Synthesis and Properties of 2′-Deoxy-2′,4′-Difluoroarabinose Modified Nucleic Acids (2′,4′-diFANA)

Saúl Martínez-Montero,[a] Glen F. Deleavey,[a] Arden Dierker-Viik,[a] Petra Lindovska,[a] Tatiana Ilina,[b] Guillem Portella,[c] Modesto Orozco,[c] Michael A. Parniak,[b] Carlos González,[d] and Masad J. Damha[a]*

[a]Department of Chemistry, McGill University, 801 Sherbrooke Street West, Montreal, QC, Canada H3A 0B8.

[b]Department of Microbiology and Molecular Genetics, School of Medicine, University of Pittsburgh, 450 Technology Drive, Pittsburgh, PA, USA.

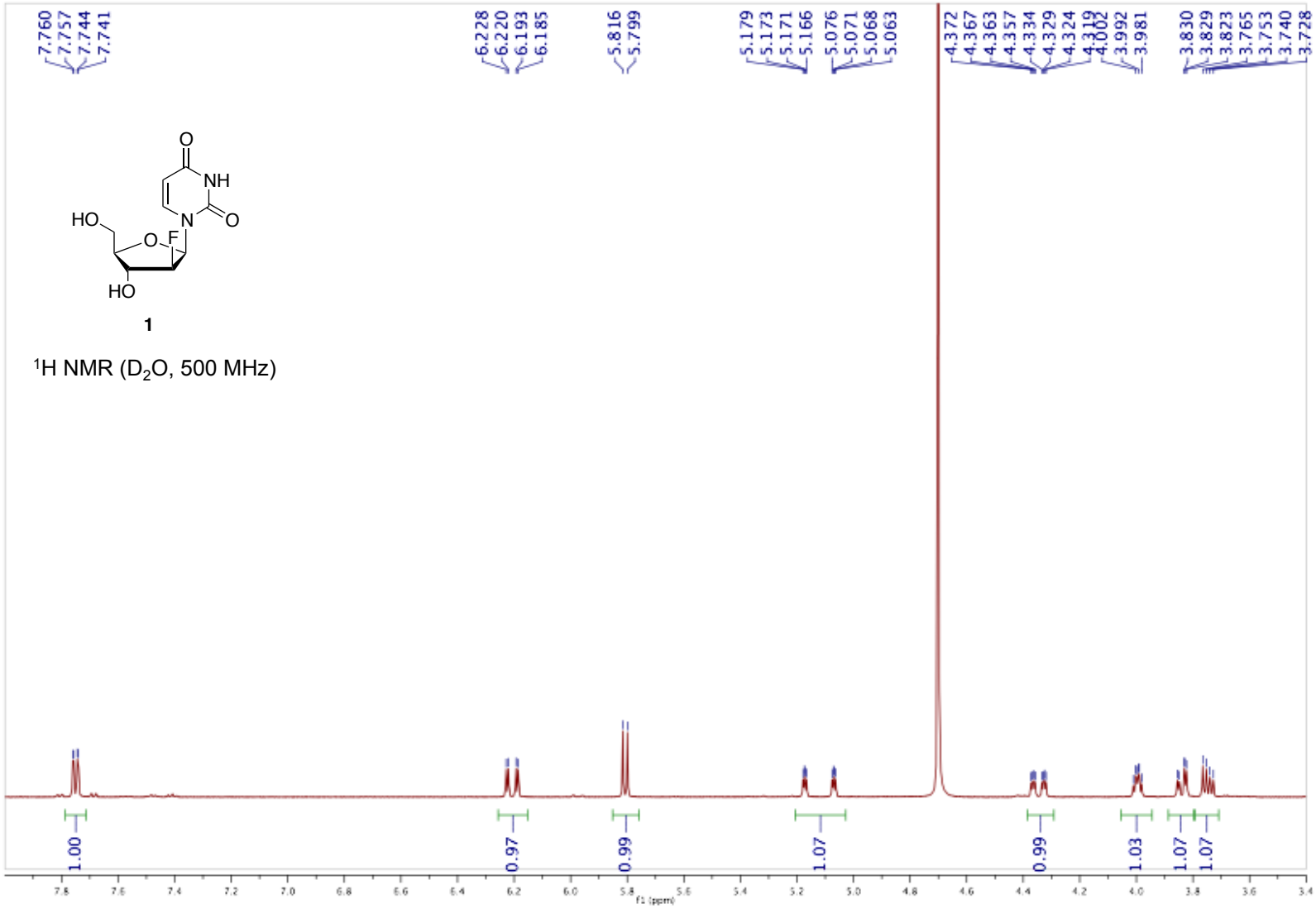
[c]Joint IRB-BSC program on Computational Biology, Institute for Research in Biomedicine, Barcelona Supercomputing Center, and Department of Biochemistry, University of Barcelona Baldir i Reixac 10–12, 08028 Barcelona, Spain.

