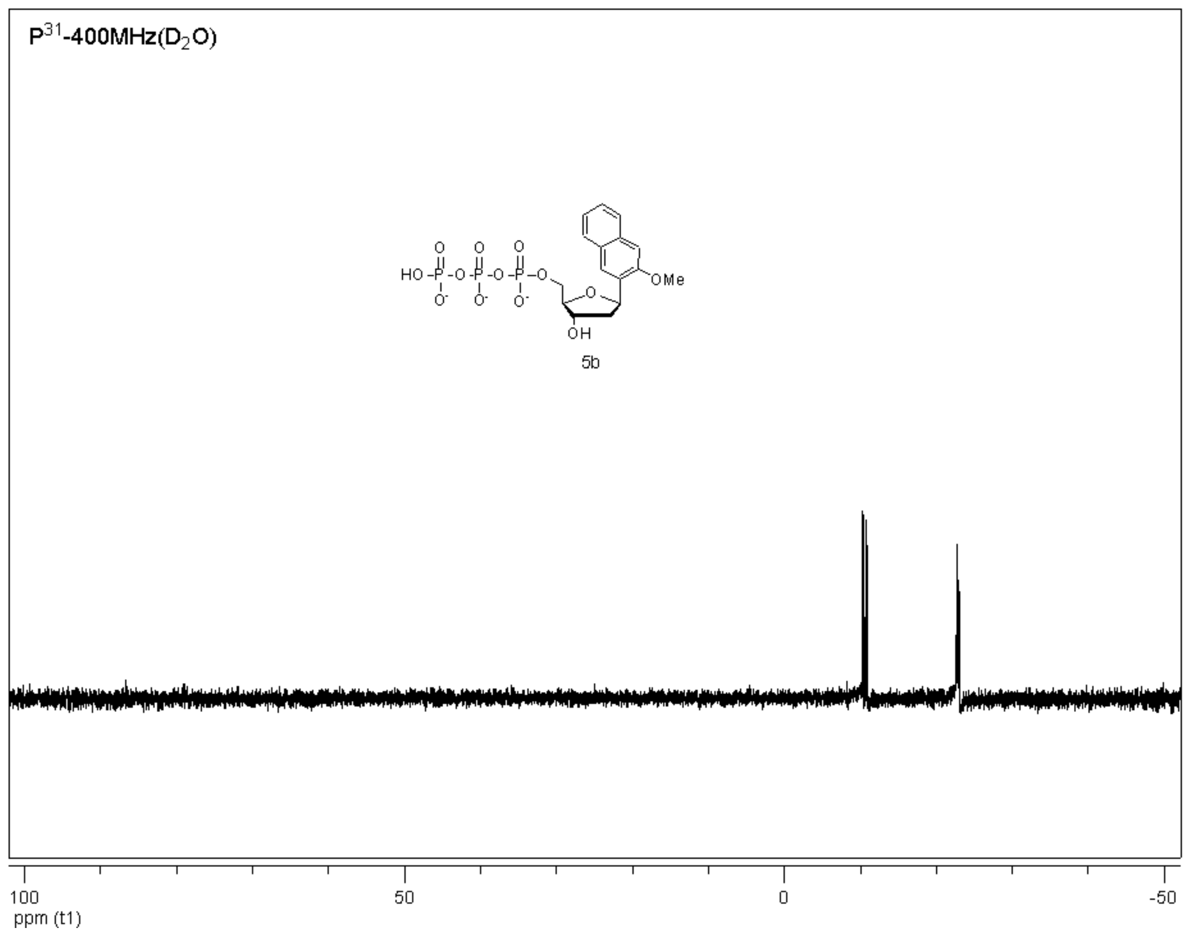
1H-300MHz(CDCl3)
DMTO
O
OMe
OH
3b
10.0
5.0
0.0
ppm (t1)

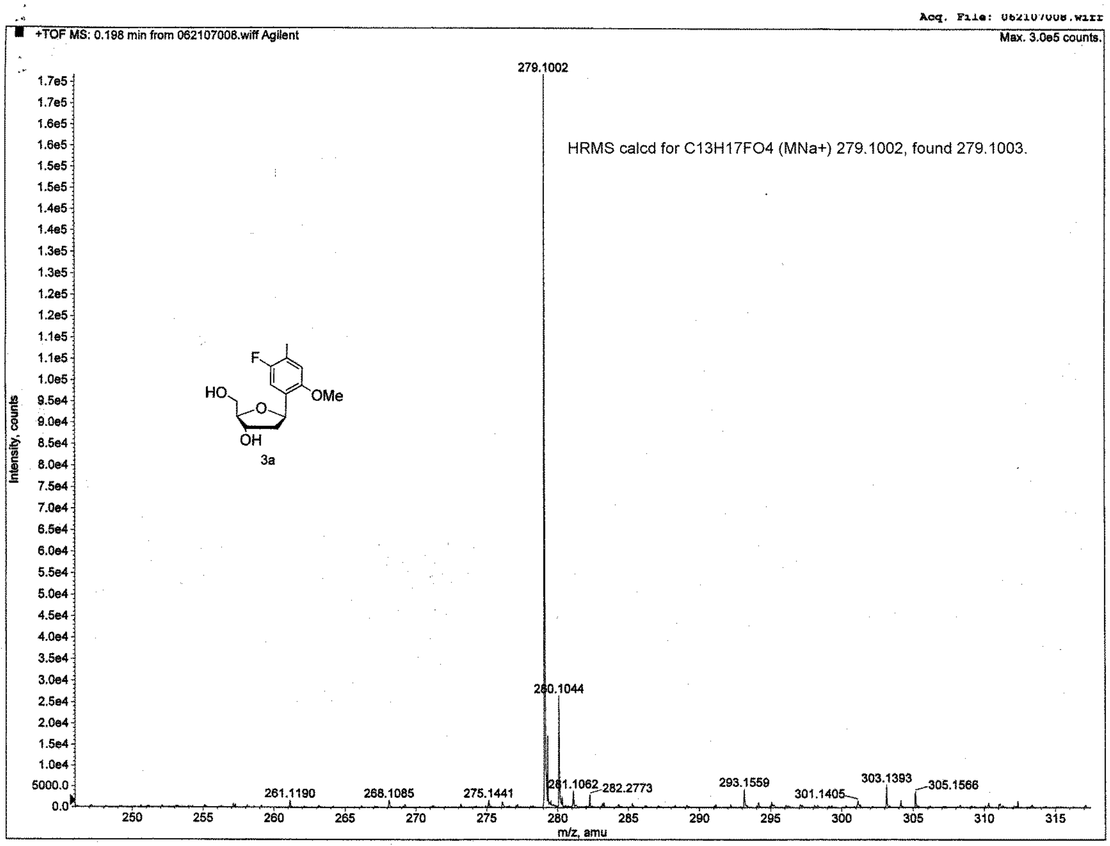

C13-300MHz(CDCl3)
DMTO
O
OMe
OH
3b
200
150
100
50
0
ppm (t1)

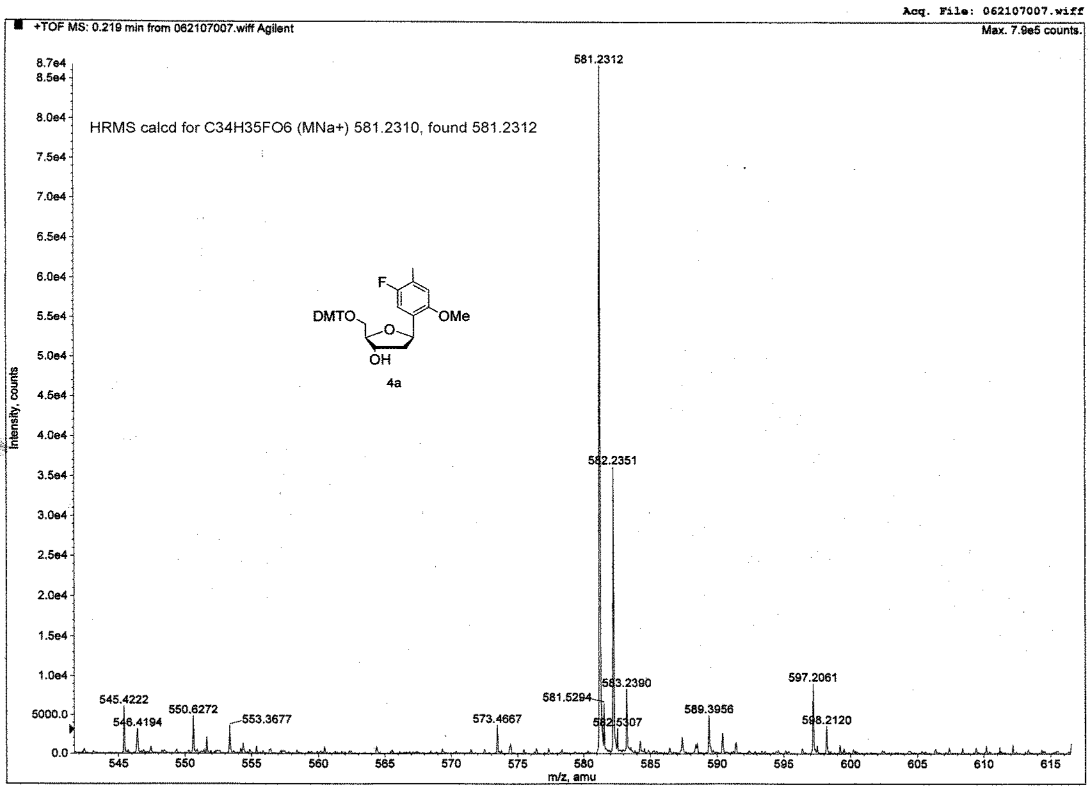


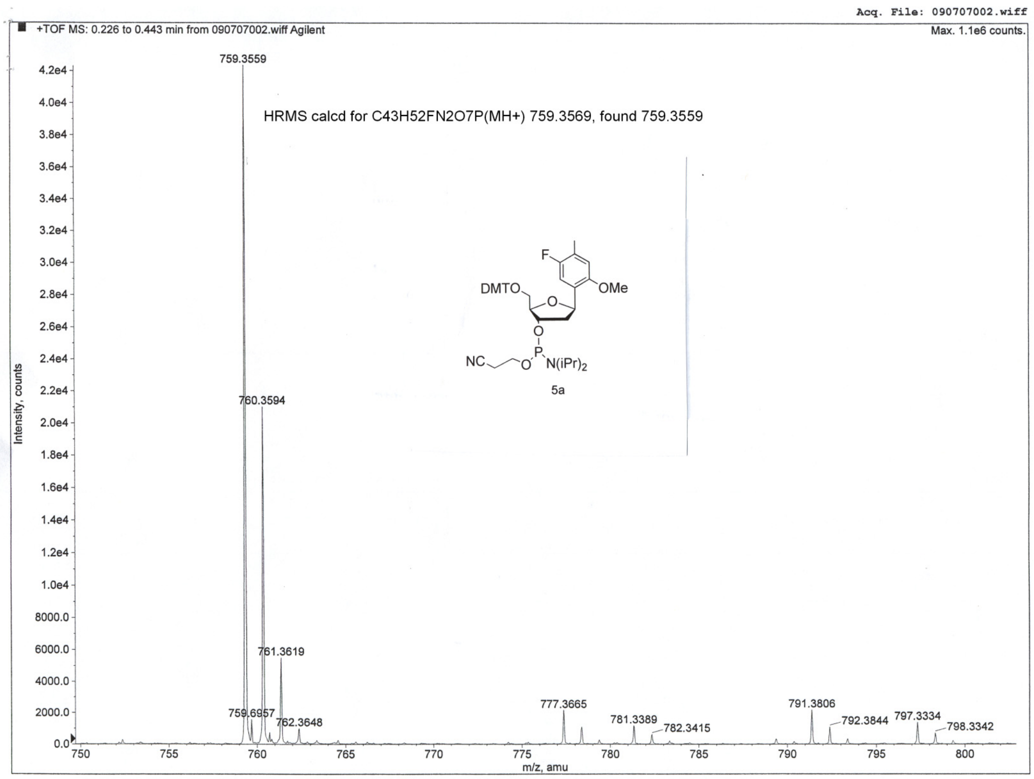
^{1}H-400MHz(CDCl$_3$)
DMTO
O
OMe
O
P
NC
O
N(iPr)$_2$
4b
10.0
5.0
0.0
ppm (t1)

