Supporting Information

Synthesis and Conformational Analysis of D-2′-deoxy-2′,2′-difluoro-4′-dihydro-4′-thionucleosides

Feng Zheng[†], Lin Fu[‡], Xingang Zhang [†], Wang Renxiao[‡] and Feng-ling Qing[*,†]

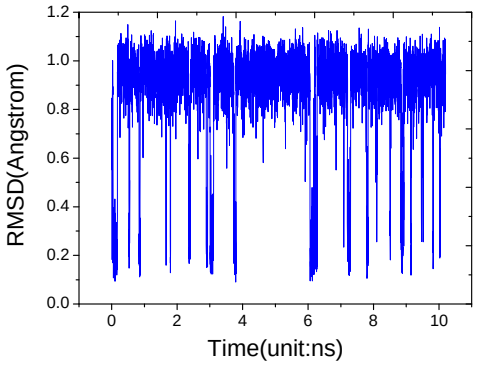
[†]Key Laboratory of Organofluorine Chemistry, Shanghai Institute of Organic Chemistry, Chinese Academy of Science, 354 Fengling Lu, Shanghai 200032, China

[‡] State Key Laboratory of Bioorganic and Natural Products Chemistry, Shanghai Institute of Organic Chemistry,Chinese Academy of Sciences,354 Fenglin Lu, Shanghai 200032, China

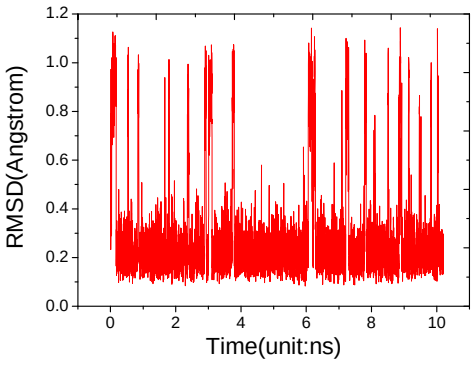
*Correspondence: flq@mail.sioc.ac.cn

Molecular dynamics simulation. Two molecular dynamics (MD) simulations were performed to study the *S*-conformer and the *N*-conformer in methanol, respectively. The electrostatic potentials of these two conformers were computed using the Gaussian 03 program at the B3LYP/6-311++G(2d,2p) level. Atomic partial charges were derived by the Restrained Electrostatic potential (RESP) method. The GAFF force field was applied to all computational jobs described below. To set up each MD job, the given conformer was soaked in a pre-optimized methanol box with a size of 20 Å × 20 Å × 20 Å with the periodic boundary condition. The entire system was then minimized at a stepwise manner. (1) 5000 steps of minimization with restraints on backbone (the harmonic force constant = 500.0 kcal/(mol·Å^2)) (2) 5000 steps of minimization with restraints on backbone (the harmonic force constant = 10.0 kcal/(mol·Å^2)). (3) 5000 steps of minimization without any restraints. At each round of minimization mentioned above, the steepest descent method was used for the first 500 cycles, and then the conjugated gradient method was used for the rest. To relax the entire system further, three rounds of position-restrained molecular dynamics simulations were conducted: (1) 30000 steps of MD simulation with restraints on backbone (the harmonic force constant = 5.0 kcal/(mol·Å^2)). (2) 40000 steps of MD simulation with restraints on backbone (the harmonic force constant = 0.5 kcal/(mol·Å^2)). (3) 50000 steps of MD simulation without restraint. After all of these preparations, a production run of 10 *ns* long was performed to sample the possible conformations of the given molecule in solution.

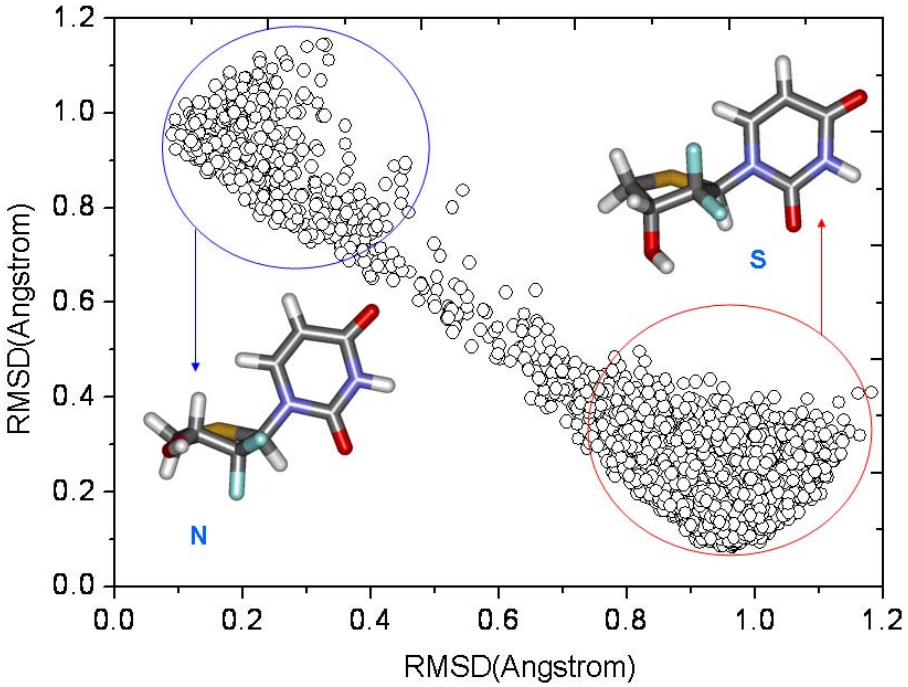
During the entire MD simulation, the time interval was set to 2 *fs*. Temperature (T = 300 *K*) was regulated by the Langevin thermostat with the collision frequency γ =2.0 ps^{-1}. Pressure (P = 1 *atm*) was controlled by the Berendsen barastat. A distance cutoff of 14 Å for non-bonded interactions was used. The Particle Mesh Ewald (PME) method was used with a grid spacing of 1Å combined with a fourth-order-B-spline interpolation to compute potentials and forces between grid points. The SHAKE algorithm was used to constrain all bonds involving hydrogen atoms. The conformational changes of the given molecule, quantified by the root-mean-square deviation (RMSD) values computed using the starting structure as reference, were monitored along the entire MD trajectory. Only the conformations sampled during the last 4 *ns* of MD simulation were considered in free energy calculation since the whole system was at good equilibrium at that stage.



(A)

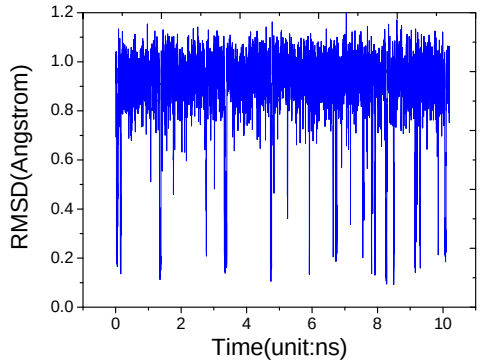


(B)

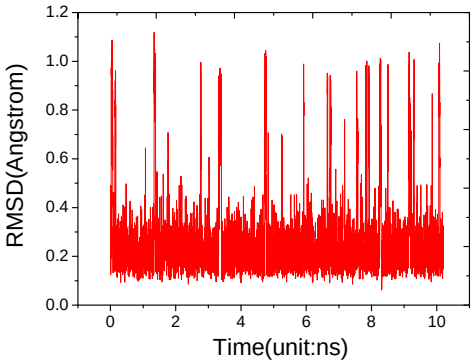


(C)

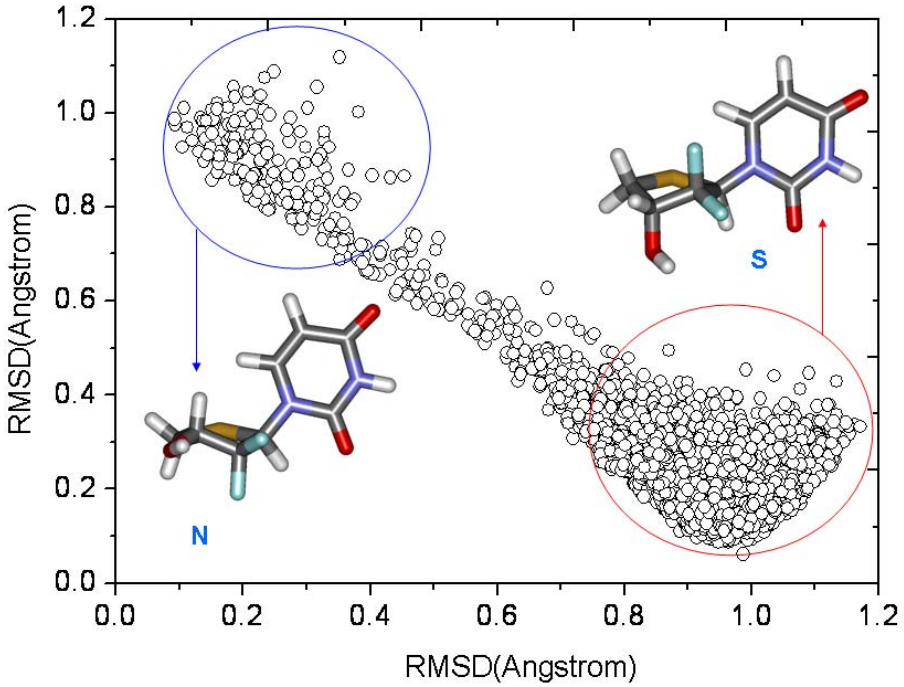
Figure 1. Conformational changes observed during the MD simulation in methanol using the *N*-conformer as the starting structure. (A) RMSD values on the MD trajectory computed using the *N*-conformer as reference. (B) RMSD values on the MD trajectory computed using the *S*-conformer as reference. (C) The *X* axis is the RMSD values computed using the *N*-conformer as reference; while the *Y* axis is the RMSD values computed using the *S*-conformer as reference. Apparently, conformations close to the *S*-conformer are much more populated during MD simulation.



(A)

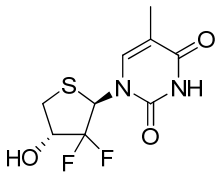


(B)



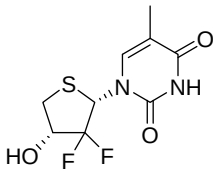
(C)

Figure 2. Conformational changes observed ′during the MD simulation in methanol using the *S*-conformer as the starting structure. (A) RMSD values on the MD trajectory computed using the *N*-conformer as reference. (B) RMSD values on the MD trajectory computed using the *S*-conformer as reference. (C) The *X* axis is the RMSD values computed using the *N*-conformer as reference; while the *Y* axis is the RMSD values computed using the *S*-conformer as reference. Apparently, conformations close to the *S*-conformer are much more populated during MD simulation.



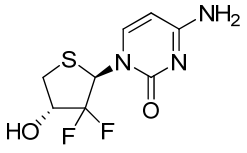
1-((2*R*,4*S*)-3,3-difluoro-4-hydroxytetrahydrothiophen–2-yl)thymine (5b)

Conversion of compound **9** (119 mg, 0.52 mmol) to **5b** and **5b′** was accomplished using the same procedure as described for compound **5a**. Compound **5b** (34 mg, 26% yield for three steps) was obtained as a white solid: $[\alpha]^{26}_{D}$ = -44.4 ˚(*c* 0.90 MeOH); ^{1}H NMR (300 MHz, MeOH-d_4) δ 7.79 (dd, *J* =2.7, 1.5 Hz, 1H), 6.50 (dd, *J* = 12.3, 9.6 Hz, 1H), 4.40 (m, 1H), 3.46 (m, 1H), 2.85 (m, 1H), 1.91 (d, *J* =0.9 Hz, 3H); ^{13}C NMR (75.5 MHz, MeOH-d_4) δ 165.9, 152.8, 139.1(d, *J* = 4.4 Hz), 126.4 (dd, *J* = 262.8, 255.8 Hz), 111.9, 72.9 (dd, *J* =30.6, 22.1 Hz), 59.5 (dd, *J* = 30.8, 18.9 Hz), 32.7, 12.4; ^{19}F NMR (282 MHz, MeOH-d_4) δ –119.5 (dd, *J* =236.6, 14.5 Hz, 1F), –123.2 (dt, *J* = 235.3, 7.9 Hz, 1F); IR (KBr) $_{max}$ 3211, 3057, 1697, 1466, 1377, 1079 cm^{-1}; MS (ESI) *m/z* 265.0 (M^++H), HRMS Calcd for $C_9H_{11}O_3N_2F_2S$ (M^+ + H): 265.0449. Found: 265.0453.