[d]Instituto de Química Física Rocasolano, CSIC, C/. Serrano 119, 28006 Madrid, Spain.

*Correspondence to: masad.damha@mcgill.ca

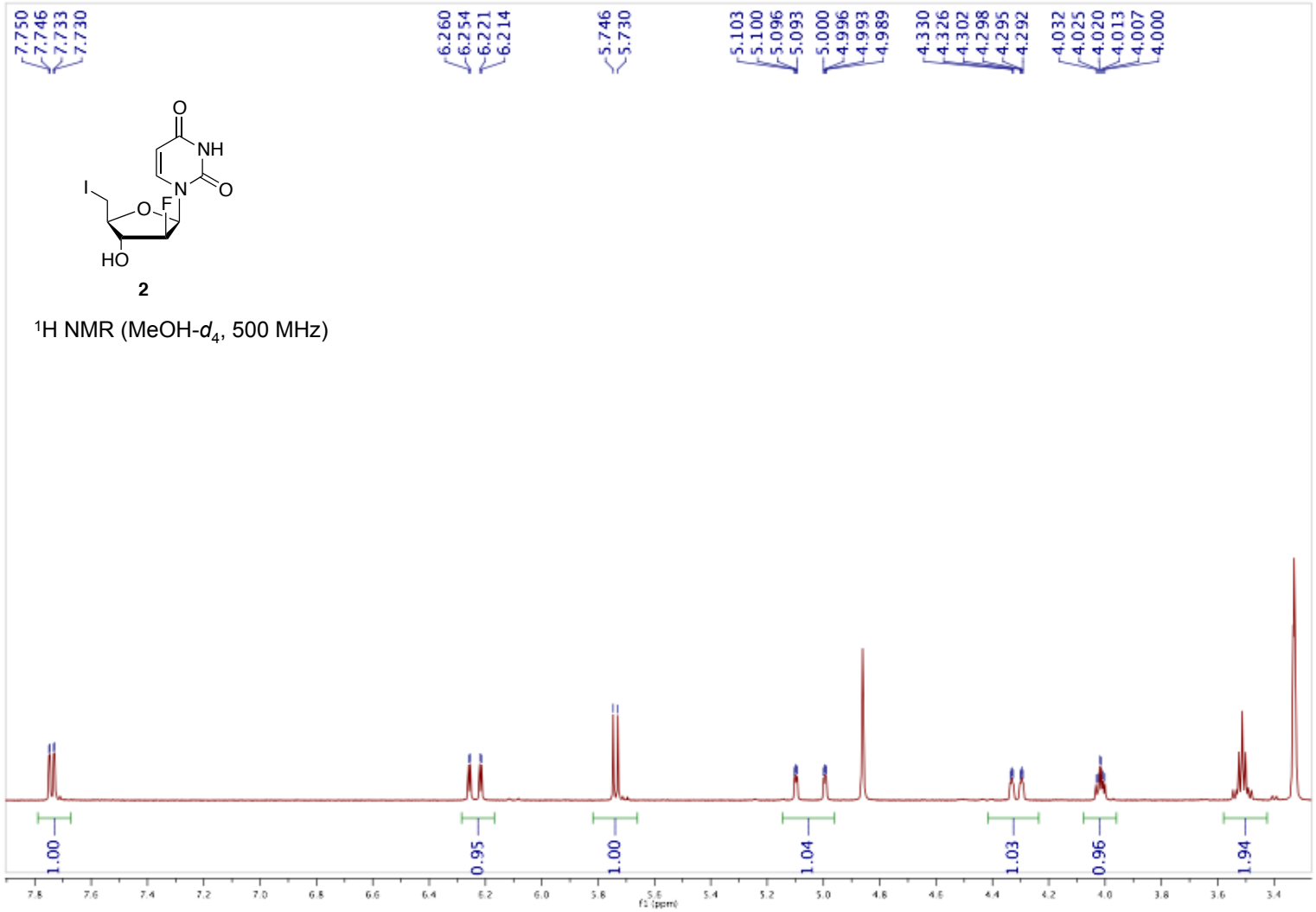
SUPPORTING INFORMATION (page 1 of 26 pages)