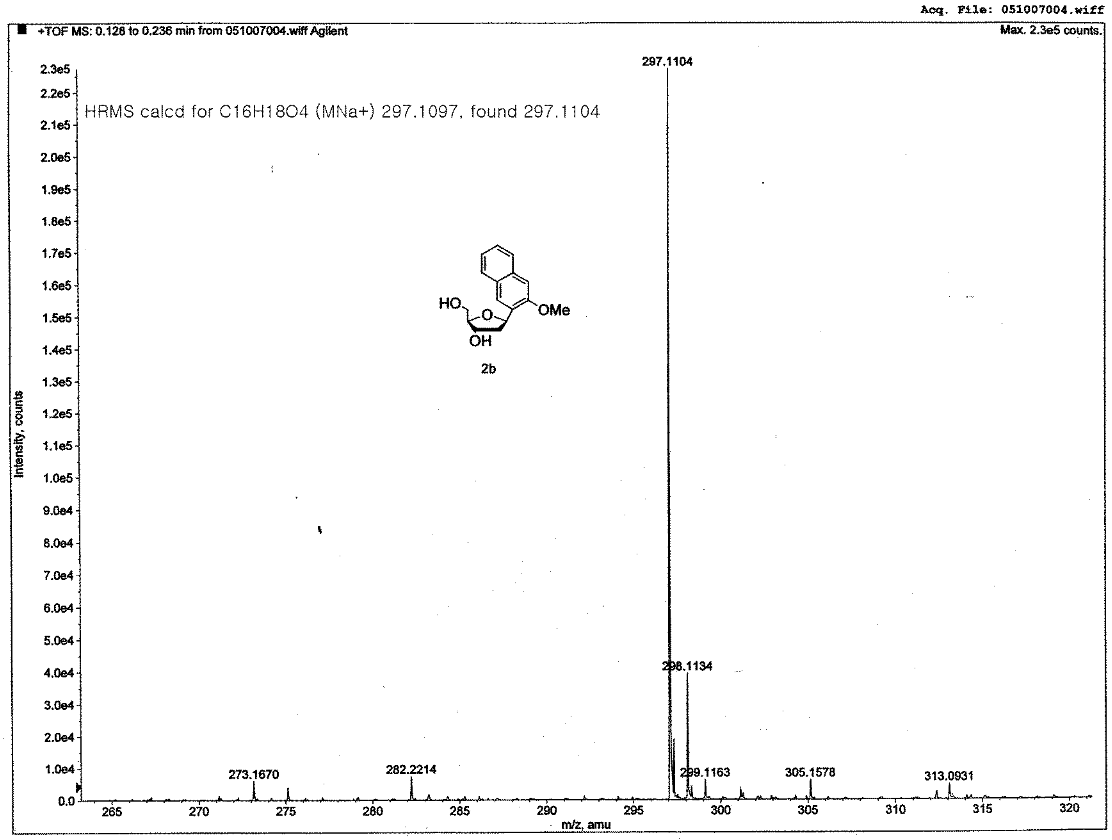

P31-400MHz(CDCl3)
DMTO
OMe
NC
N(iPr)2
4b
200
150
100
50
0
ppm (t1)

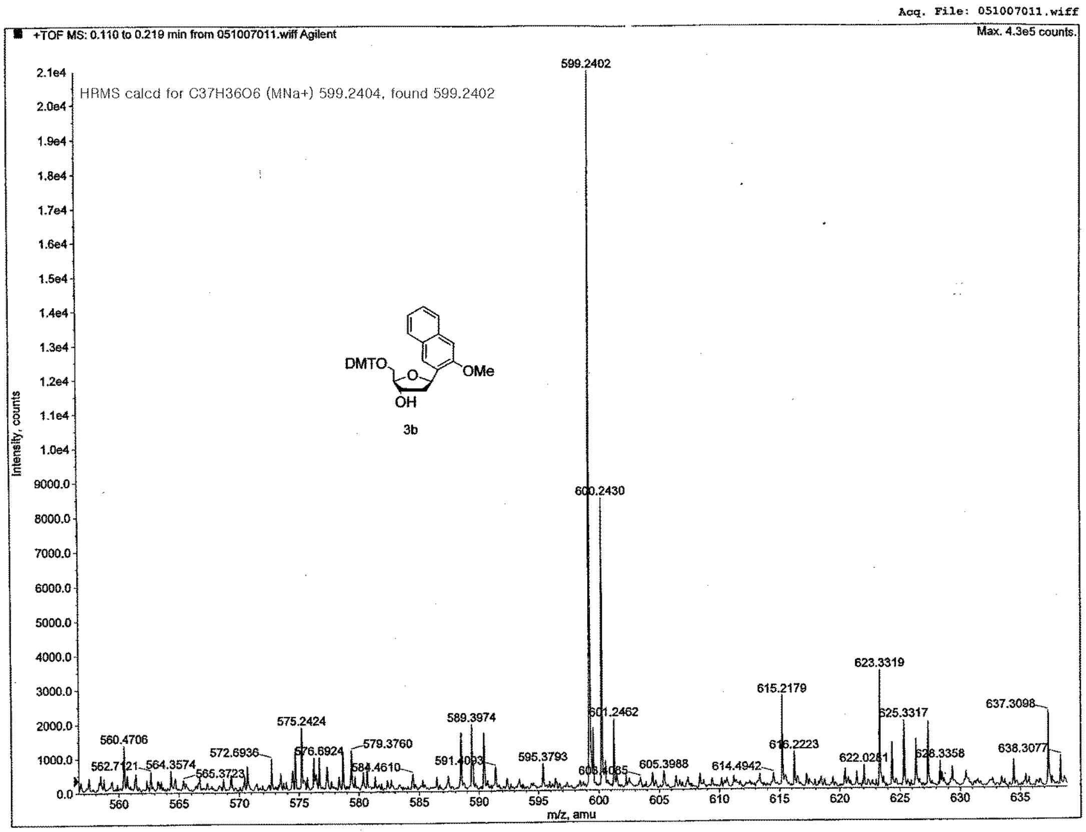

P31-400MHz(D2O)
HO
O
P
O-
OMe
OH
5b
100
50
0
-50
ppm (t1)


Acq. File: 062107008.wiff
+TOF MS: 0.198 min from 062107008.wiff Agilent
Max. 3.0e5 counts.
279.1002
HRMS calcd for C13H17FO4 (MNa+) 279.1002, found 279.1003.
HO
F
OMe
OH
3a
280.1044
261.1190
268.1085
275.1441
281.1062
282.2773
293.1559
301.1405
303.1393
305.1566
Intensity, counts
m/z, amu

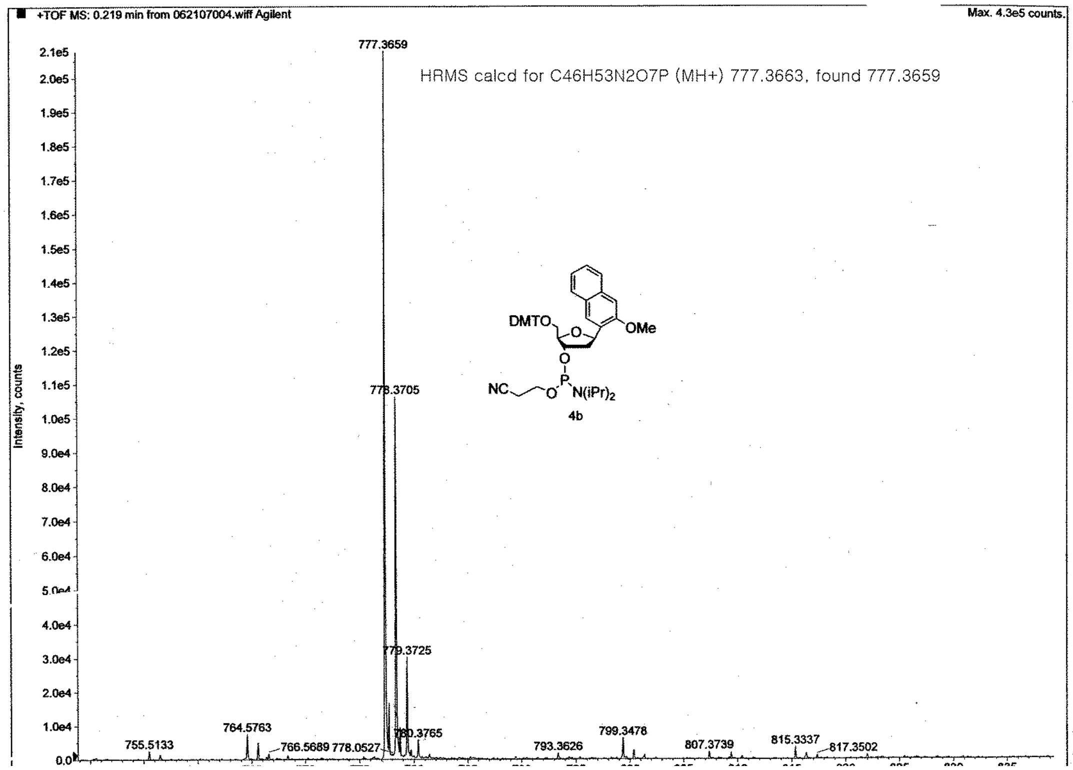

Acq. File: 062107007.wiff
+TOF MS: 0.219 min from 062107007.wiff Agilent
Max. 7.9e5 counts.
HRMS calcd for C34H35FO6 (MNa+) 581.2310, found 581.2312
DMTO
F
OMe
O
OH
4a
581.2312
582.2351
583.2390
581.5294
582.5307
597.2061
598.2120
589.3956
573.4667
545.4222
546.4194
550.6272
553.3677
Intensity, counts
m/z, amu

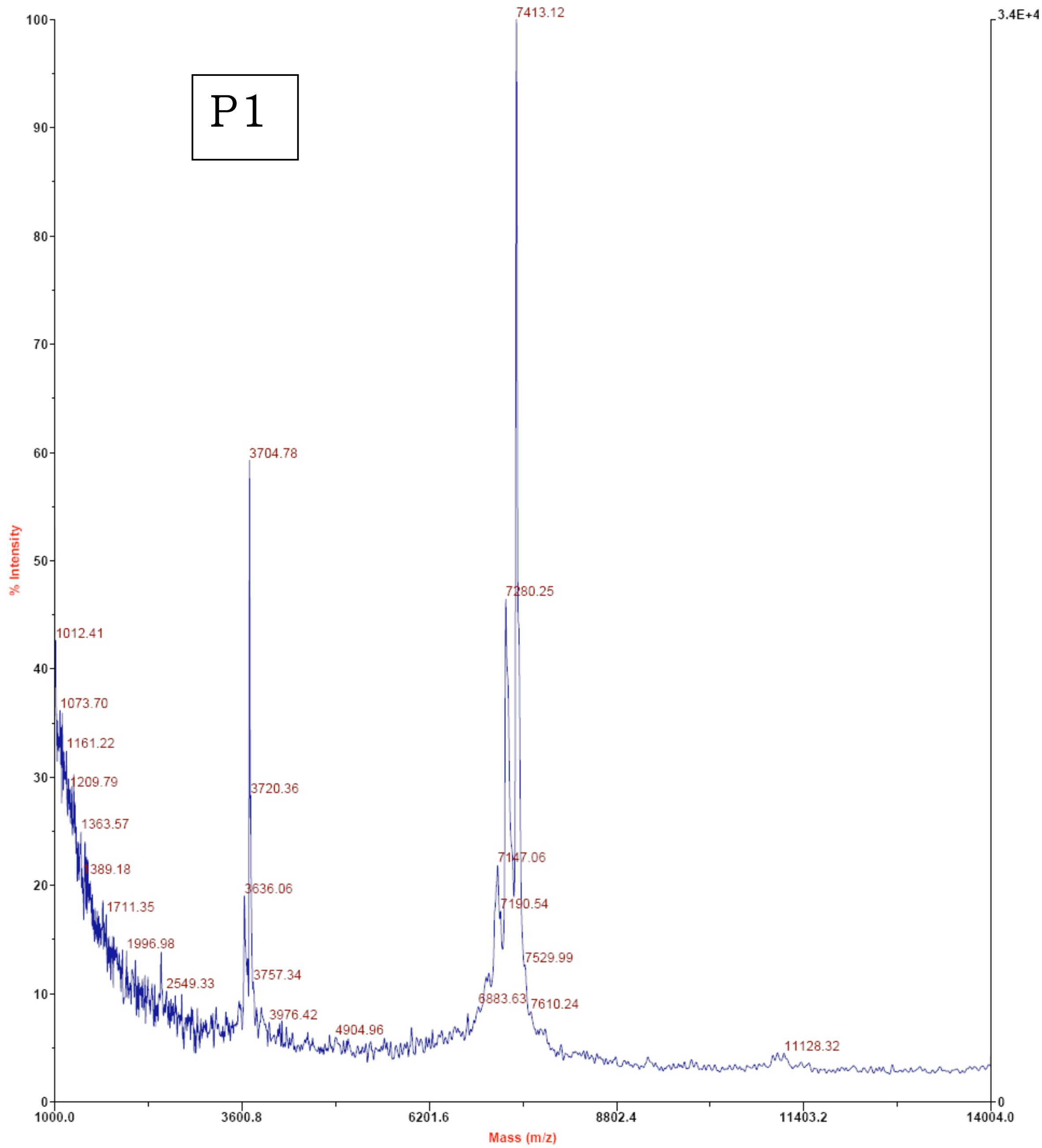

Acq. File: 090707002.wiff
+TOF MS: 0.226 to 0.443 min from 090707002.wiff Agilent
Max. 1.1e6 counts.
HRMS calcd for C43H52FN2O7P(MH+) 759.3569, found 759.3559
759.3559
760.3594
761.3619
759.6957
762.3648
777.3665
781.3389
782.3415
791.3806
792.3844
797.3334
798.3342
F
OMe
DMTO
O
P
NC
N(iPr)2
5a
Intensity, counts
m/z, amu