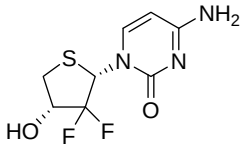
1-((2*S*,4*S*)-3,3-difluoro-4-hydroxytetrahydrothiophen–2-yl)thymine (5b′)

Compound **5b′** (19 mg, 26% yield for three steps) was obtained as a white solid : $[\alpha]^{26}_{D}$ = 17.8 ˚(*c* 0.85 MeOH); ^{1}H NMR (300 MHz, MeOH-d_4) δ 8.03 (s, 1H), 6.37 (dd, *J* = 13.8, 4.8 Hz, 1H), 4.43 (m, 1H), 3.18 (m, 1H), 3.10 (m, 1H), 1.90 (s, 3H); ^{13}C NMR (75.5 MHz, MeOH-d_4) δ 166.0, 153.1, 140.3, 126.6 (dd, *J* = 265.2, 254.7 Hz), 110.9, 73.1 (dd, *J* =31.0, 22.1 Hz), 62.1 (dd, *J* = 38.9, 20.6 Hz), 33.5, 12.5; ^{19}F NMR (282 MHz, MeOH-d_4) δ –108.3 (dd, *J* =247.6, 7.6 Hz, 1F), –124.0 (d, *J* = 244.5 Hz, 1F); IR (KBr) $_{max}$ 3213, 3074, 1693, 1466, 1389, 1083 cm^{-1}; MS (ESI) *m/z* 265.0 (M^++H), HRMS Calcd for $C_9H_{11}O_3N_2F_2S$ (M^+ + H): 265.0446. Found: 265.0446.



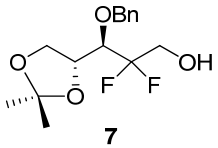
1-((2*R*,4*S*)-3,3-difluoro-4-hydroxytetrahydrothiophen–2-yl)cytosine (5c) Conversion of compound **9** (131 mg, 0.57 mmol) to **5c** and **5c′**was accomplished using the similar procedure as described for compound **5a.** After removal of the benzyl group, the benzoyl group was removed using NH_3/CH_3OH**.** Compound **5c** (29 mg, 21% yield for four steps) was

obtained as a white solid: $[\alpha]^{26}_{D}$ = -32.7 ˚(*c* 0.90 MeOH); ^{1}H NMR (300 MHz, MeOH-d_4) δ 8.00 (dd, *J* = 7.5, 2.4 Hz, 1H), 6.67 (dd, *J* = 12.3, 9.6 Hz, 1H), 5.97 (d, *J* = 7.5 Hz, 1H), 4.38 (m, 1H), 3.43 (m, 1H), 2.85 (m, 1H); ^{13}C NMR (75.5 MHz, MeOH-d_4) δ 167.1, 158.5, 144.5, 126.4 (dd, *J* = 262.7, 256.3 Hz), 96.7, 73.1 (dd, *J* =30.5, 22.1 Hz), 60.1 (dd, *J* = 31.9, 19.5 Hz), 32.5; ^{19}F NMR (282 MHz, MeOH-d_4) δ –119.0 (dd, *J* =235.5, 11.8 Hz, 1F), –122.4 (dt, *J* = 234.6, 6.8 Hz, 1F); IR (KBr) $_{max}$ 3328, 3197, 1649, 1492, 1393, 1077 cm^{-1}; MS (ESI) *m/z* 250.0 (M^++H), HRMS Calcd for $C_8H_{10}O_2N_3F_2S$ (M^+ + H): 250.0451. Found: 250.0456.



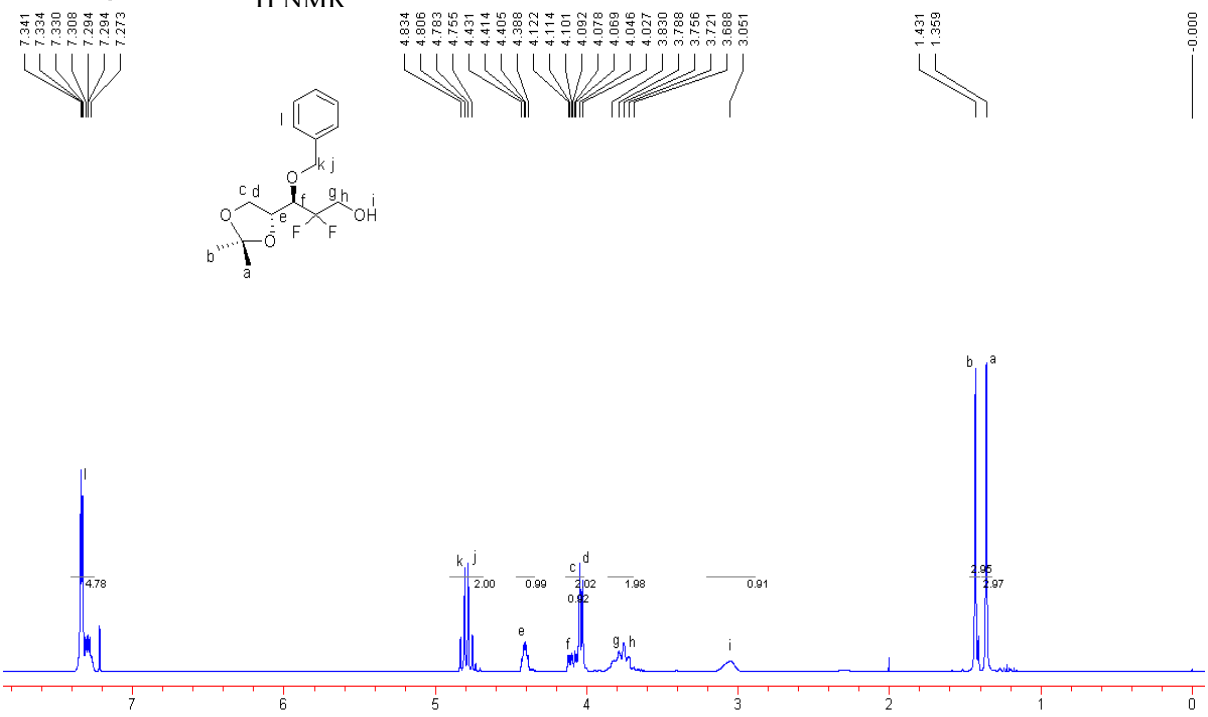
1-((2*S*,4*S*)-3,3-difluoro-4-hydroxytetrahydrothiophen–2-yl)cytosine (5c′)

Compound **5c′** (14 mg, 11% yield for three steps) was obtained as a white solid :$[\alpha]^{26}_{D}$ = 5.6 ˚(*c* 0.60 MeOH); ^{1}H NMR (300 MHz, MeOH-d_4) δ 8.19 (d, *J* = 8.1 Hz, 1H), 6.53 (dd, *J* = 13.8, 5.1 Hz, 1H), 5.92 (d, *J* = 8.1 Hz, 1H), 4.42 (m, 1H), 3.17 (m, 1H), 3.09 (m, 1H); ^{13}C NMR (75.5 MHz, MeOH-d_4) δ 167.5, 158.7, 145.5, 126.6 (t, *J* = 258.1 Hz), 96.0, 73.4 (dd, *J* =30.0, 17.1 Hz), 60.6 (dd, *J* = 38.9, 20.1 Hz), 33.4; ^{19}F NMR (282 MHz, MeOH-d_4) δ –108.5 (dd, *J* =237.2, 9.0 Hz, 1F), –124.3 (d, *J* = 238.0 Hz, 1F); IR (KBr) $_{max}$ 3320, 3190, 1650, 1493, 1398, 1053 cm^{-1}; MS (ESI) *m/z* 250.0 (M^++H), HRMS Calcd for $C_8H_{10}O_2N_3F_2S$ (M^+ + H): 250.0447. Found: 250.0456.