Table of Contents



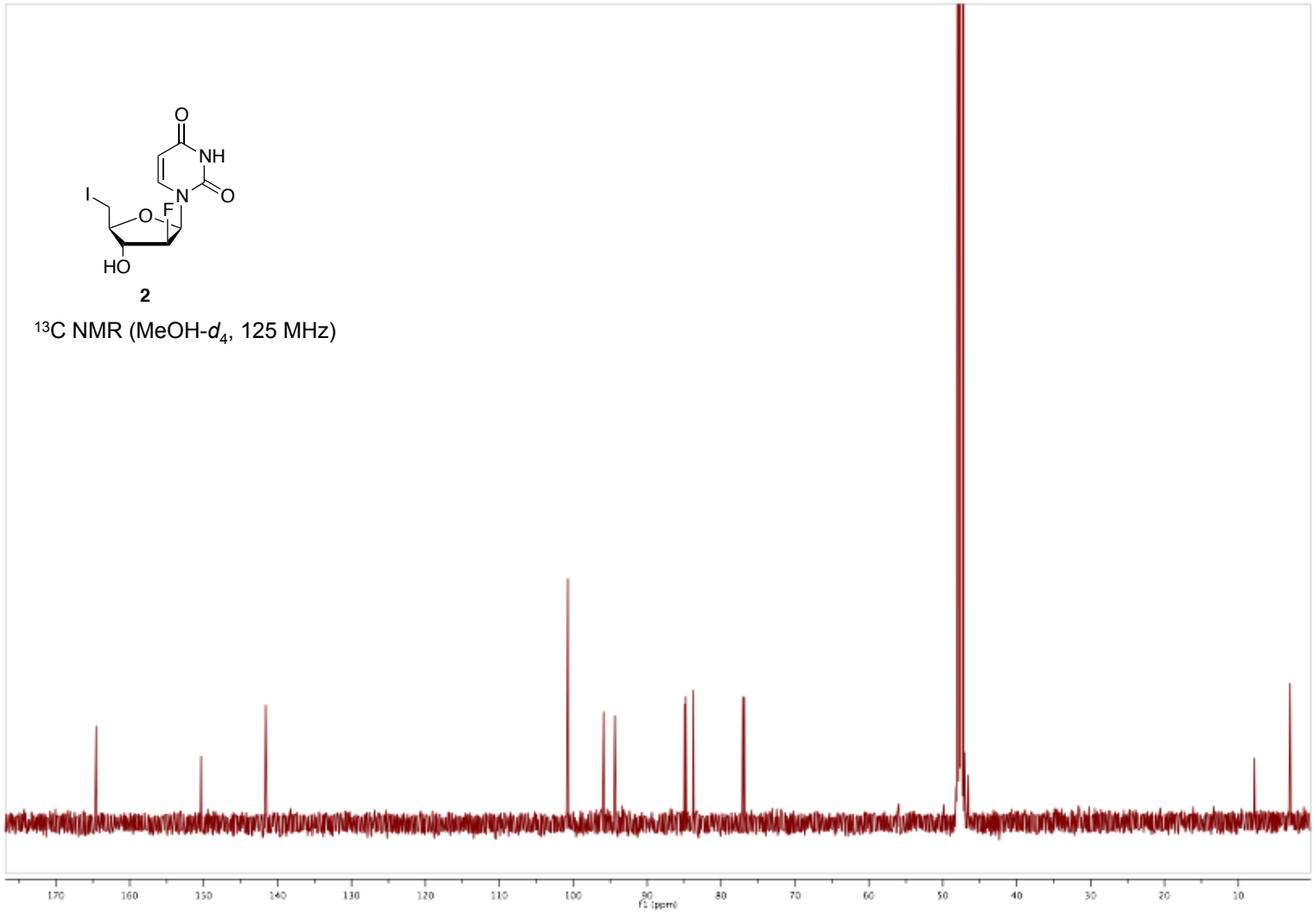
1

^{1}H NMR (D_2O, 500 MHz)