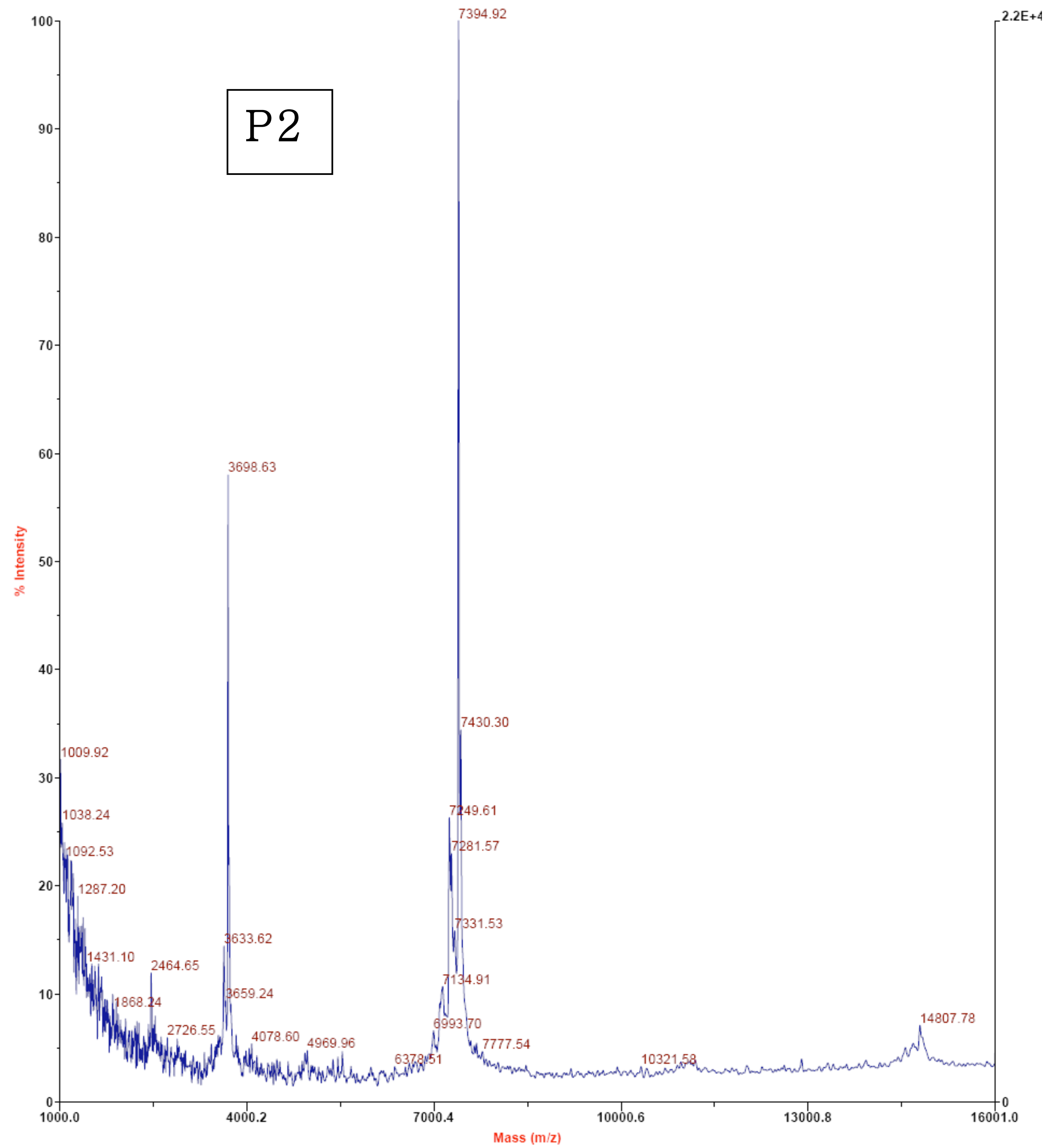

Acq. File: 051007004.wiff
+TOF MS: 0.128 to 0.236 min from 051007004.wiff Agilent
Max. 2.3e5 counts.
HRMS calcd for C16H18O4 (MNa+) 297.1097, found 297.1104
297.1104
298.1134
299.1163
273.1670
282.2214
305.1578
313.0931
HO
O
OMe
OH
2b
Intensity, counts
m/z, amu

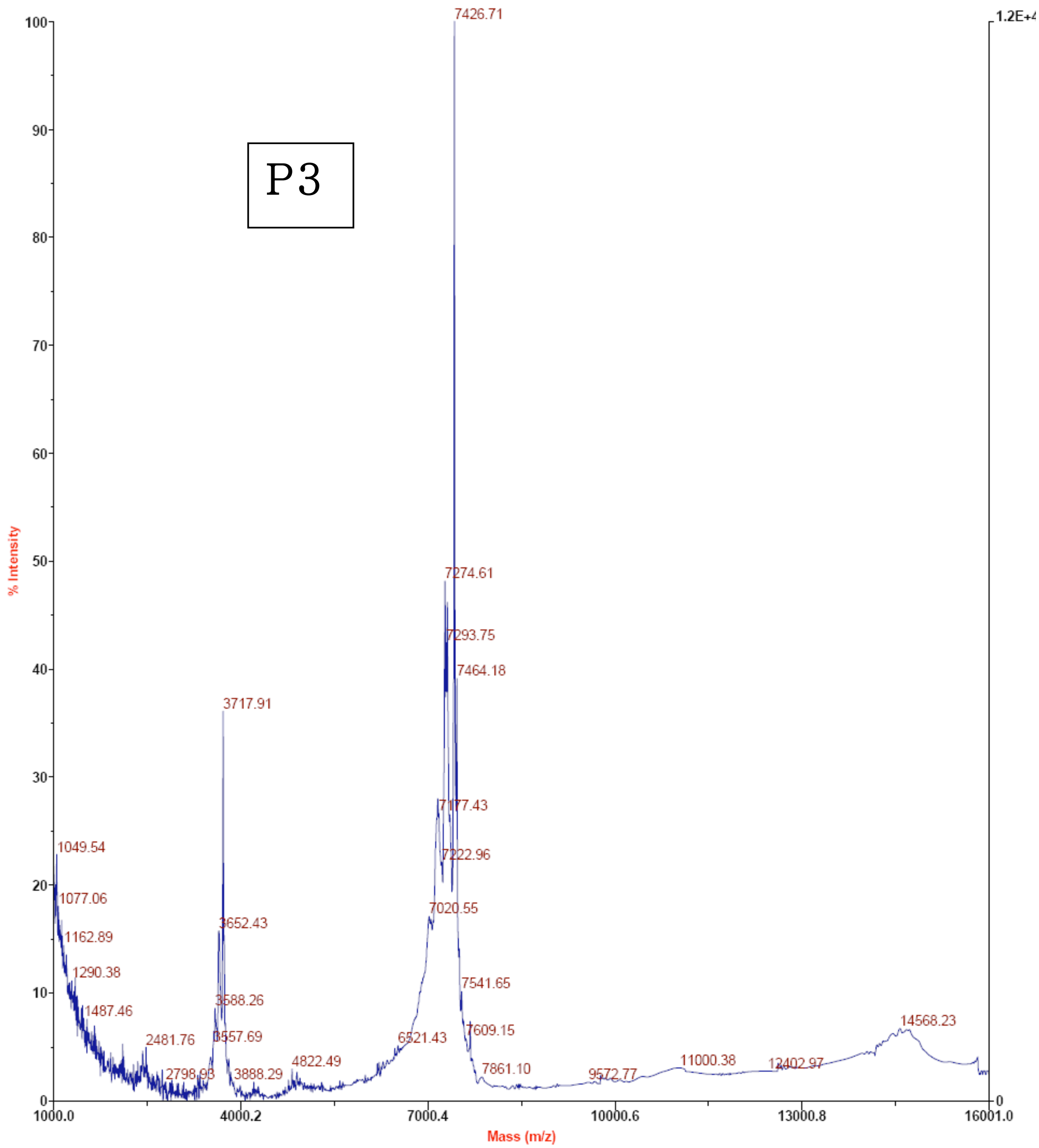

Acq. File: 051007011.wiff
+TOF MS: 0.110 to 0.219 min from 051007011.wiff Agilent
Max. 4.3e5 counts.
HRMS calcd for C37H36O6 (MNa+) 599.2404, found 599.2402
DMTO
OMe
OH
3b
599.2402
600.2430
623.3319
615.2179
601.2462
625.3317
637.3098
Intensity, counts
m/z, amu

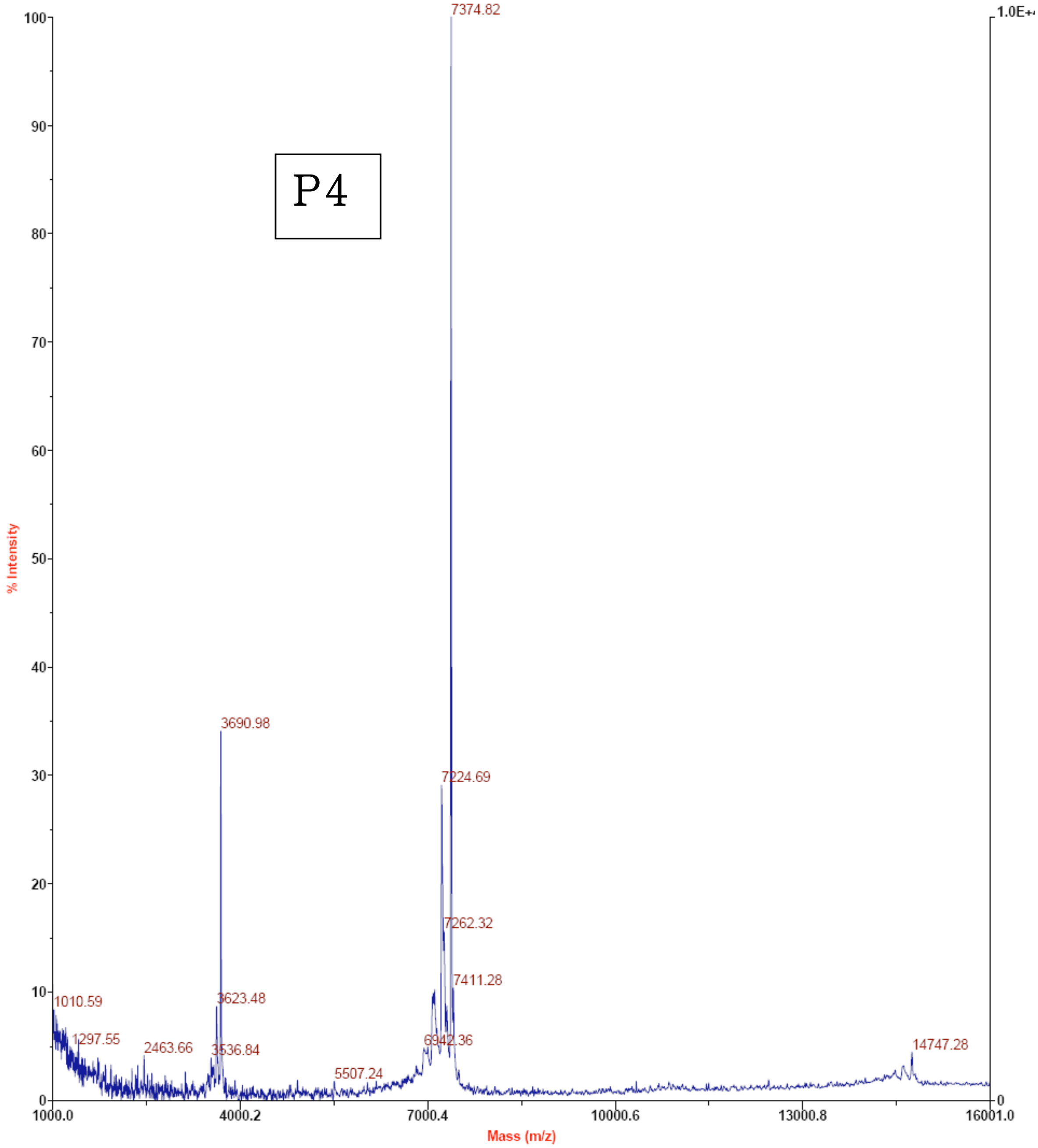
Polarity/Scan Type: Positive   Sample Name: rbys5757   Acq. I   lay, June 21, 2007
le: 062107004.wiff