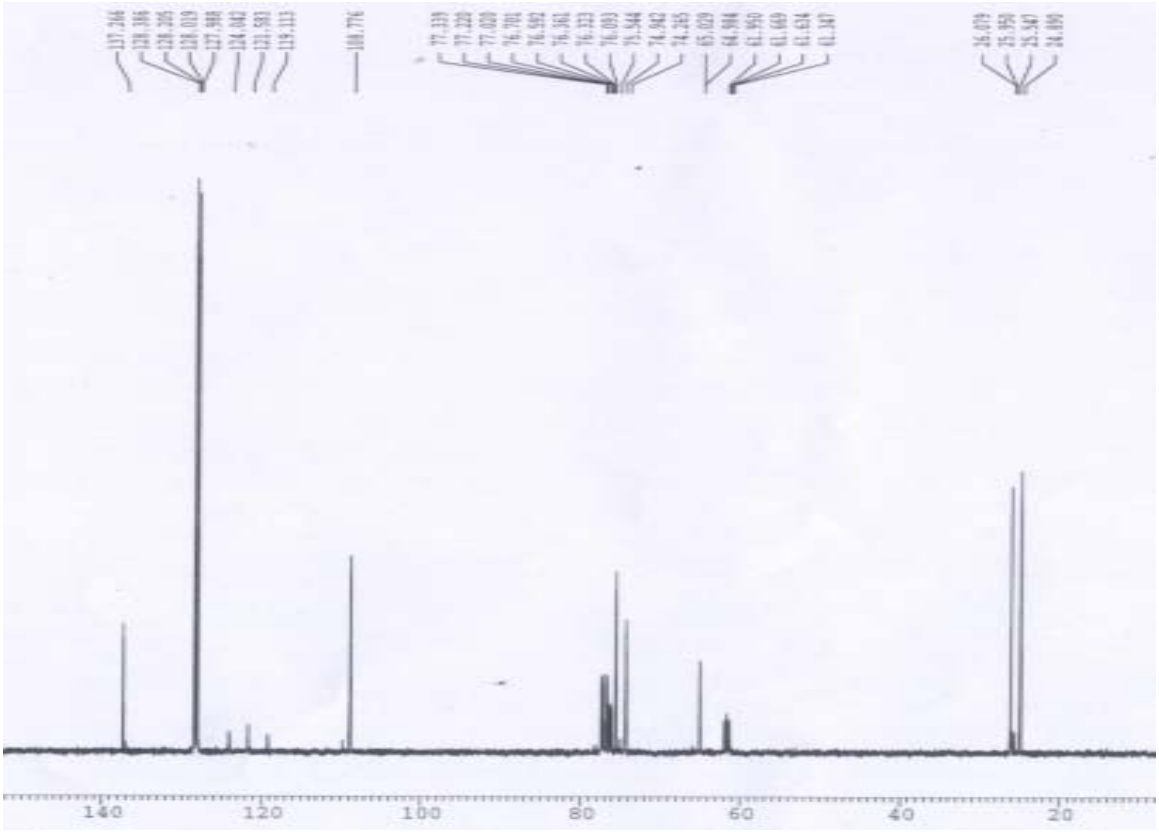


7

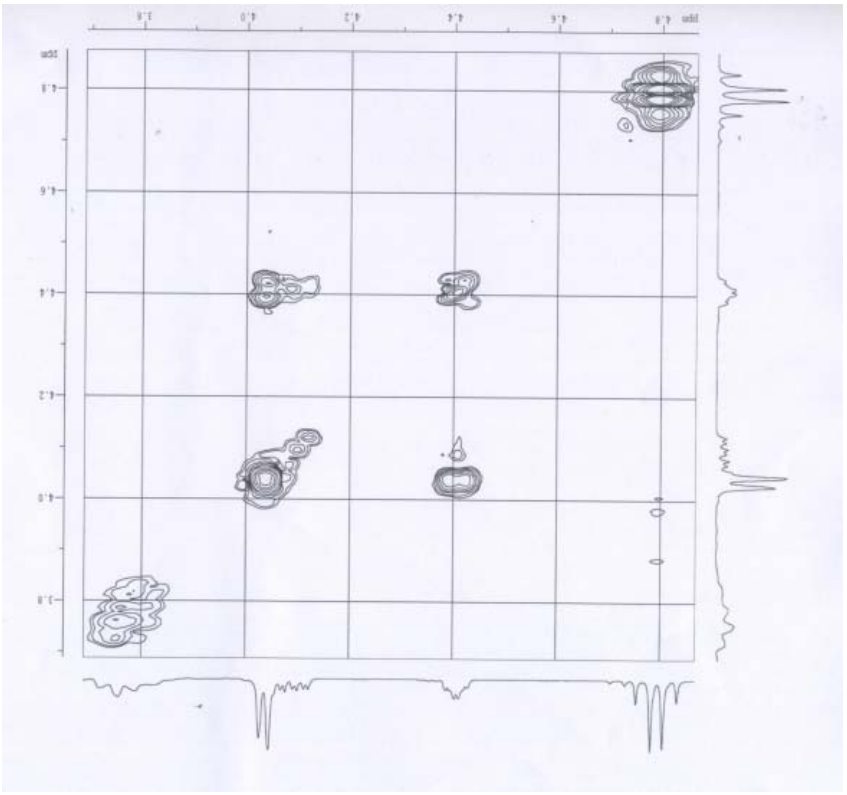
^{1}H NMR



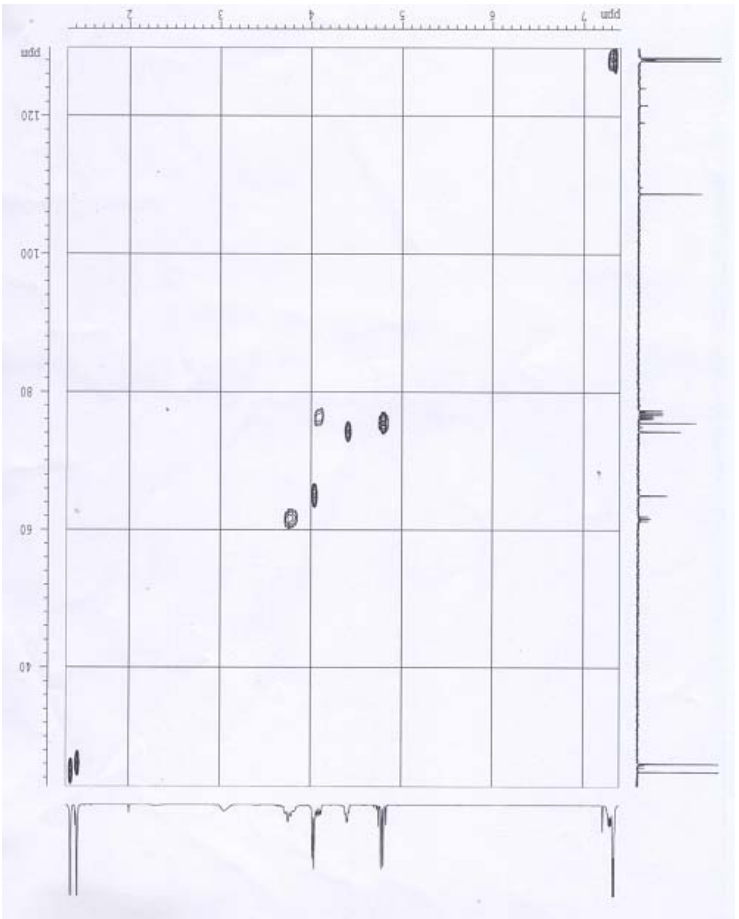
^{13}C NMR

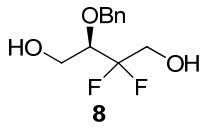


COSY



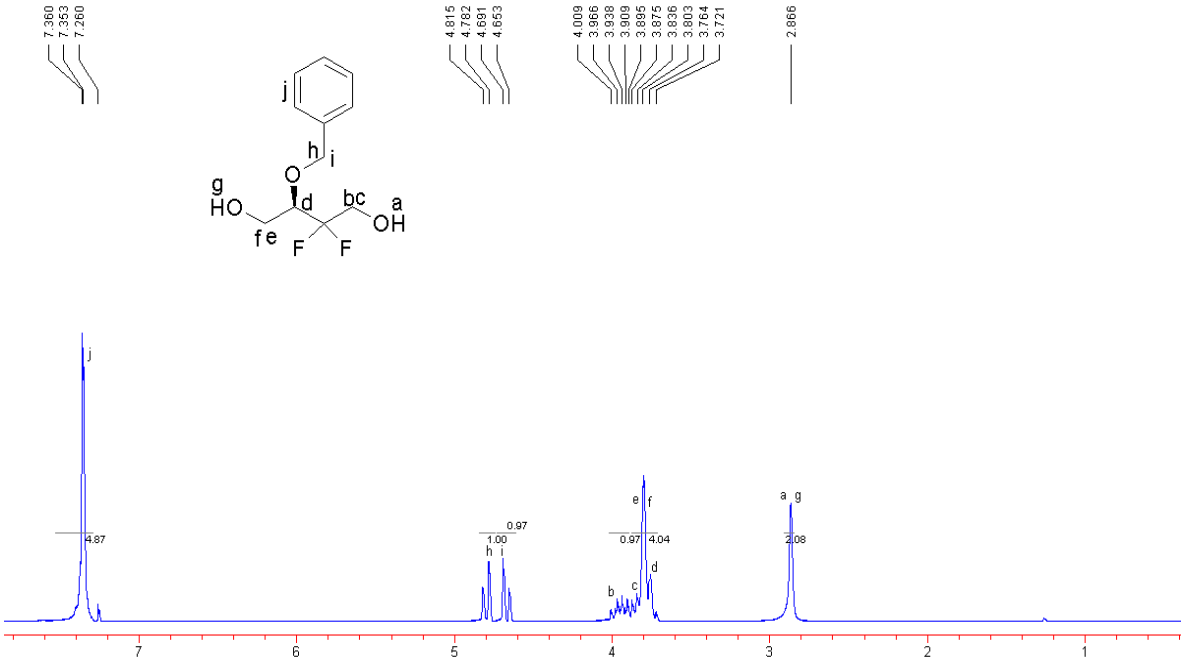
HSQC



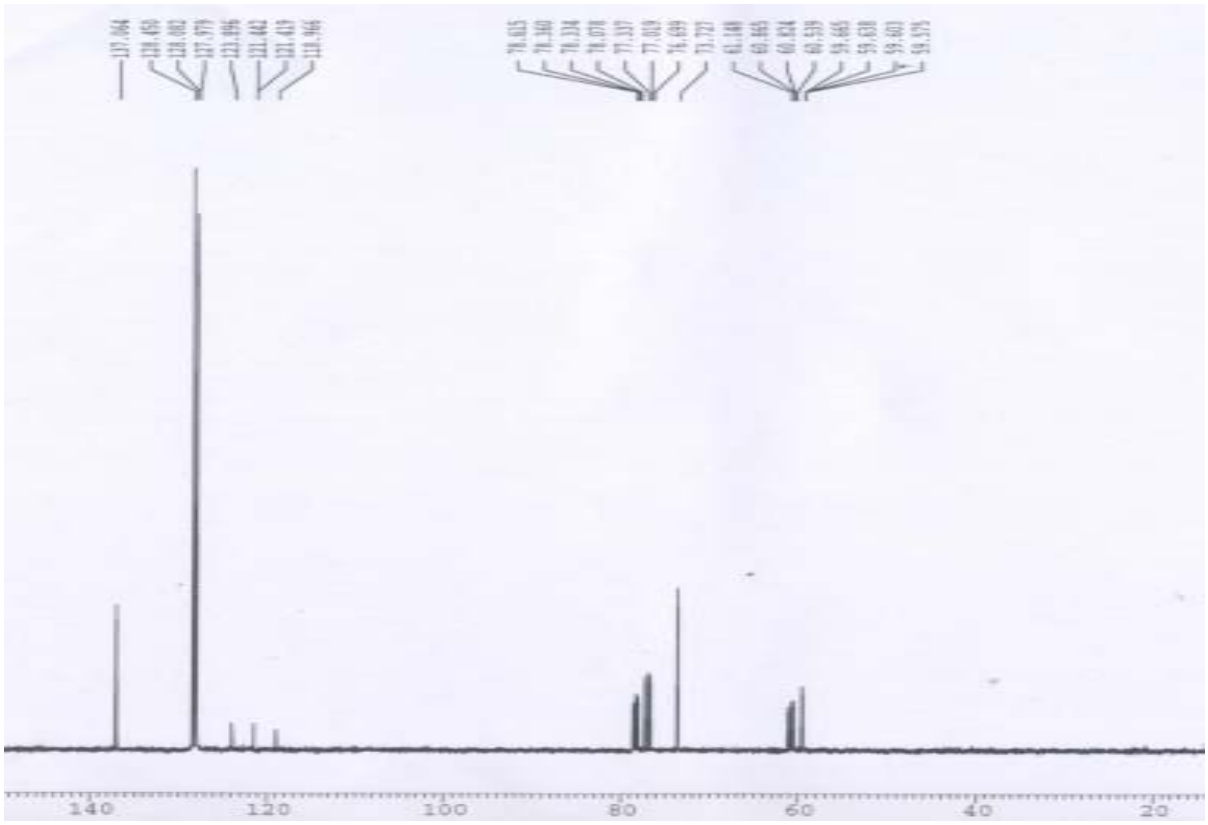


8

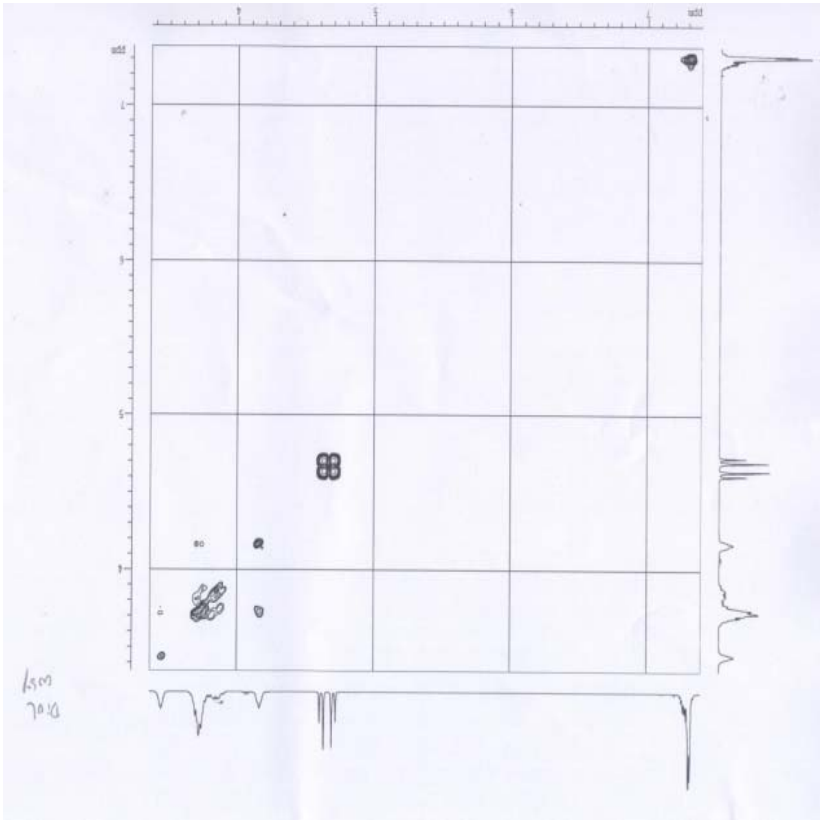
^{1}H NMR



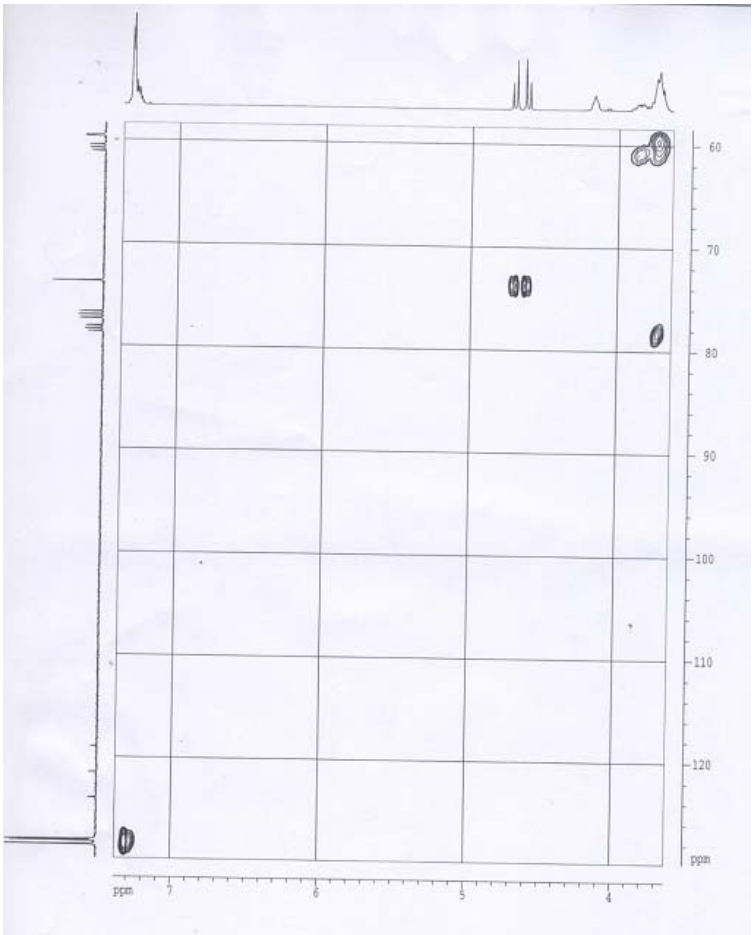
^{13}C NMR



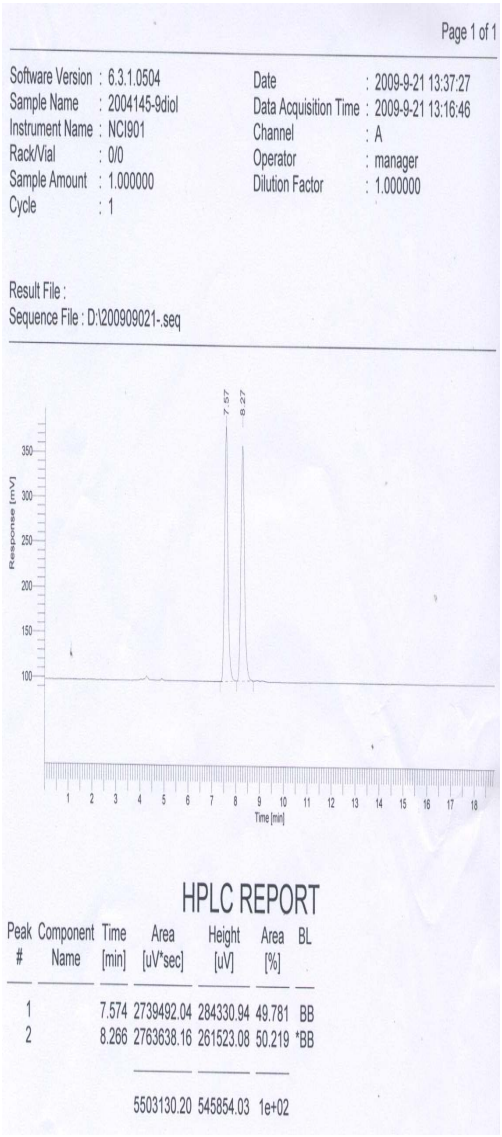
COSY



HSQC



Chiral HPLC analytical data of compound **8**



Page 1 of 1

Software Version : 6.3.1.0504
Sample Name : 2004145-9diol
Instrument Name : NCI901
Rack/Vial : 0/0
Sample Amount : 1.000000
Cycle : 1

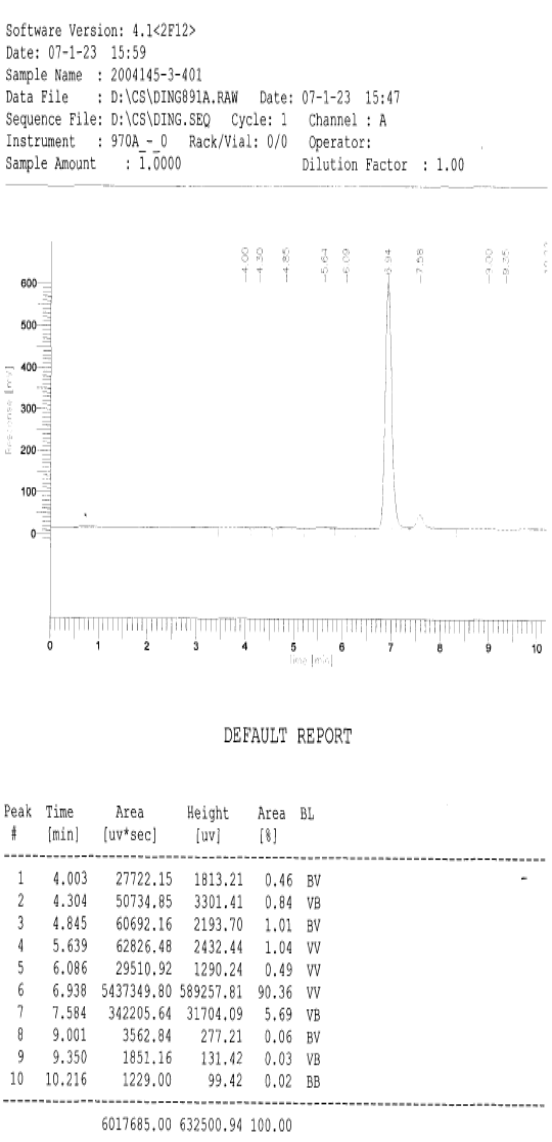
Date : 2009-9-21 13:37:27
Data Acquisition Time : 2009-9-21 13:16:46
Channel : A
Operator : manager
Dilution Factor : 1.000000

Result File :
Sequence File : D:\200909021-.seq

HPLC REPORT

Peak #	Component Name	Time [min]	Area [uV*sec]	Height [uV]	Area [%]	BL