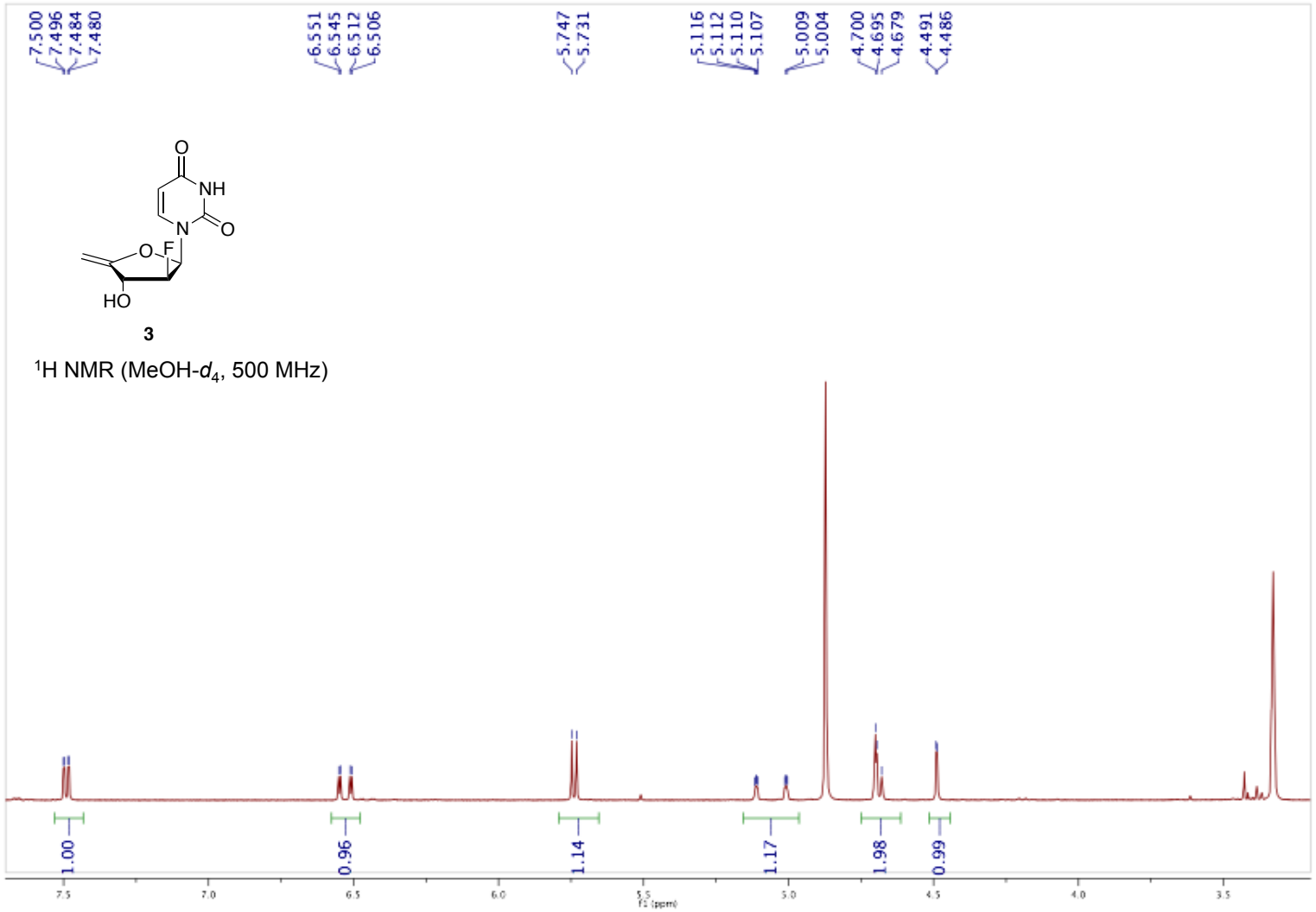
2

^{1}H NMR (MeOH-d_4, 500 MHz)



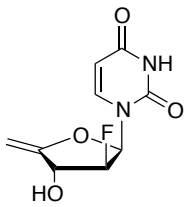
2

^{13}C NMR (MeOH-d_4, 125 MHz)