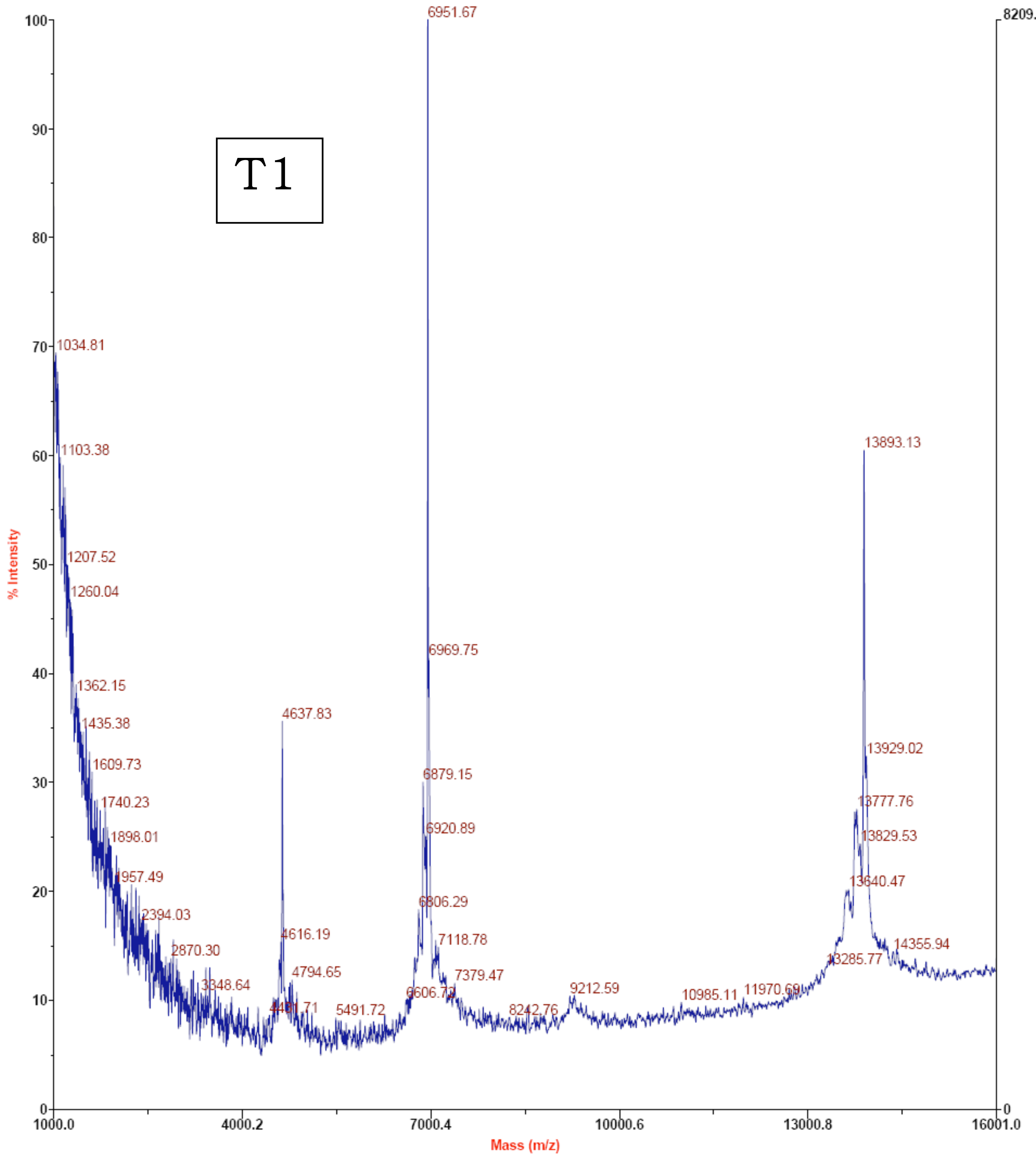

P1
7413.12
3704.78
7280.25
1012.41
1073.70
1161.22
1209.79
1363.57
1389.18
1711.35
1996.98
2549.33
3720.36
3636.06
3757.34
3976.42
4904.96
7147.06
7190.54
7529.99
6883.63
7610.24
11128.32
% Intensity
Mass (m/z)
3.4E+4
1000.0
3600.8
6201.6
8802.4
11403.2
14004.0

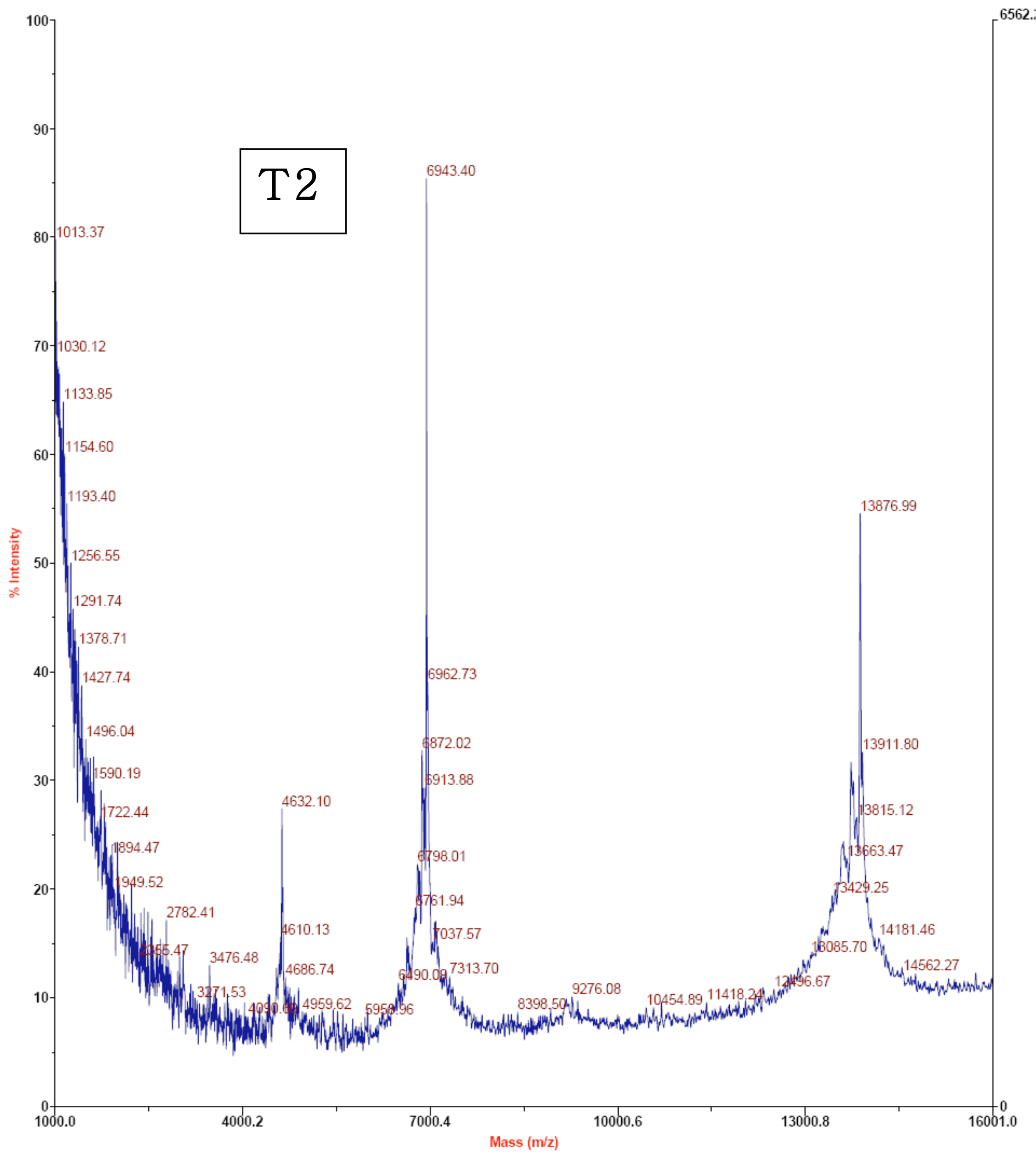

P2
7394.92
3698.63
7430.30
1009.92
1038.24
7249.61
1092.53
7281.57
1287.20
3633.62
7331.53
1431.10
2464.65
7134.91
3659.24
1868.24
2726.55
4078.60
6993.70
4969.96
7777.54
6378.51
10321.58
14807.78
% Intensity
Mass (m/z)
2.2E+4
100
90
80
70
60
50
40
30
20
10
0
1000.0
4000.2
7000.4
10000.6
13000.8
16001.0


P3
7426.71
7274.61
7293.75
7464.18
3717.91
7177.43
7222.96
1049.54
1077.06
7020.55
3652.43
1162.89
1290.38
7541.65
3588.26
1487.46
7609.15
14568.23
2481.76
3557.69
6521.43
4822.49
11000.38
12402.97
2798.93
3888.29
7861.10
9572.77
% Intensity
Mass (m/z)
1.2E+4
1000.0
4000.2
7000.4
10000.6
13000.8
16001.0