1		7.574	2739492.04	284330.94	49.781	BB
2		8.266	2763638.16	261523.08	50.219	*BB
			5503130.20	545854.03	1e+02	

racemic

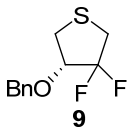


Software Version: 4.1<2F12>
Date: 07-1-23 15:59
Sample Name : 2004145-3-401
Data File : D:\CS\DING891A.RAW Date: 07-1-23 15:47
Sequence File: D:\CS\DING.SEQ Cycle: 1 Channel : A
Instrument : 970A - 0 Rack/Vial: 0/0 Operator:
Sample Amount : 1.0000 Dilution Factor : 1.00

DEFAULT REPORT

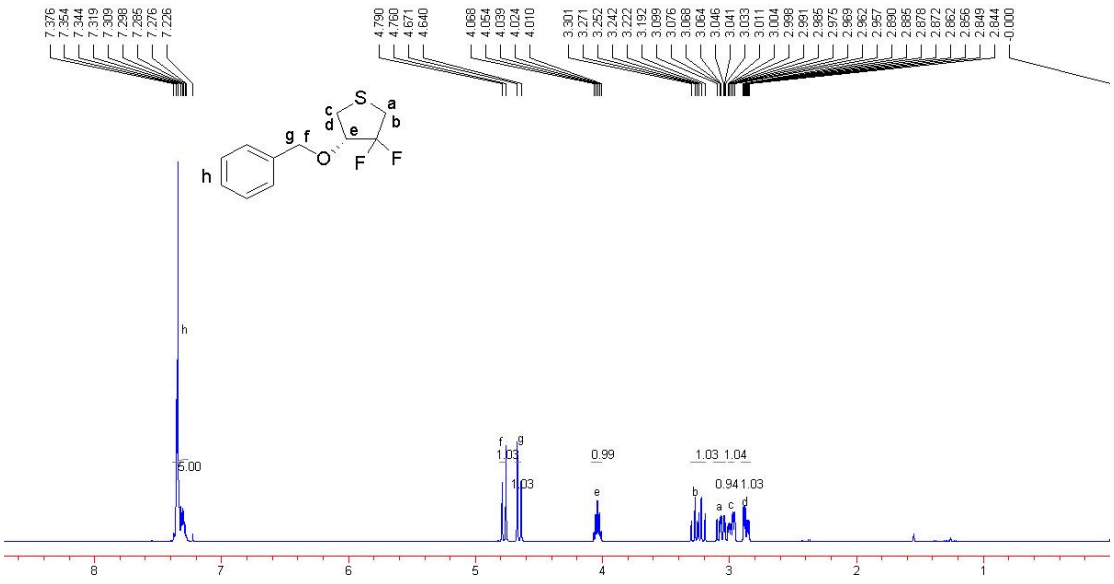
Peak #	Time [min]	Area [uv*sec]	Height [uv]	Area [%]	BL
1	4.003	27722.15	1813.21	0.46	BV
2	4.304	50734.85	3301.41	0.84	VB
3	4.845	60692.16	2193.70	1.01	BV
4	5.639	62826.48	2432.44	1.04	VV
5	6.086	29510.92	1290.24	0.49	VV
6	6.938	5437349.80	589257.81	90.36	VV
7	7.584	342205.64	31704.09	5.69	VB
8	9.001	3562.84	277.21	0.06	BV
9	9.350	1851.16	131.42	0.03	VB
10	10.216	1229.00	99.42	0.02	BB
		6017685.00	632500.94	100.00	