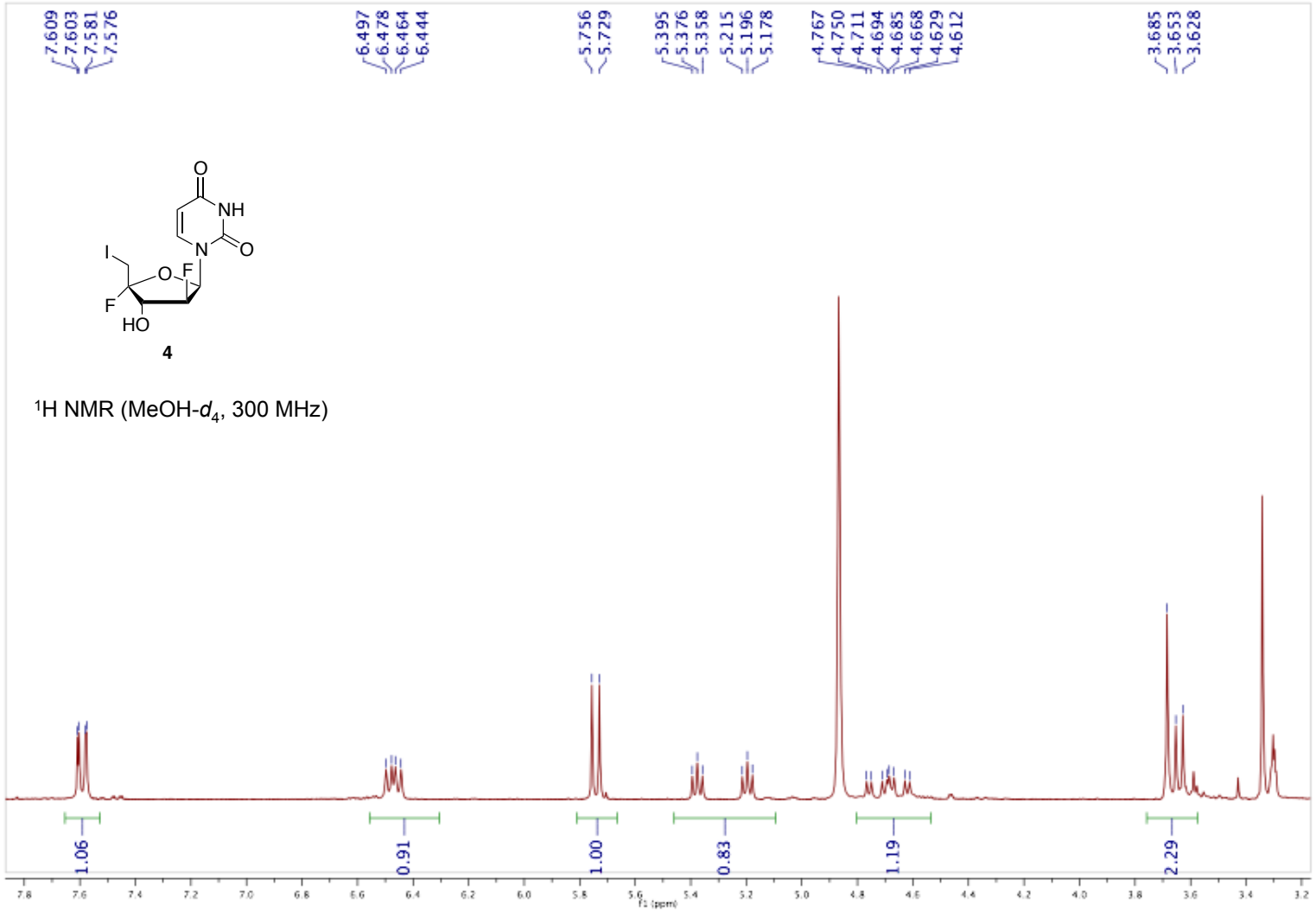
3

^{1}H NMR (MeOH-d_4, 500 MHz)



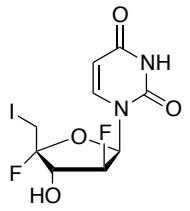
3

^{13}C NMR (MeOH-d_4, 125 MHz)