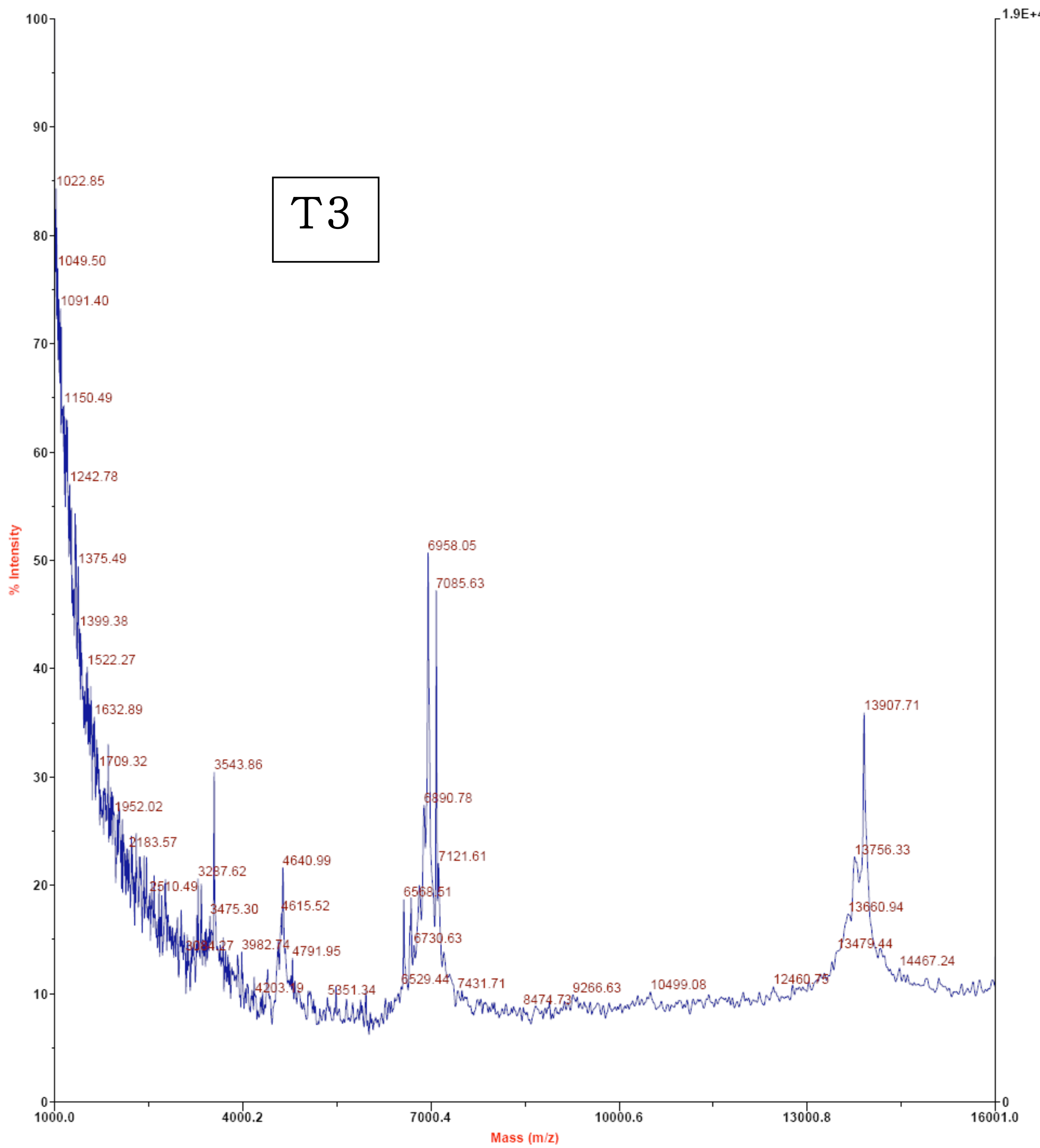


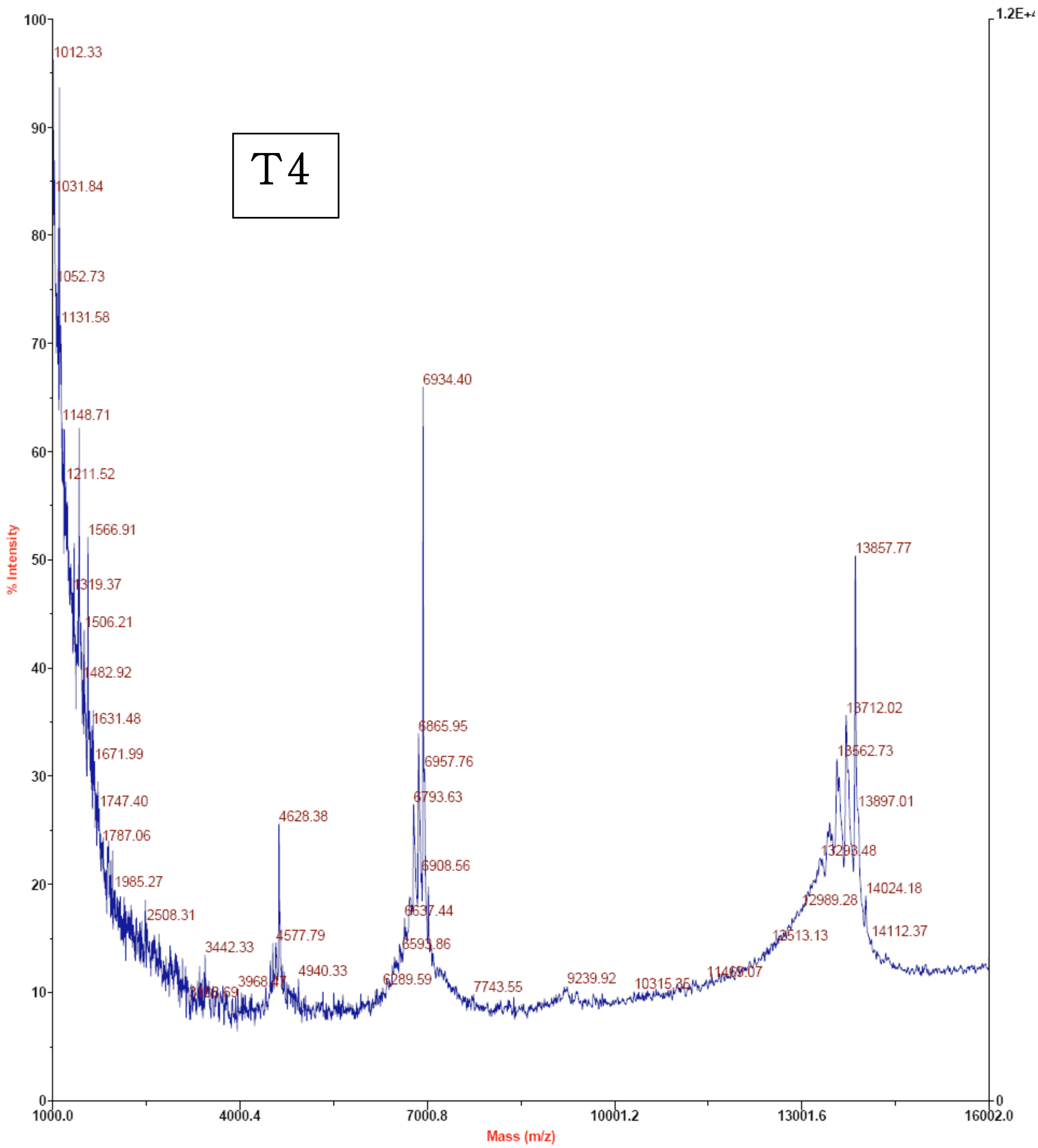
P4
% Intensity
Mass (m/z)
7374.82
3690.98
7224.69
7262.32
7411.28
3623.48
1010.59
1297.55
2463.66
3536.84
6942.36
5507.24
14747.28
1.0E+4
100
90
80
70
60
50
40
30
20
10
0
1000.0
4000.2
7000.4
10000.6
13000.8
16001.0

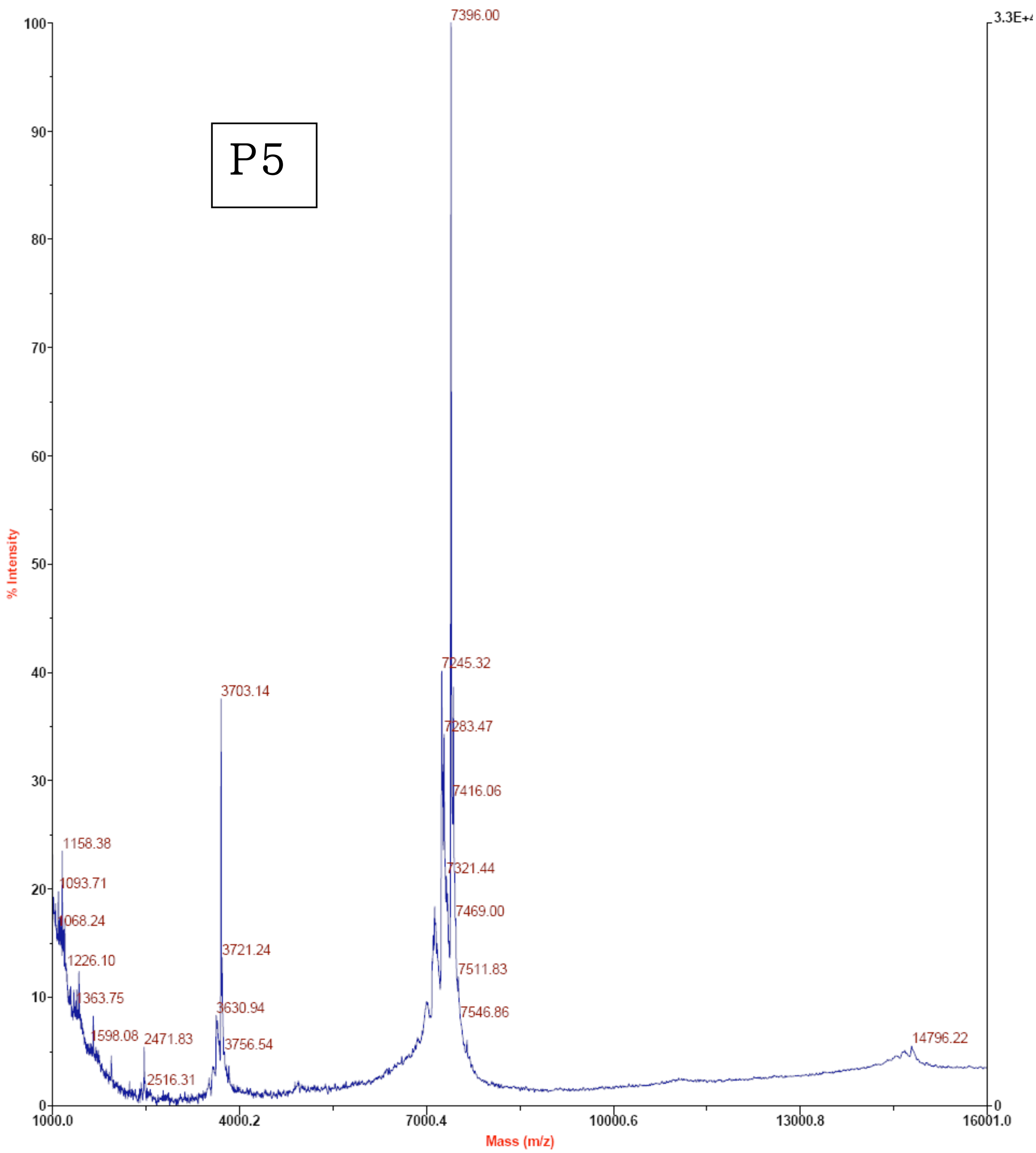

T1
% Intensity
Mass (m/z)
6951.67
13893.13
1034.81
1103.38
1207.52
1260.04
6969.75
1362.15
1435.38
4637.83
1609.73
13929.02
6879.15
1740.23
13777.76
6920.89
1898.01
13829.53
1957.49
13640.47
6806.29
2394.03
4616.19
7118.78
14355.94
2870.30
13285.77
4794.65
7379.47
3348.64
9212.59
6606.72
10985.11
11970.69
5491.72
8242.76
8209.
1000.0
4000.2
7000.4
10000.6
13000.8
16001.0



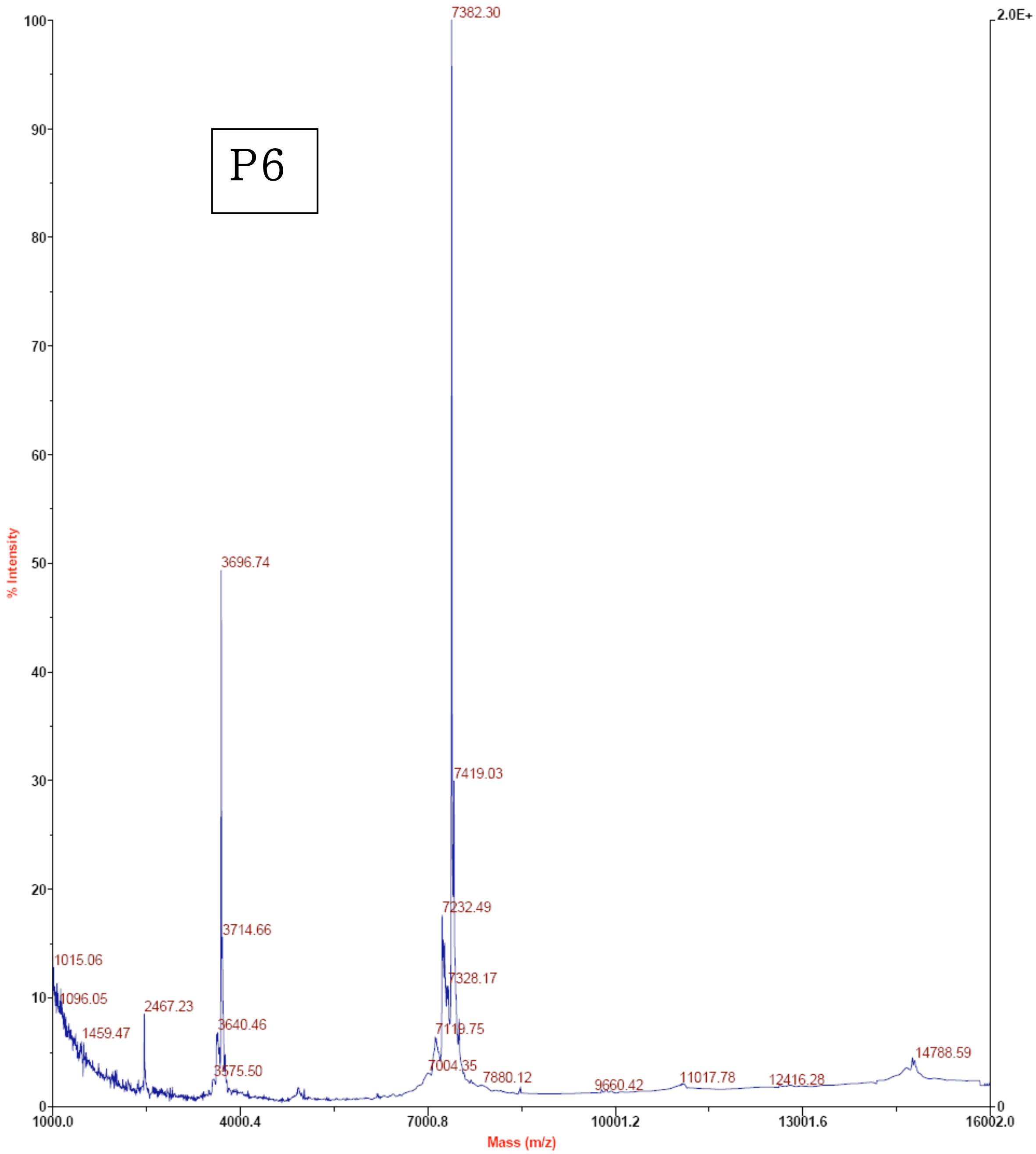
T2
% Intensity
Mass (m/z)
6943.40
13876.99
1013.37
4632.10