88% ee

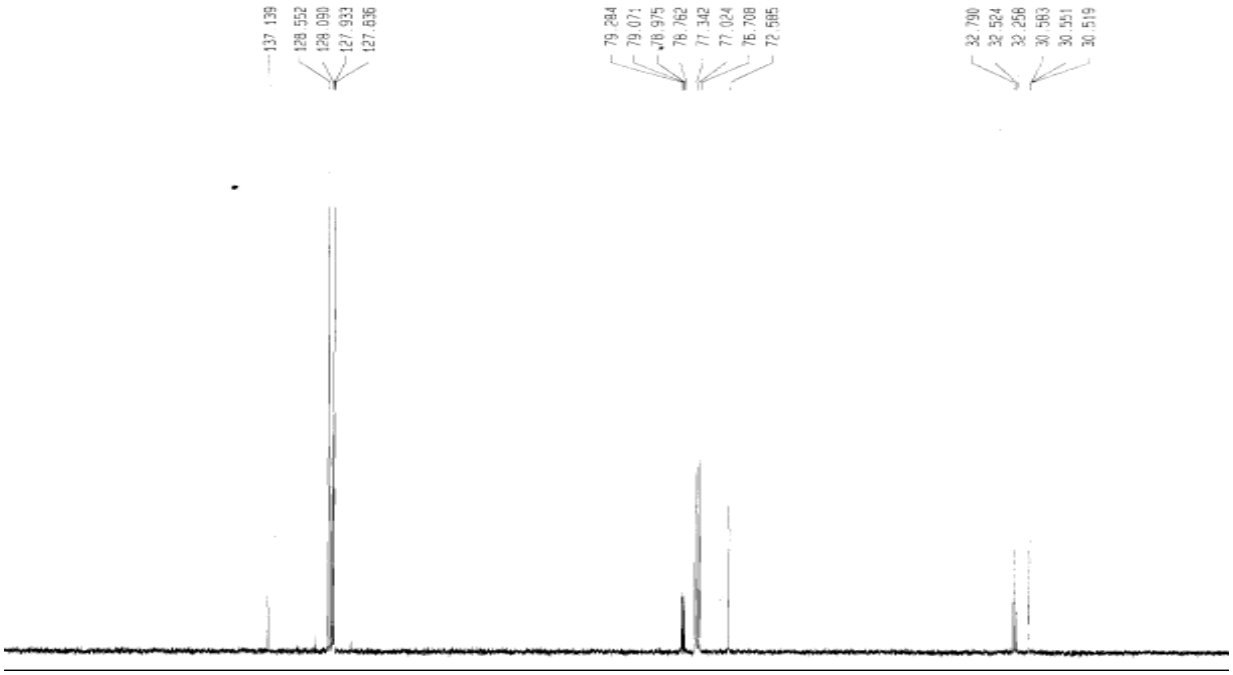


9

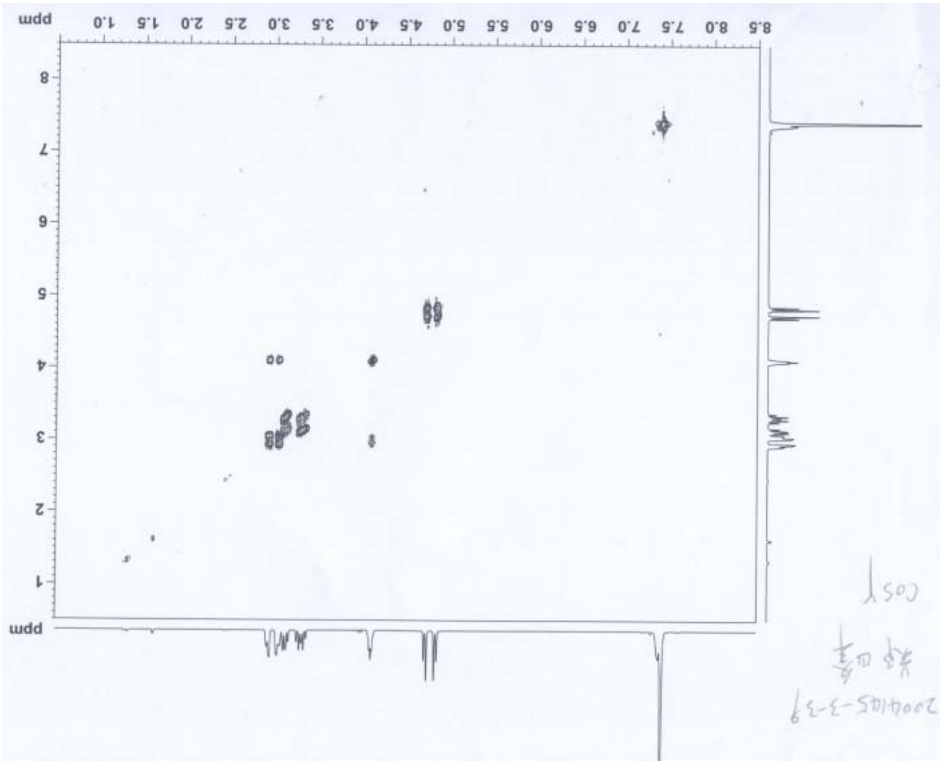
^{1}H NMR



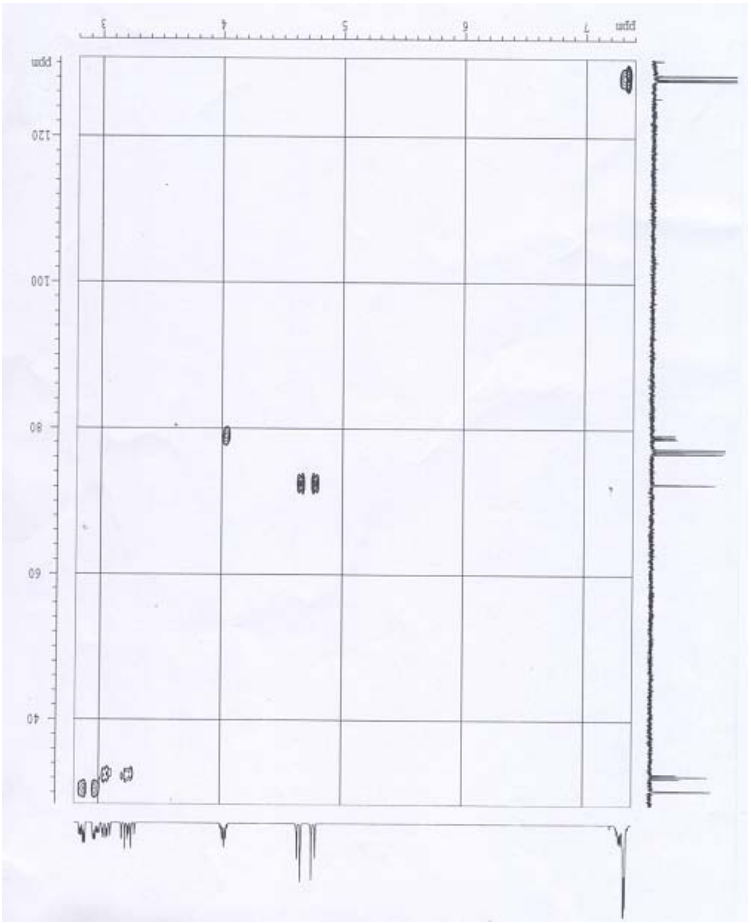
^{13}C NMR

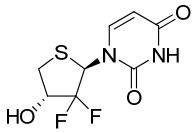


COSY



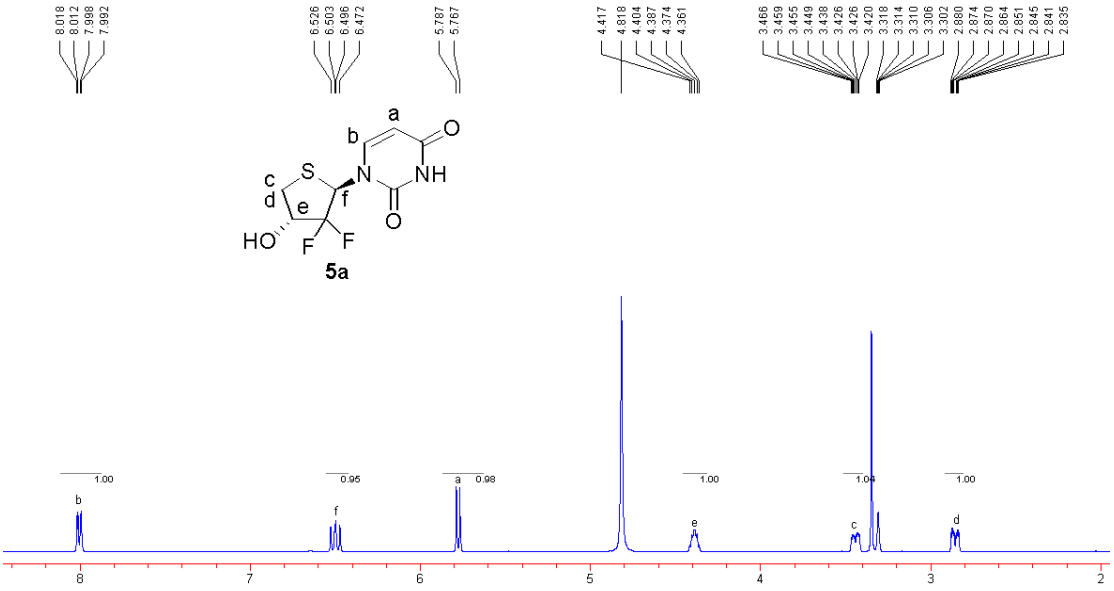
HSQC





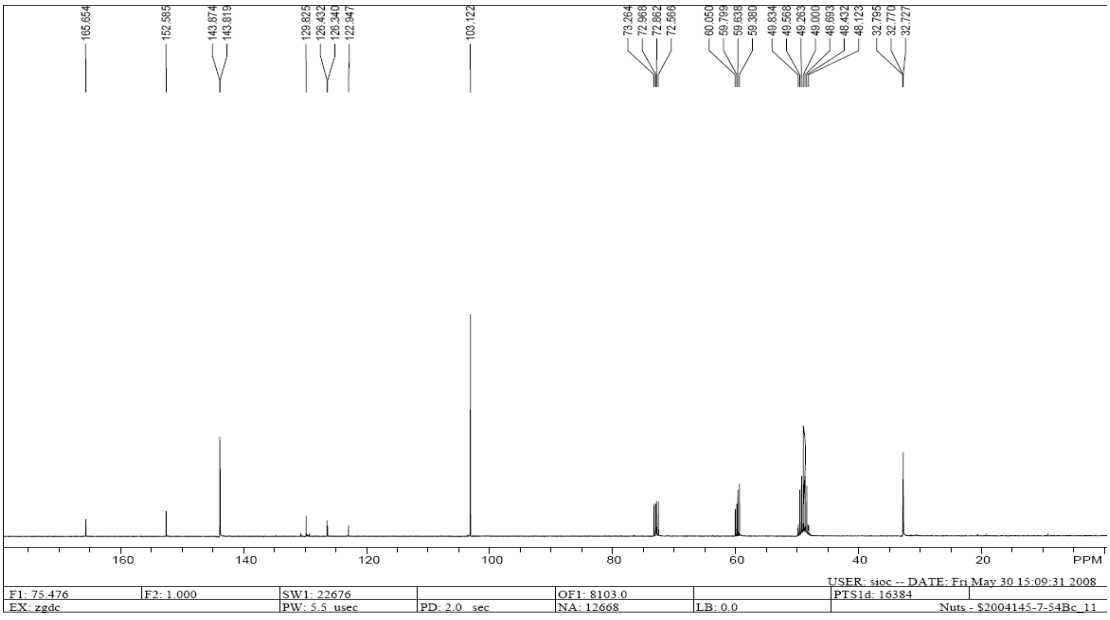
5a

^{1}H NMR

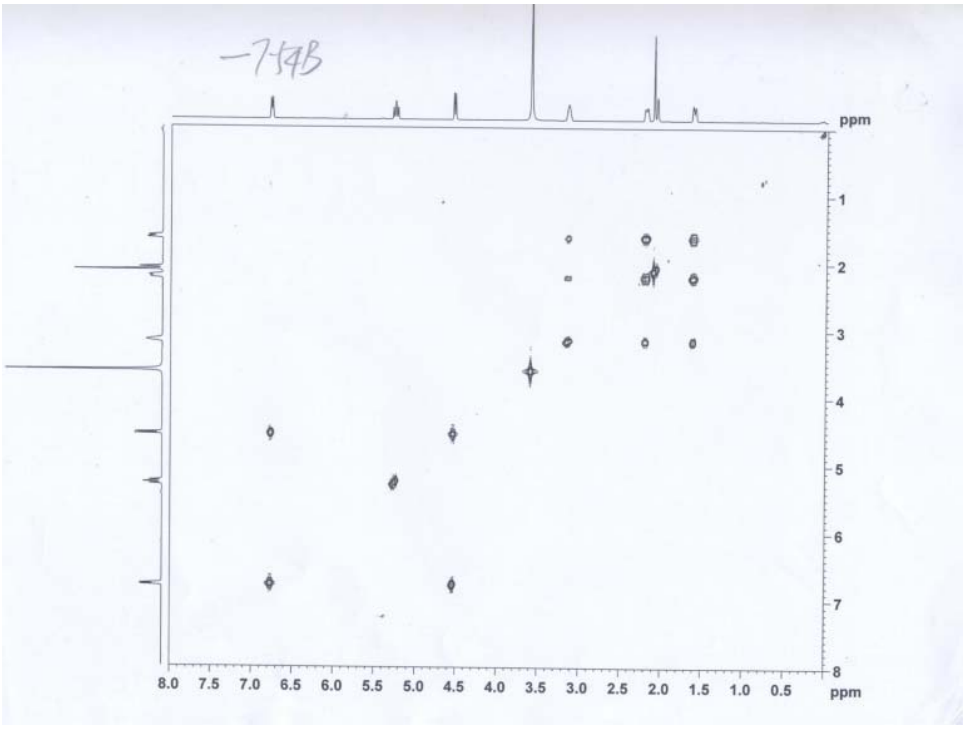


5a

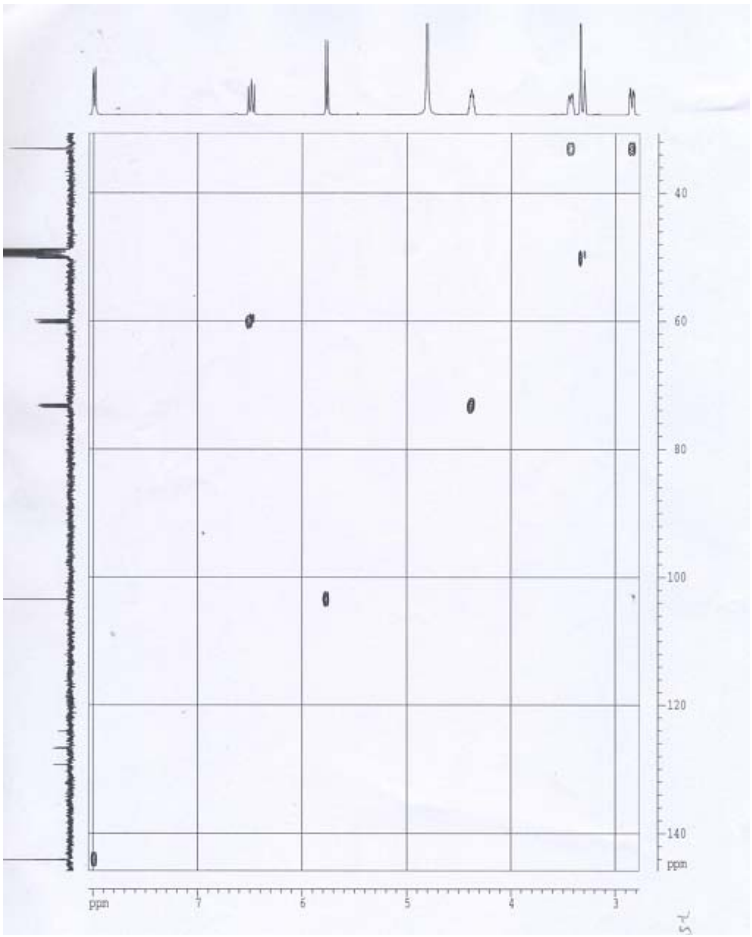
^{13}C NMR



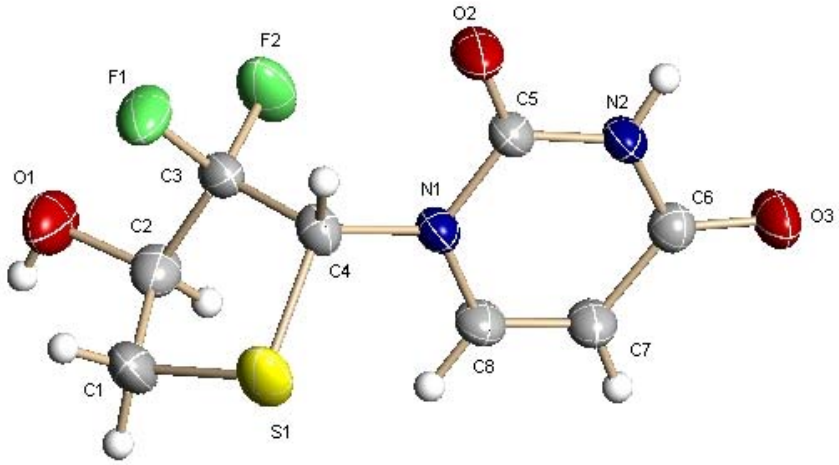
COSY


ppm
1
2
3
4
5
6
7
8