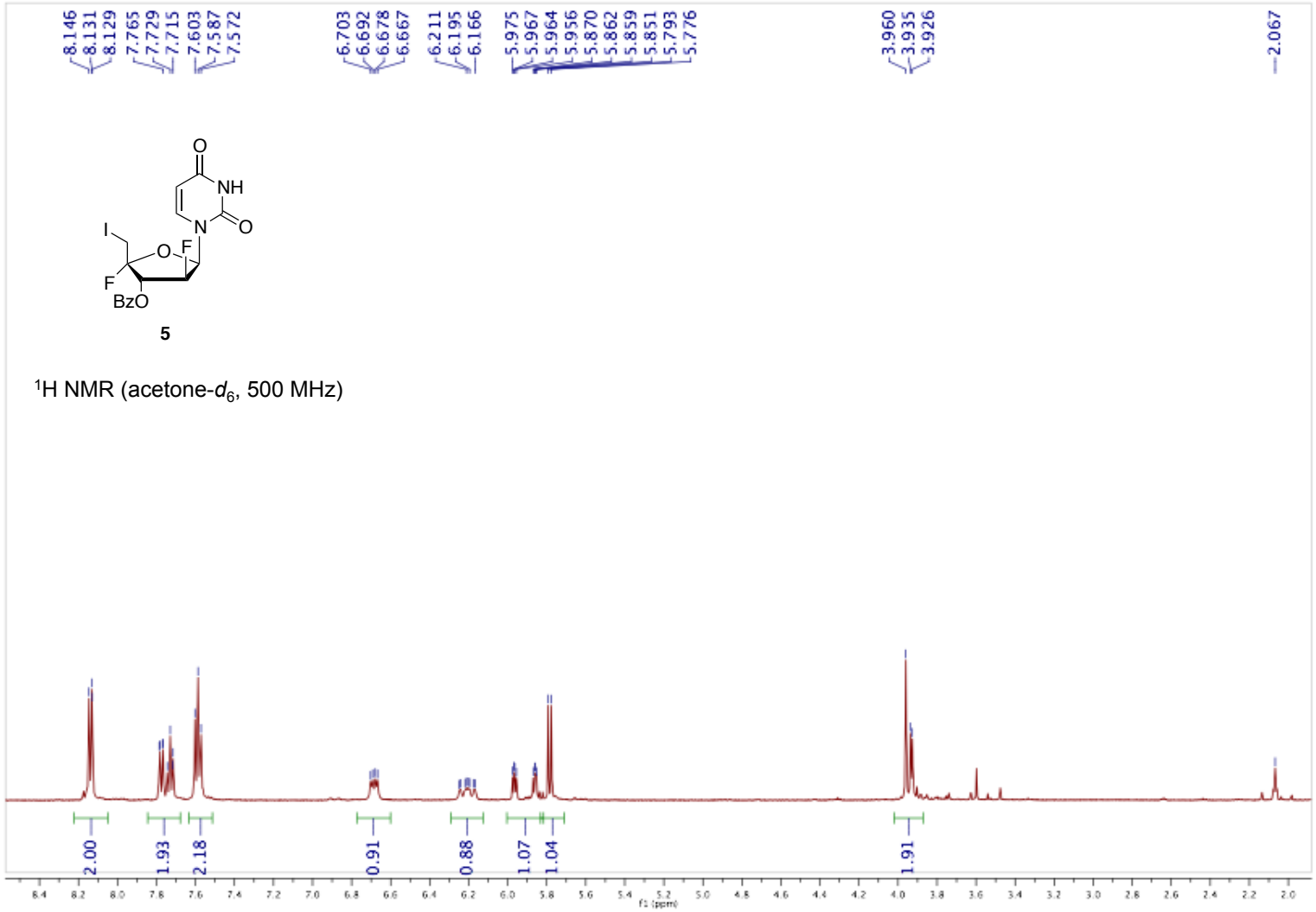
4

^{1}H NMR (MeOH-d_4, 300 MHz)