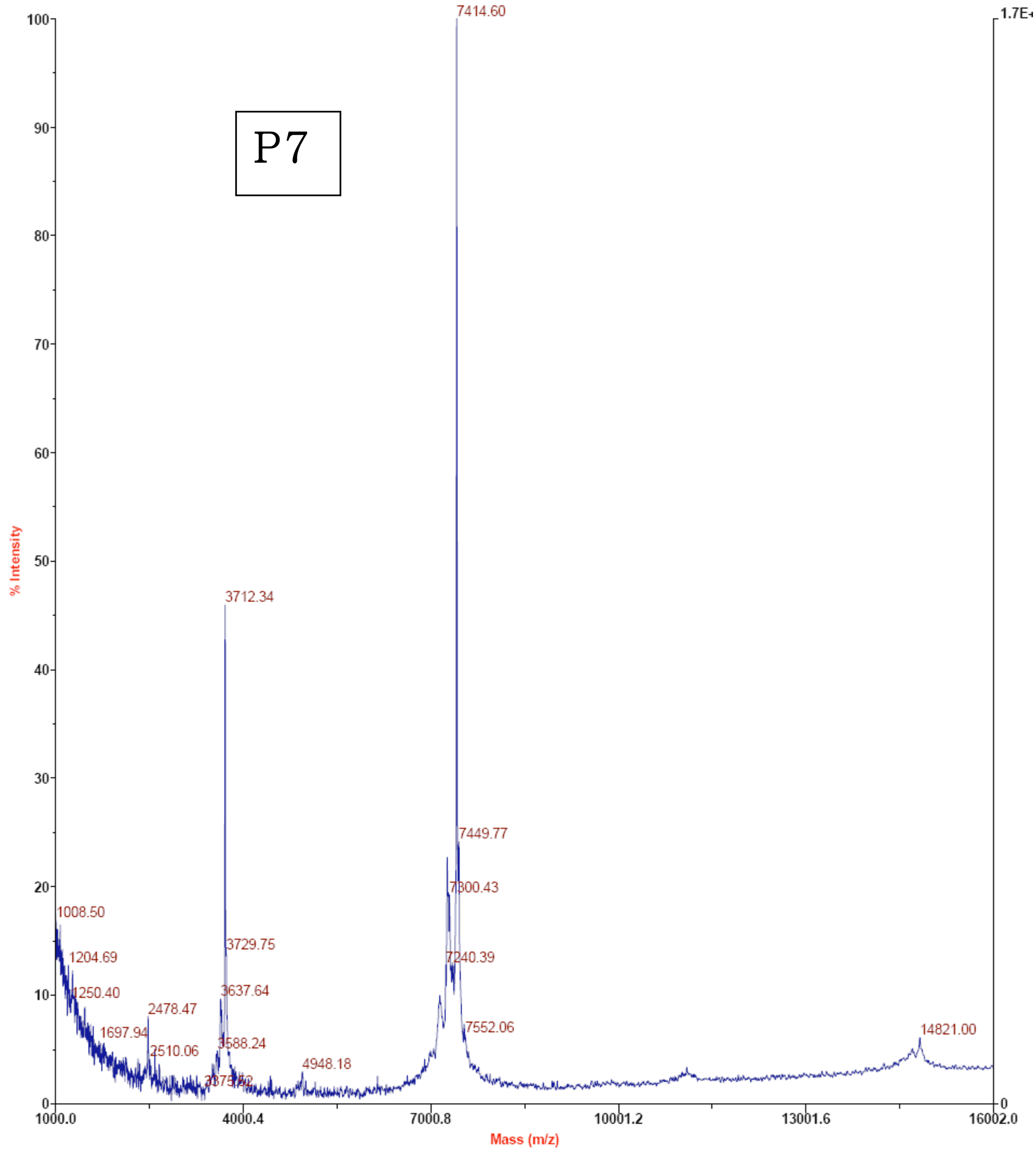

T3
% Intensity
Mass (m/z)
1.9E+4
1022.85
1049.50
1091.40
1150.49
1242.78
1375.49
1399.38
1522.27
1632.89
1709.32
1952.02
2183.57
2510.49
3287.62
3475.30
3543.86
3982.74
4203.79
4615.52
4640.99
4791.95
5351.34
6529.44
6568.51
6730.63
6890.78
6958.05
7085.63
7121.61
7431.71
8474.73
9266.63
10499.08
12460.75
13479.44
13660.94
13756.33
13907.71
14467.24
1000.0
4000.2
7000.4
10000.6
13000.8
16001.0


T4
% Intensity
Mass (m/z)
1.2E+4
1012.33
1031.84
1052.73
1131.58
1148.71
1211.52
1566.91
1319.37
1506.21
1482.92
1631.48
1671.99
1747.40
1787.06
1985.27
2508.31
3442.33
3968.47
4628.38
4577.79
4940.33
6934.40
6865.95
6957.76
6793.63
6908.56
6637.44
6593.86
6289.59
7743.55
9239.92
10315.35
13857.77
13712.02
13562.73
13897.01
13293.48
14024.18
12989.28
12513.13
14112.37
1000.0
4000.4
7000.8
10001.2
13001.6
16002.0