8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5
ppm

HSQC

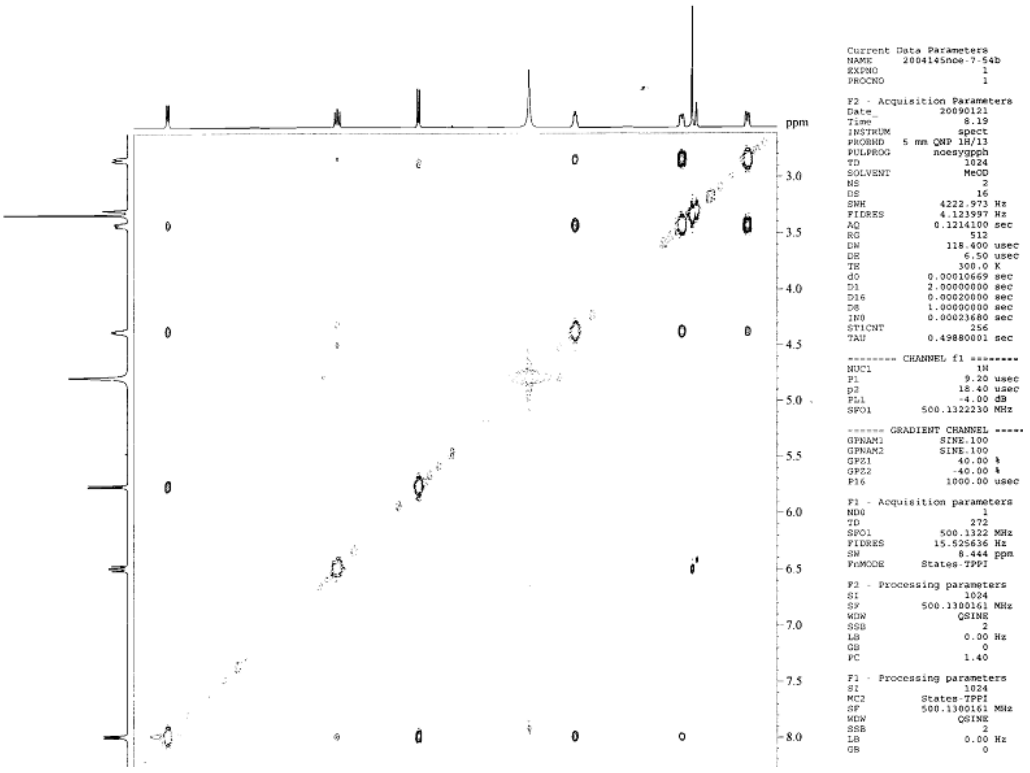

40
60
80
100
120
140
ppm
ppm
7
6
5
4
3

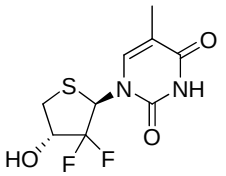
X-ray crystal structure of compound **5a**



Crystal data: for **5a**, $C_8H_8F_2N_2O_3S$, M = 250.22, Monoclinic, a = 6.1849(8), b = 18.082(2), c = 8.9343(11) Å, α=90, β=100.122(2), γ=90 deg., V = 983.6(2) Å^3, T = 293K, space group $P2(1)$, Z=4, 5594 reflections collected, 2129 unique (R_{int} = 0.0244), R_1 = 0.0433[I>2δ(I)], wR_2 = 0.1110.