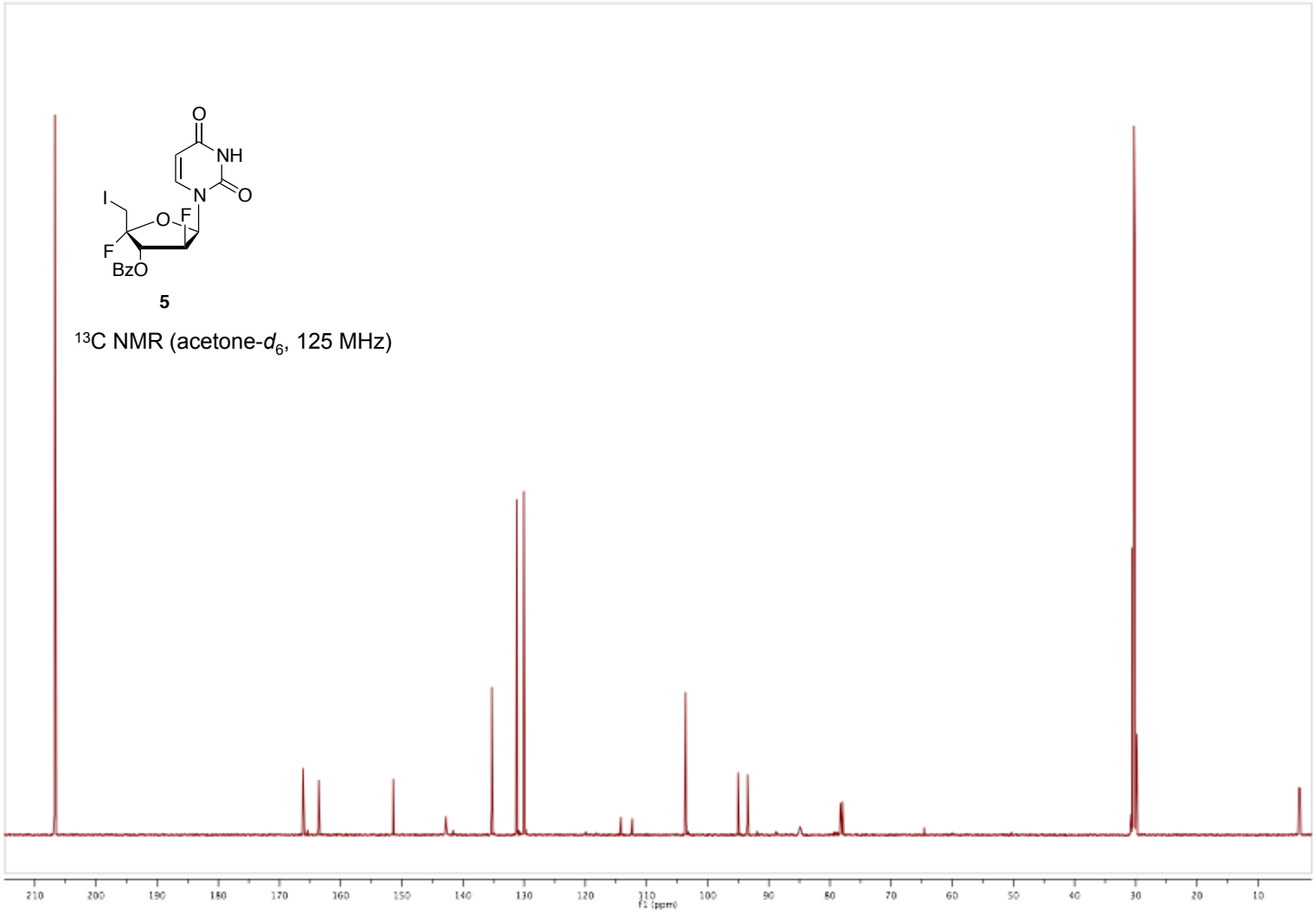
4

^{13}C NMR (MeOH-d_4, 125 MHz)



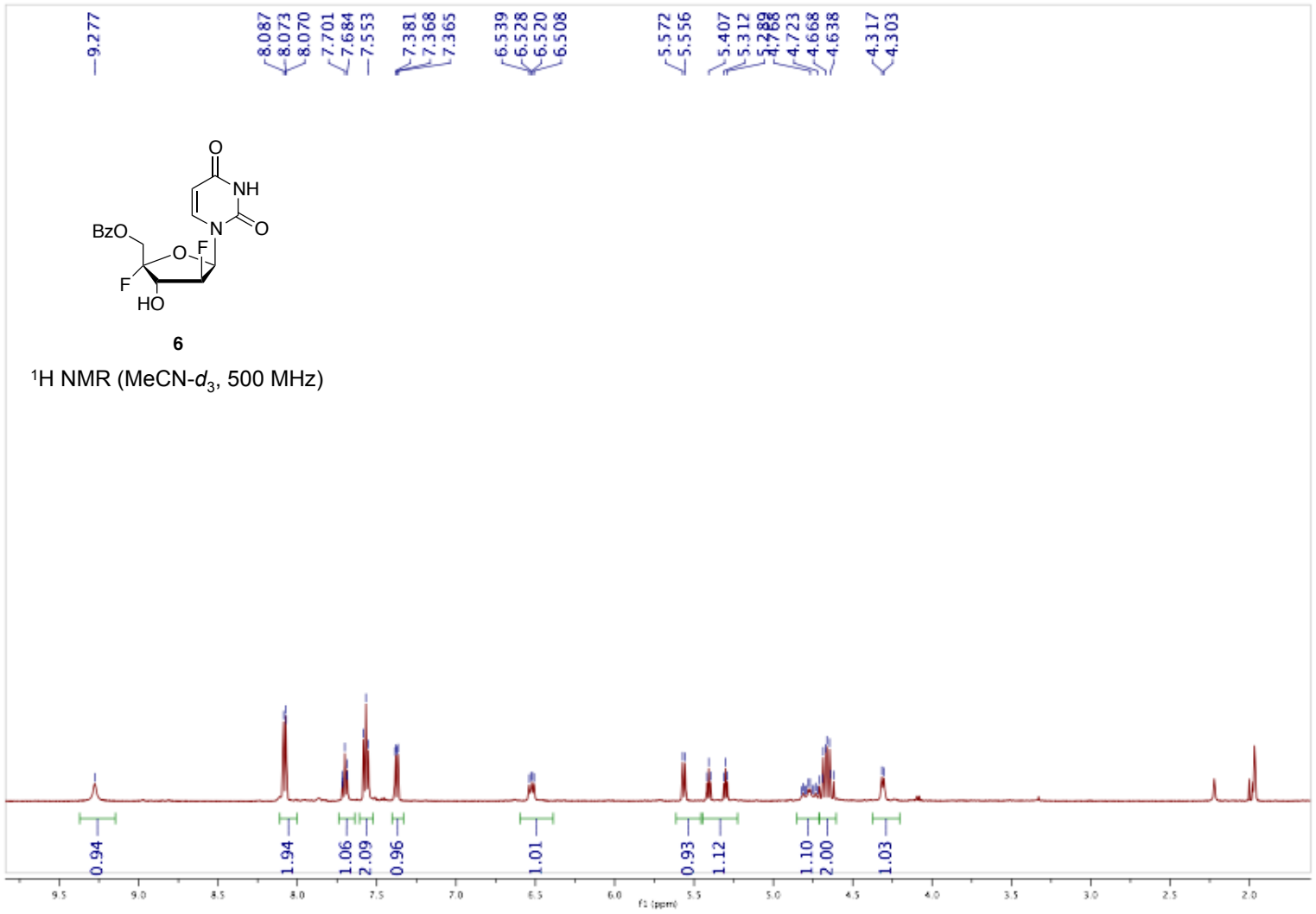
5

^{1}H NMR (acetone-d_6, 500 MHz)