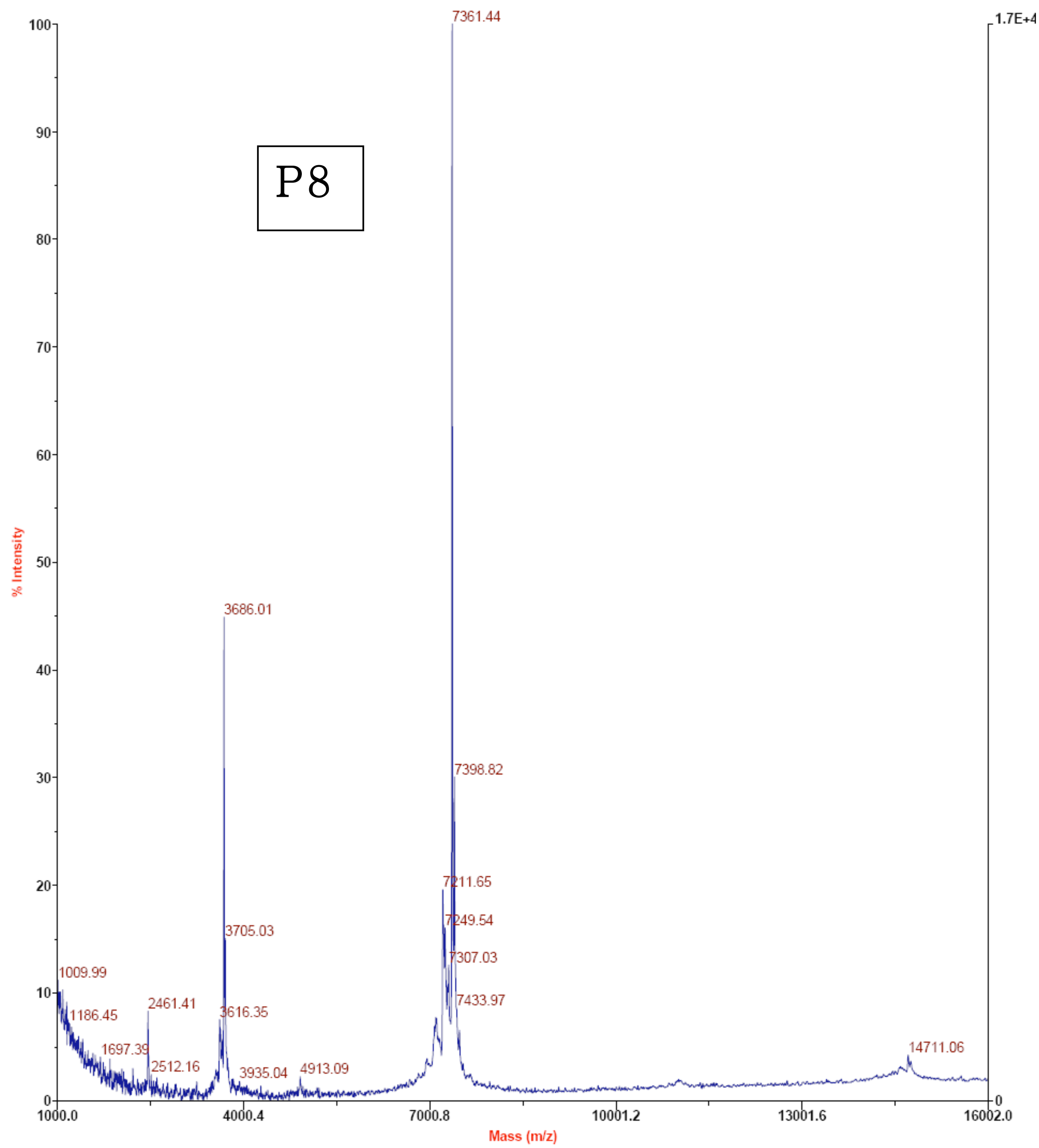


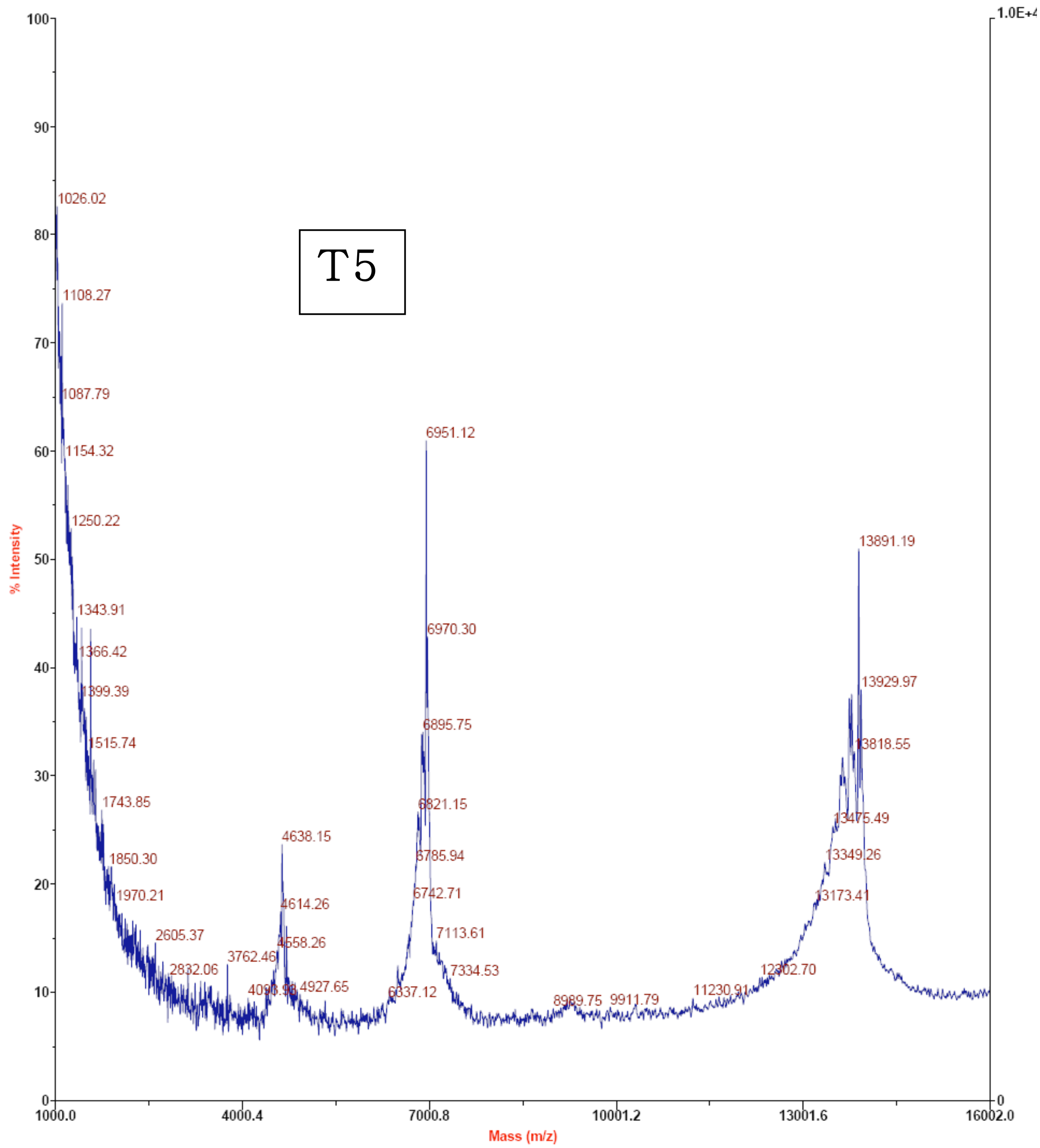
P5
7396.00
3.3E+4
100
90
80
70
60
50
40
30
20
10
0
% Intensity
7245.32
3703.14
7283.47
7416.06
1158.38
7321.44
1093.71
7469.00
1068.24
3721.24
1226.10
7511.83
1363.75
3630.94
7546.86
1598.08
2471.83
3756.54
14796.22
2516.31
1000.0
4000.2
7000.4
10000.6
13000.8
16001.0
Mass (m/z)

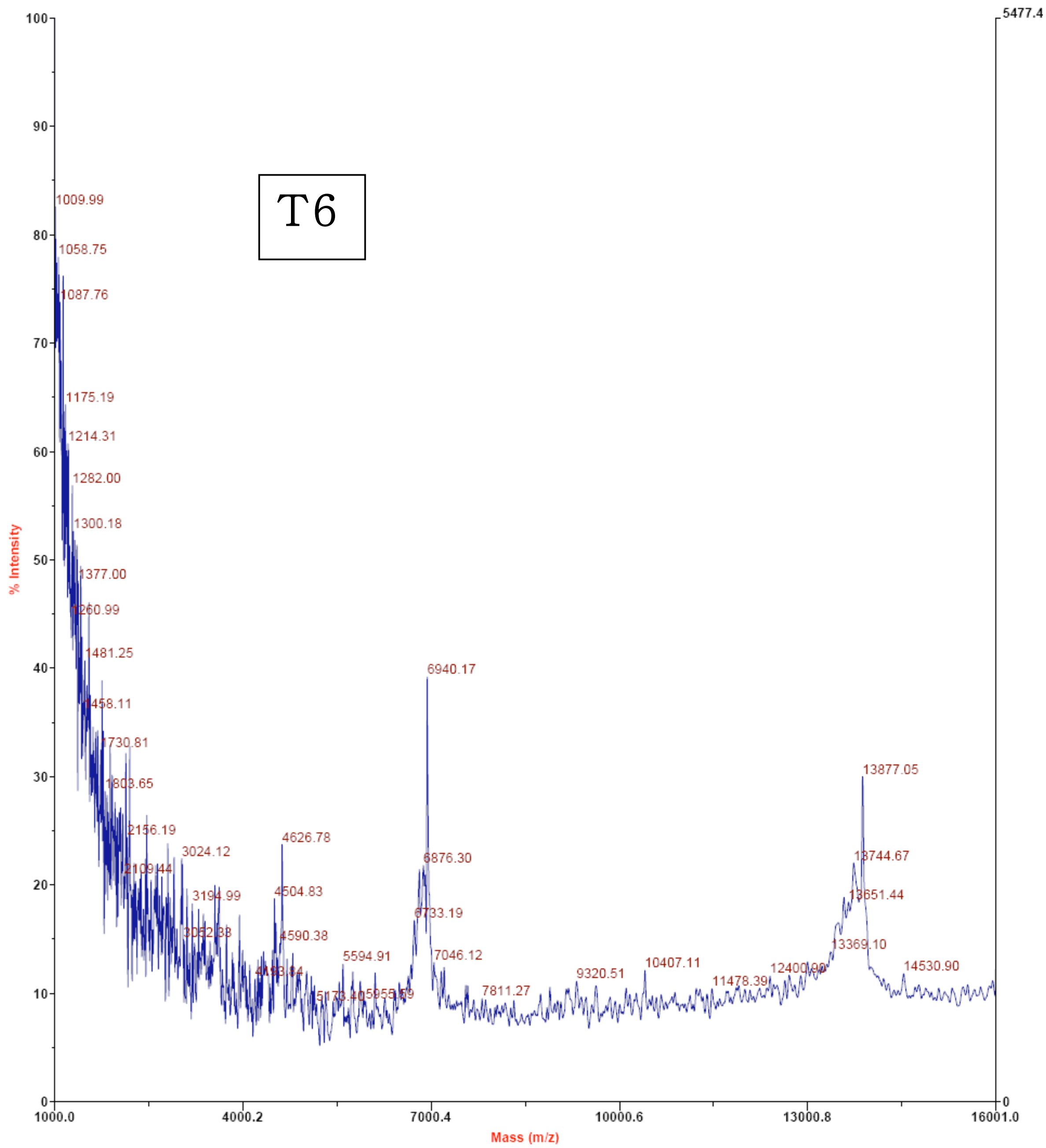

P6
7382.30
3696.74
7419.03
7232.49
3714.66
1015.06
7328.17
1096.05
2467.23
3640.46
1459.47
7119.75
14788.59
7004.35
3575.50
7880.12
11017.78
12416.28
9660.42
% Intensity
Mass (m/z)
2.0E+
100
90
80
70
60
50
40
30
20
10
0
1000.0
4000.4
7000.8
10001.2
13001.6
16002.0

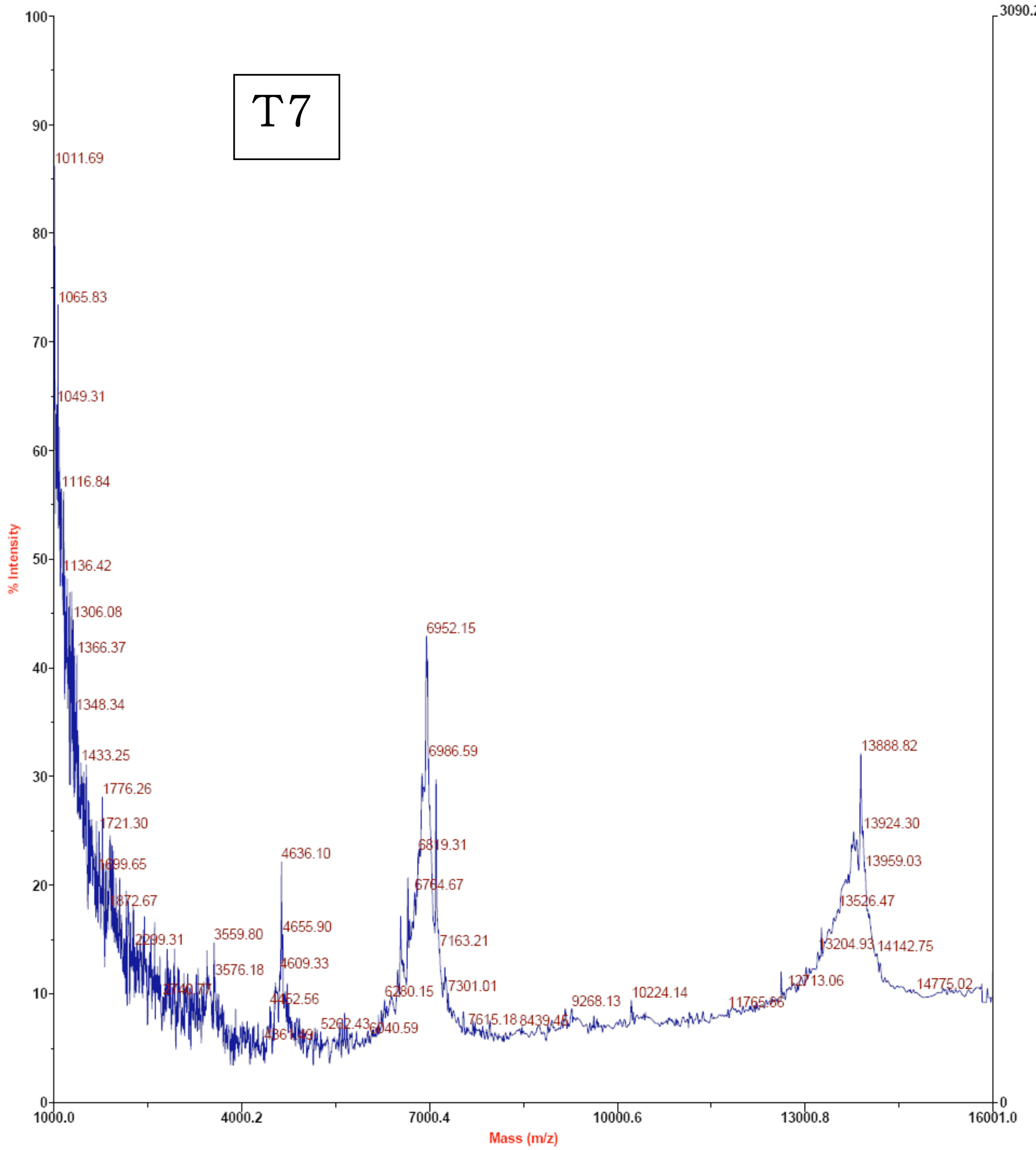


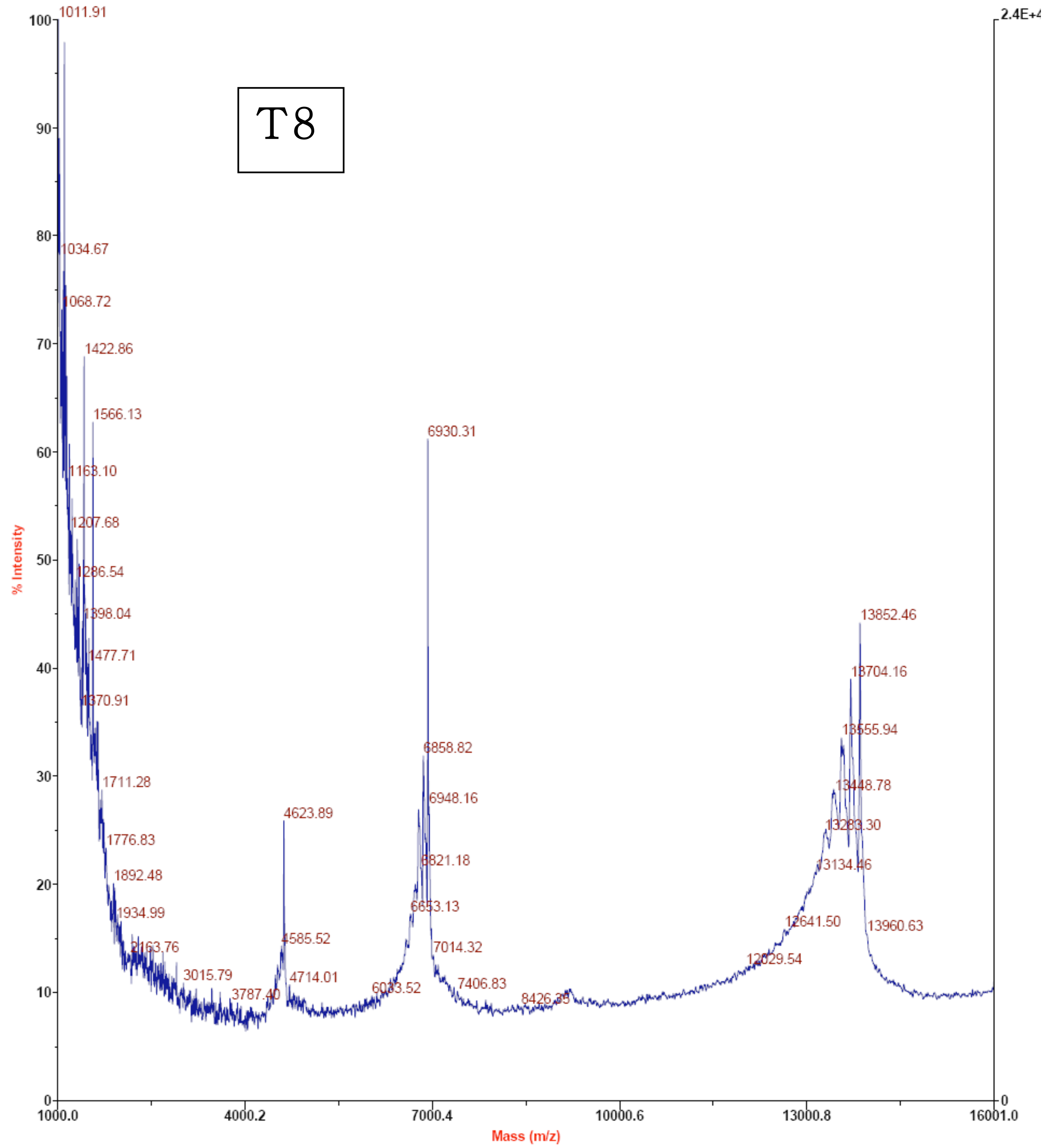
P7
7414.60
1.7E4
3712.34
7449.77
7300.43
1008.50
3729.75
1204.69
7240.39
3637.64
1250.40
2478.47
1697.94
7552.06
14821.00
3588.24
2510.06
4948.18
% Intensity
Mass (m/z)
1000.0
4000.4
7000.8
10001.2
13001.6
16002.0