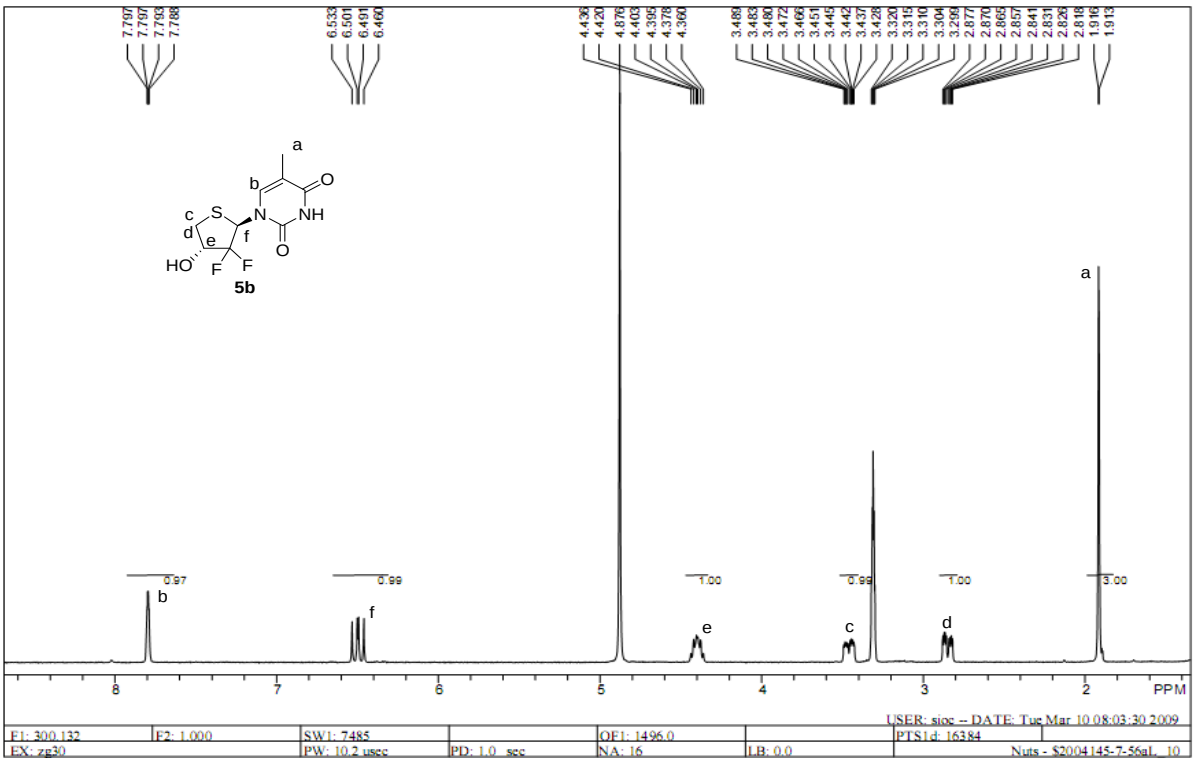
NOESY NMR spectrum of compound **5a**



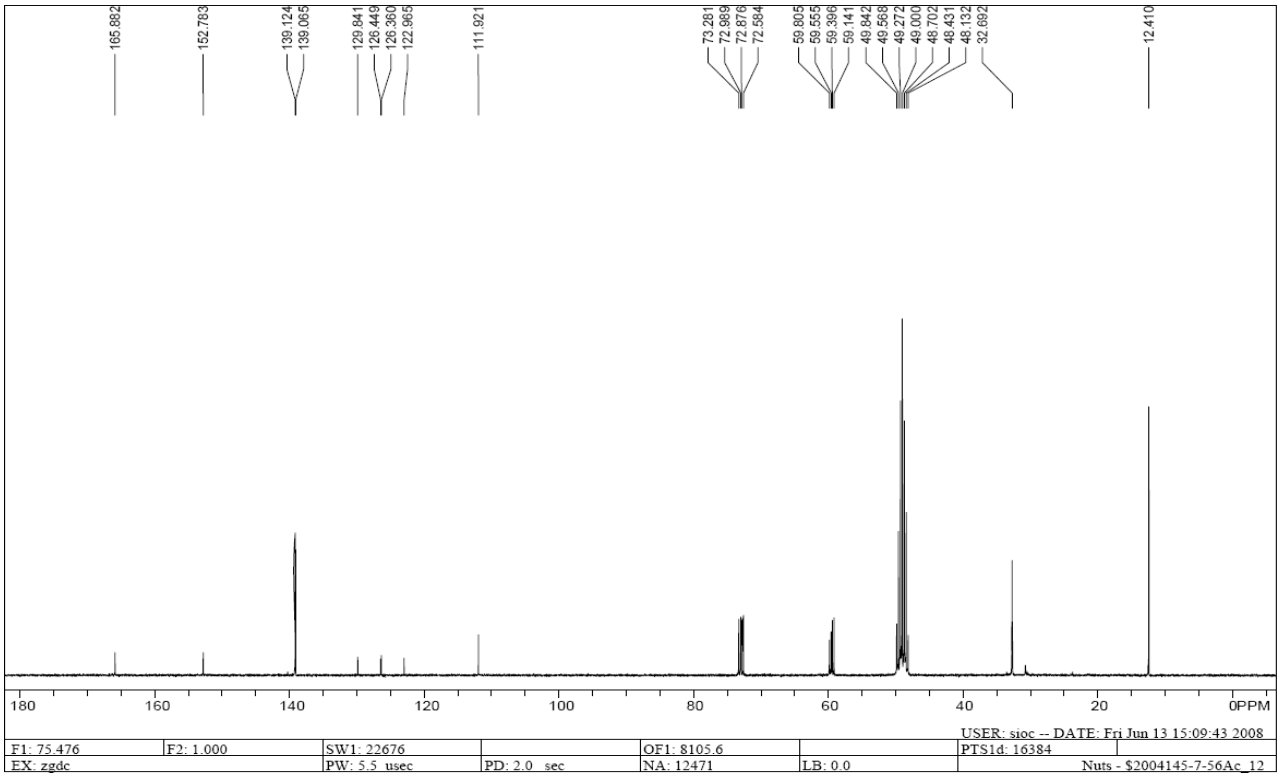


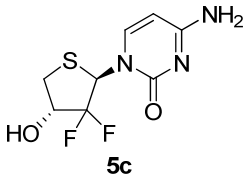
5b

^{1}H NMR



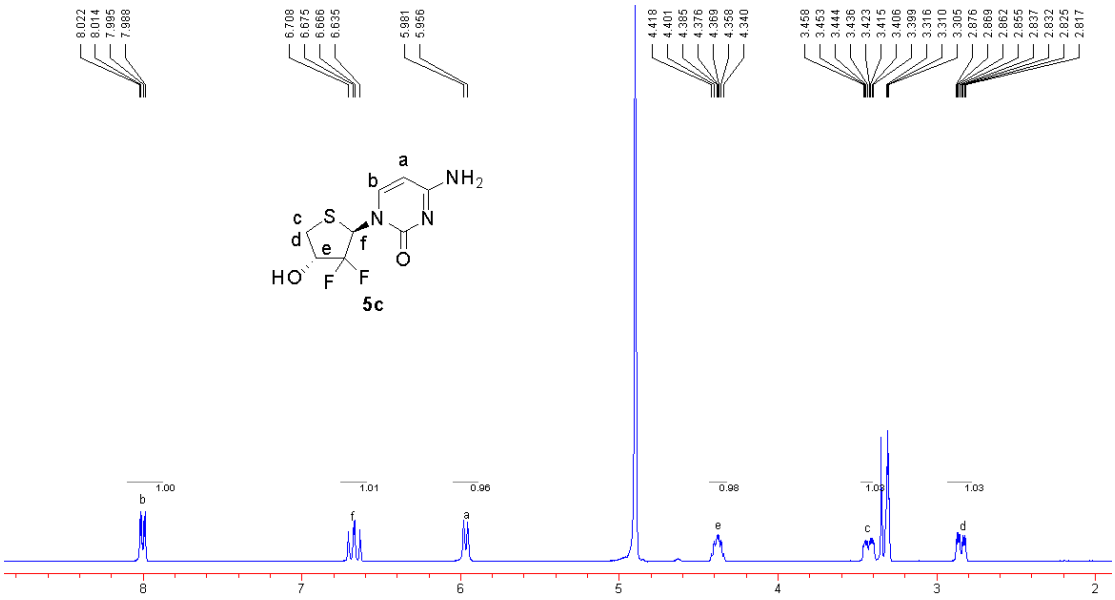
^{13}C NMR



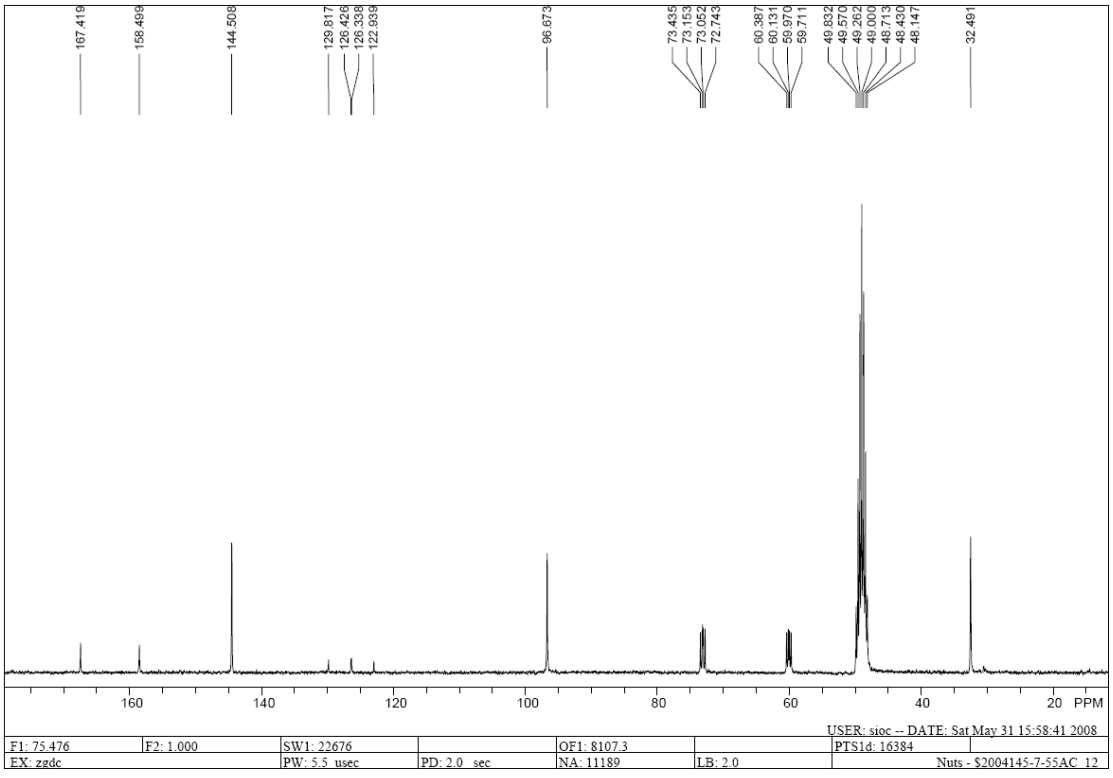


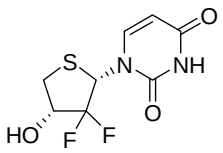
5c

^{1}H NMR



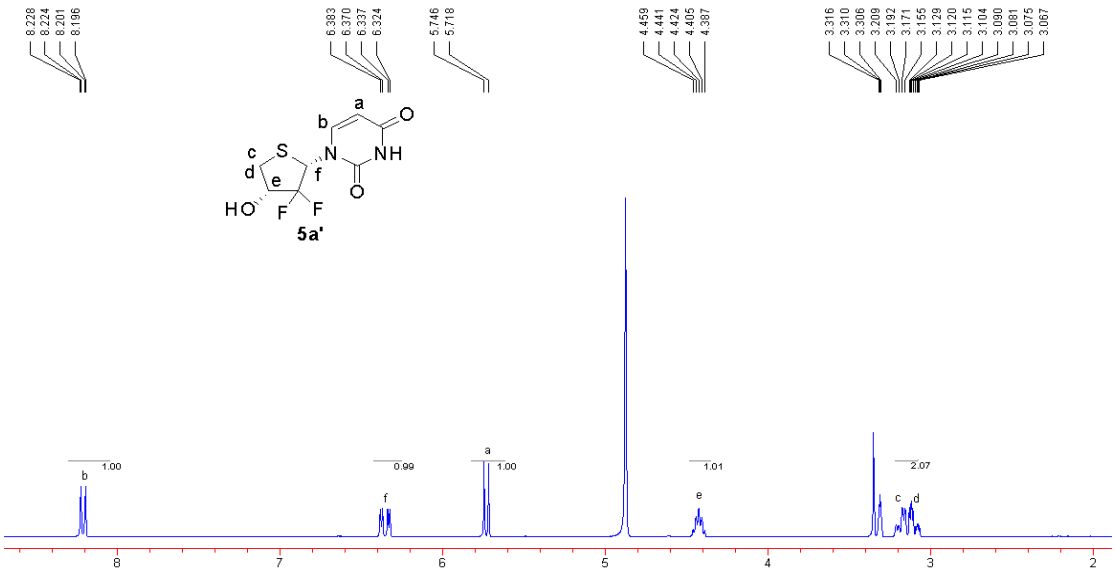
^{13}C NMR



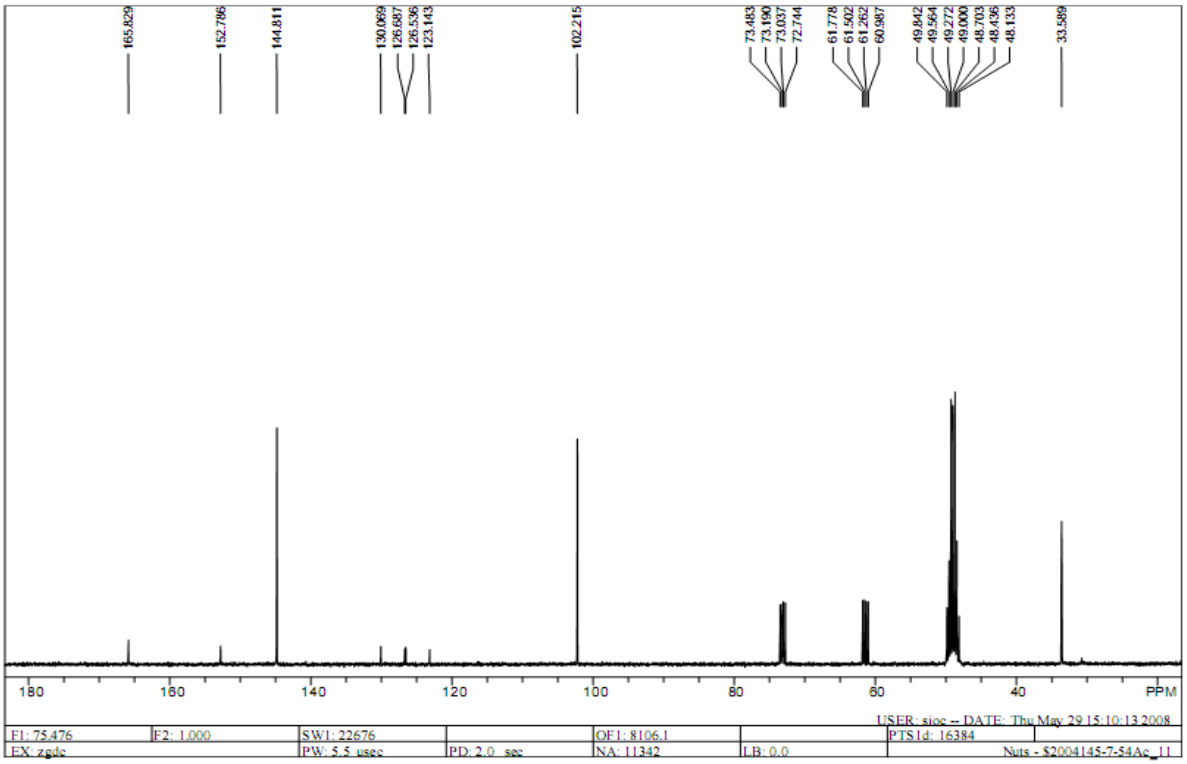


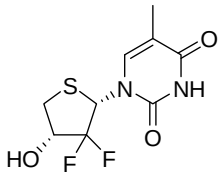
5a'

^{1}H NMR



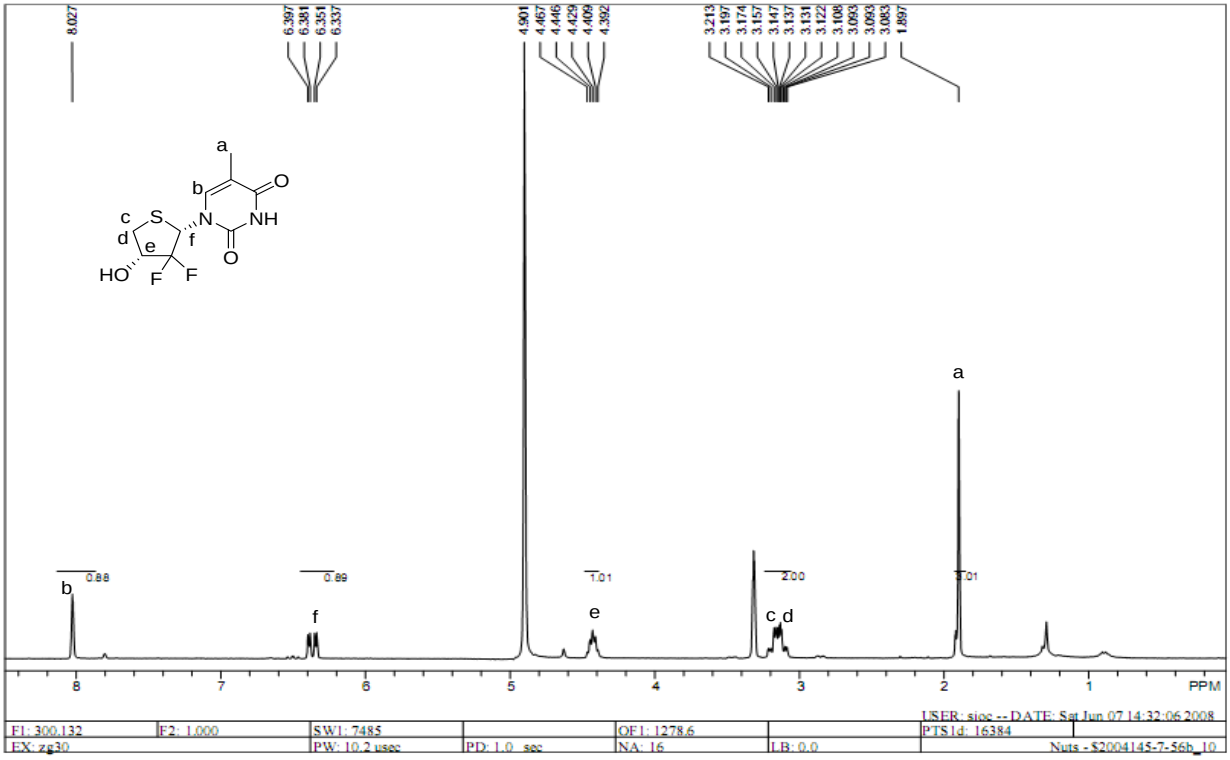
^{13}C NMR



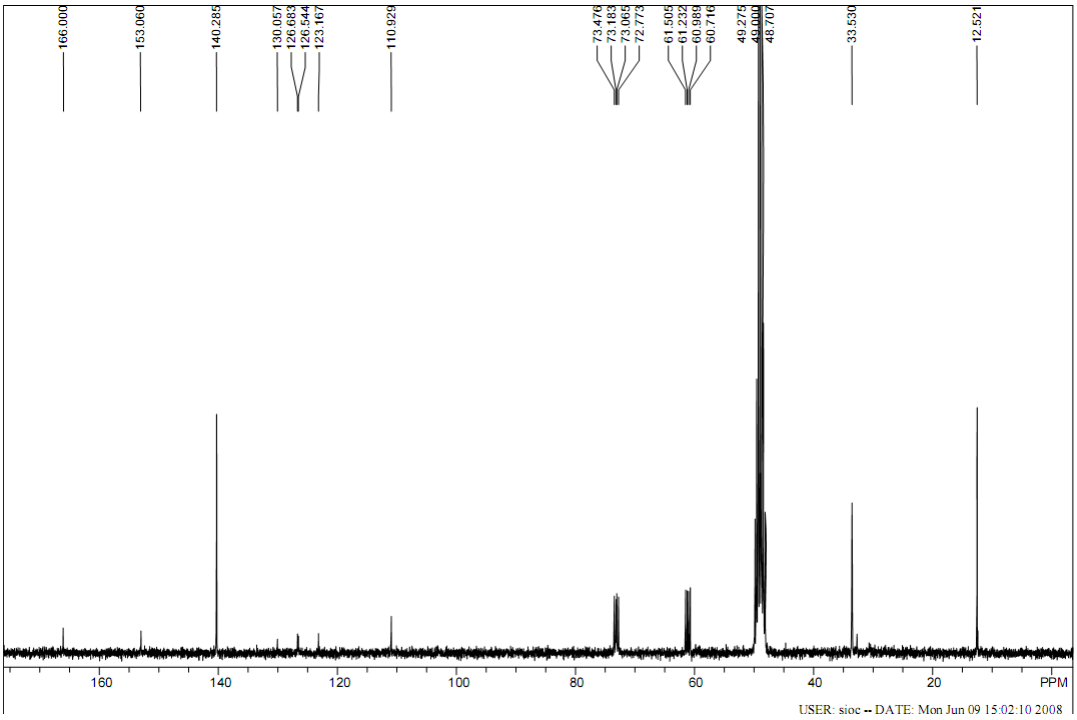


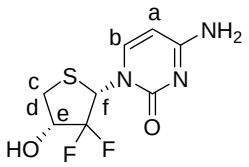
5b'

^{1}H NMR

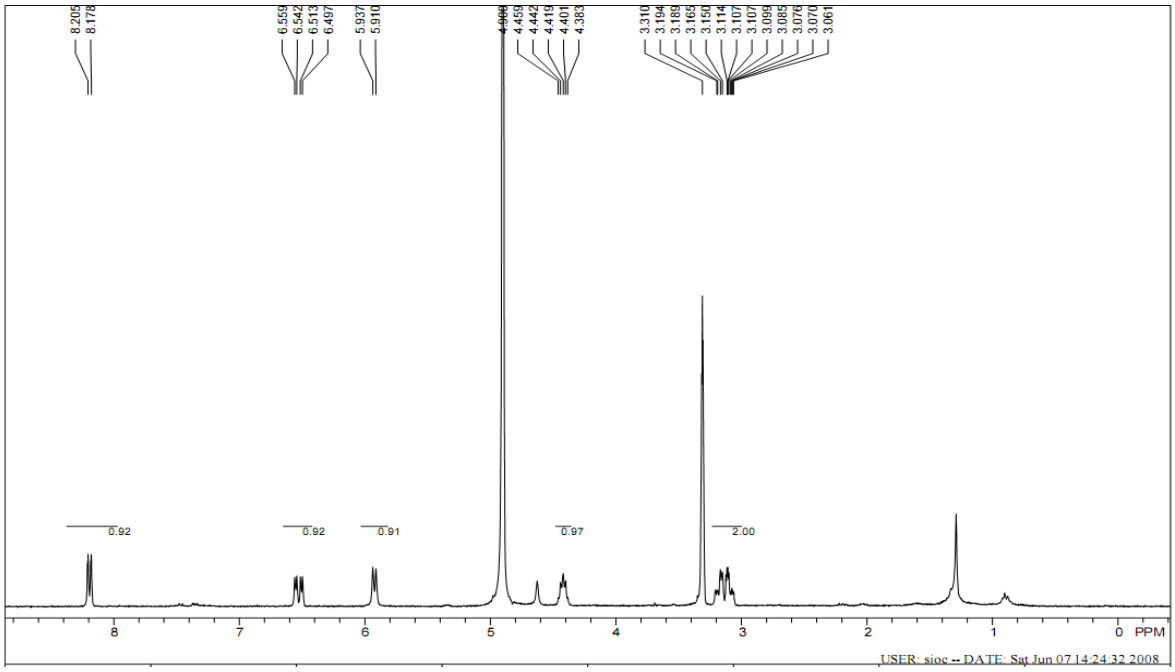


^{13}C NMR

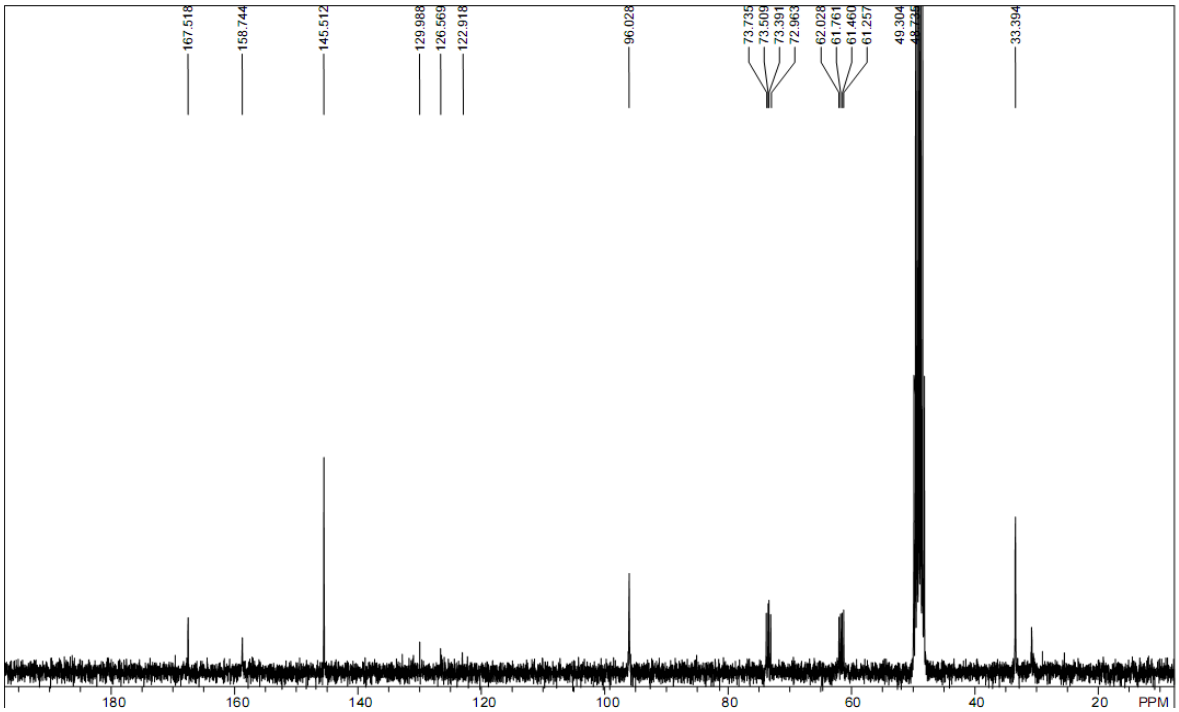



a
b
NH2
c
S
N
N
d
e
f
HO
F
F
O

5c'


8.205
8.178
6.559
6.542
6.513
6.497
5.937
5.910
4.459
4.442
4.419
4.401
4.383
3.310
3.194
3.189
3.165
3.150
3.114
3.107
3.107
3.099
3.085
3.076
3.070
3.061
0.92
0.92
0.91
0.97
2.00
8
7
6
5
4
3
2
1
0 PPM
USER: sioc -- DATE: Sat Jun 07 14:24:32 2008

^{13}C NMR


167.518
158.744
145.512
129.988
126.569
122.918
96.028
73.735
73.509
73.391
72.963
62.028
61.761
61.460
61.257
49.304
33.394
180
160
140
120
100
80
60
40
20
PPM