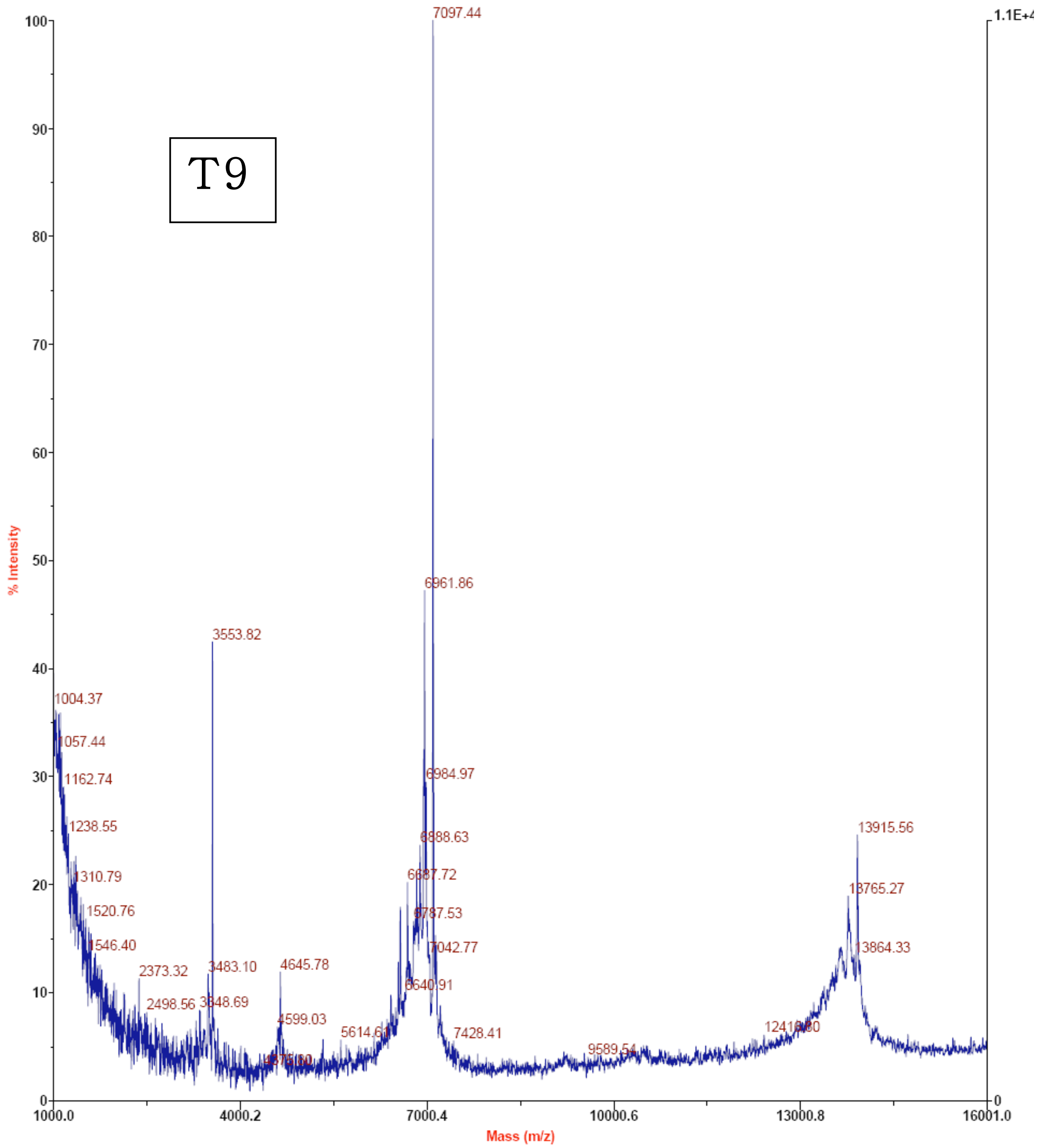

P8
7361.44
1.7E+4
3686.01
7398.82
7211.65
7249.54
3705.03
7307.03
1009.99
7433.97
2461.41
1186.45
3616.35
1697.39
14711.06
2512.16
3935.04
4913.09
% Intensity
Mass (m/z)
1000.0
4000.4
7000.8
10001.2
13001.6
16002.0

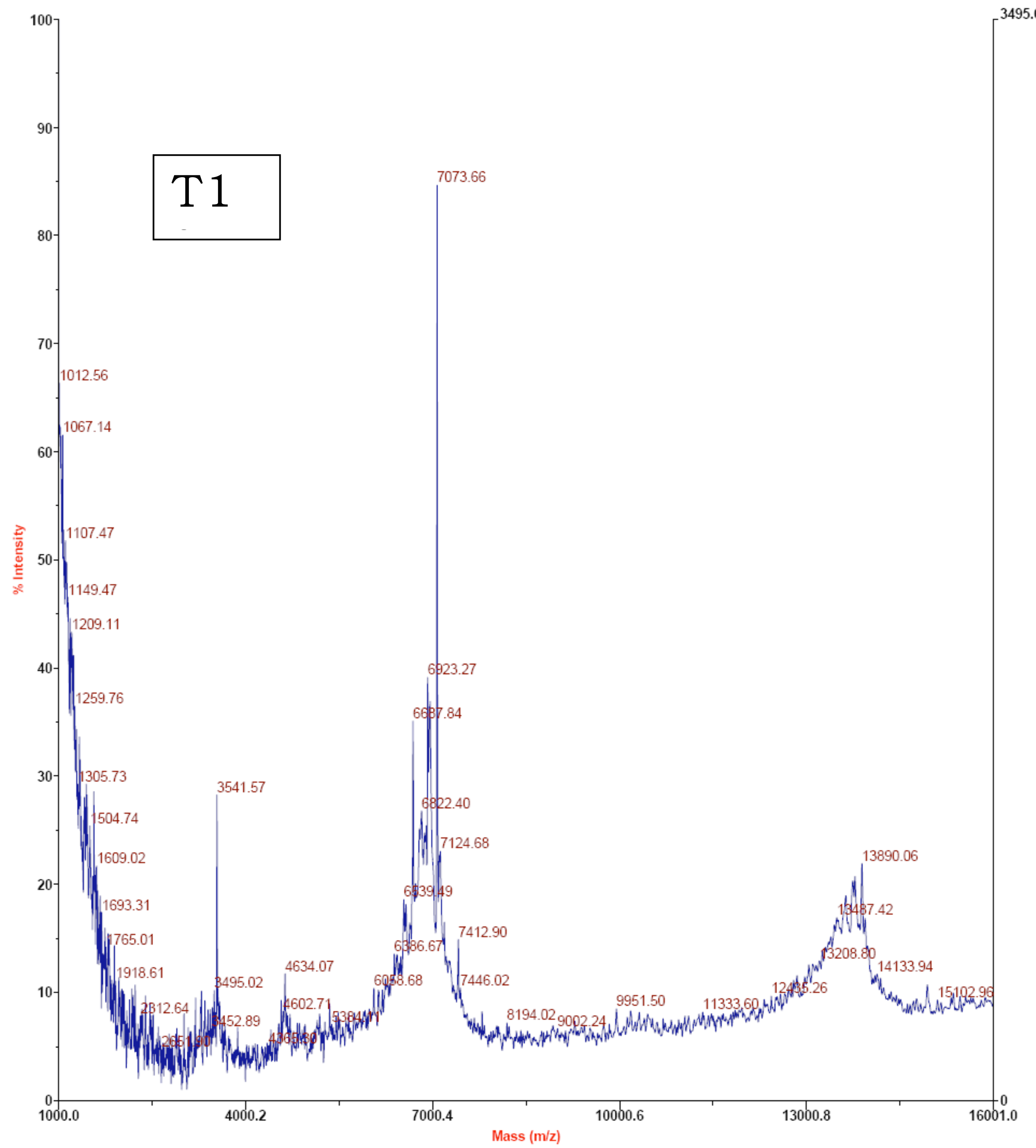

T5
% Intensity
Mass (m/z)
1.0E+4
1026.02
1108.27
1087.79
1154.32
1250.22
1343.91
1366.42
1399.39
1515.74
1743.85
1850.30
1970.21
2605.37
2832.06
3762.46
4638.15
4614.26
4558.26
4927.65
6951.12
6970.30
6895.75
6821.15
6785.94
6742.71
7113.61
7334.53
6337.12
8989.75
9911.79
11230.91
12802.70
13891.19
13929.97
13818.55
13475.49
13349.26
13173.41
1000.0
4000.4
7000.8
10001.2
13001.6
16002.0


T6
% Intensity
Mass (m/z)
6940.17
13877.05
4626.78
1009.99
5477.4


T7
% Intensity
Mass (m/z)
6952.15
13888.82
4636.10
1011.69
1065.83
1049.31
1116.84
1136.42
1306.08
1366.37
1348.34
1433.25
1776.26
1721.30
6986.59
13924.30
13959.03
13526.47
14142.75
14775.02
10224.14
9268.13
7163.21
7301.01
3559.80
3576.18
4655.90
4609.33
2299.31
12713.06
13204.93
3090.2


T8
% Intensity
Mass (m/z)
2.4E+4
1011.91
1034.67
1068.72
1422.86
1566.13
1163.10
1207.68
1286.54
1398.04
1477.71
1370.91
1711.28
1776.83
1892.48
1934.99
2163.76
3015.79
3787.40
4623.89
4585.52
4714.01
6033.52
6930.31
6858.82
6948.16
6821.18
6653.13
7014.32
7406.83
8426.39
13852.46
13704.16
13555.94
13448.78
13283.30
13134.46
12641.50
12029.54
13960.63
1000.0
4000.2
7000.4
10000.6
13000.8
16001.0


T9
7097.44
6961.86
3553.82
1004.37
1057.44
1162.74
1238.55
1310.79
1520.76
1546.40
2373.32
2498.56
3483.10
3548.69
4645.78
4599.03
5614.61
6984.97
6888.63
6687.72
6787.53
7042.77
6640.91
7428.41
9589.54
13915.56
13765.27
13864.33
% Intensity
Mass (m/z)
1.1E+4
1000.0
4000.2
7000.4
10000.6
13000.8
16001.0


T1
% Intensity
Mass (m/z)
3495.6
7073.66
6923.27
3541.57
13890.06
1012.56
1067.14
1107.47
1149.47
1209.11
1259.76
1305.73
1504.74
1609.02
1693.31
1765.01
1918.61
2812.64
3495.02
4634.07
4602.71
6058.68
7124.68
7412.90
7446.02
8194.02
9002.24
9951.50
11333.60
13487.42
13208.80
14133.94
15102.96