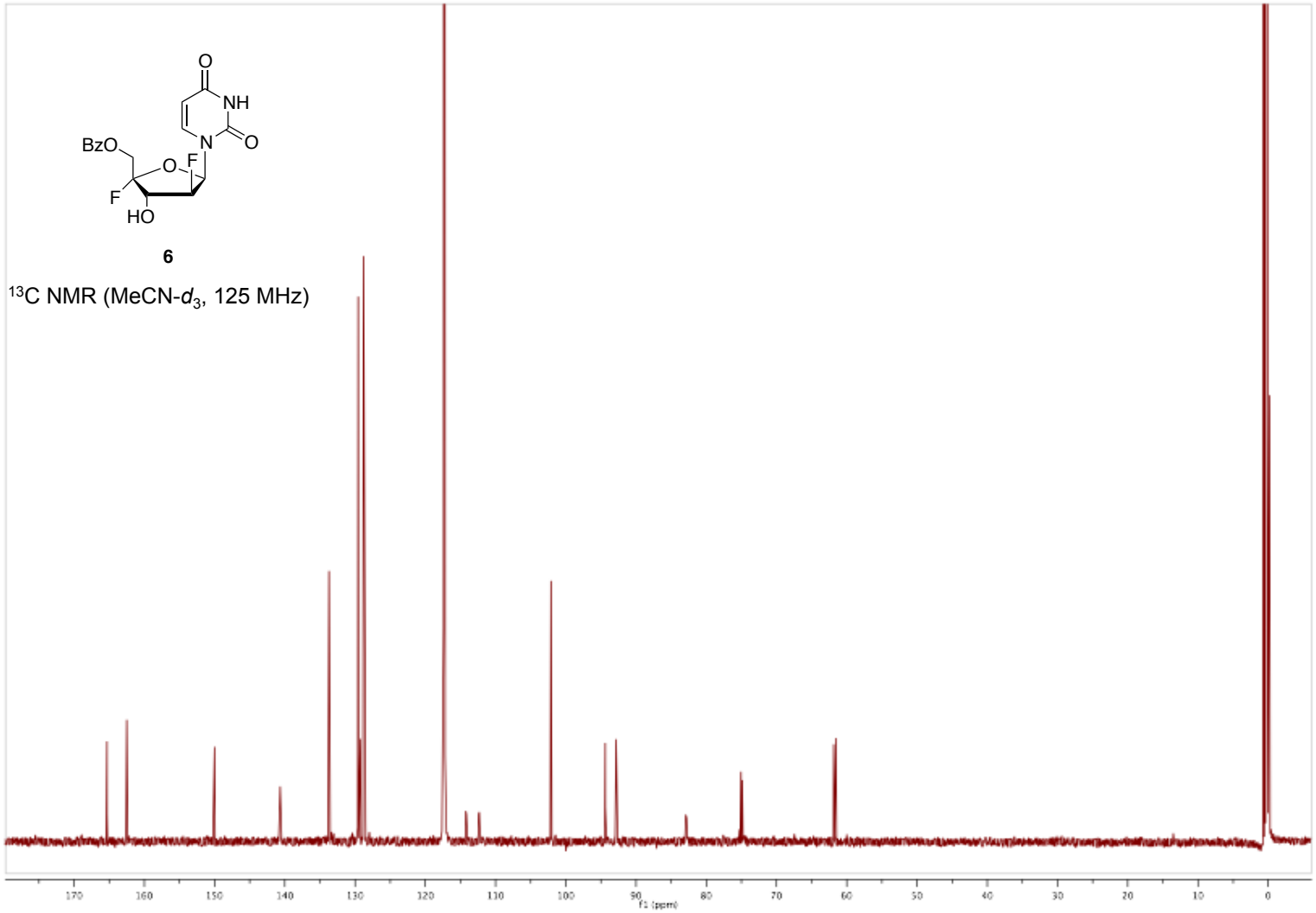
5

^{13}C NMR (acetone-d_6, 125 MHz)



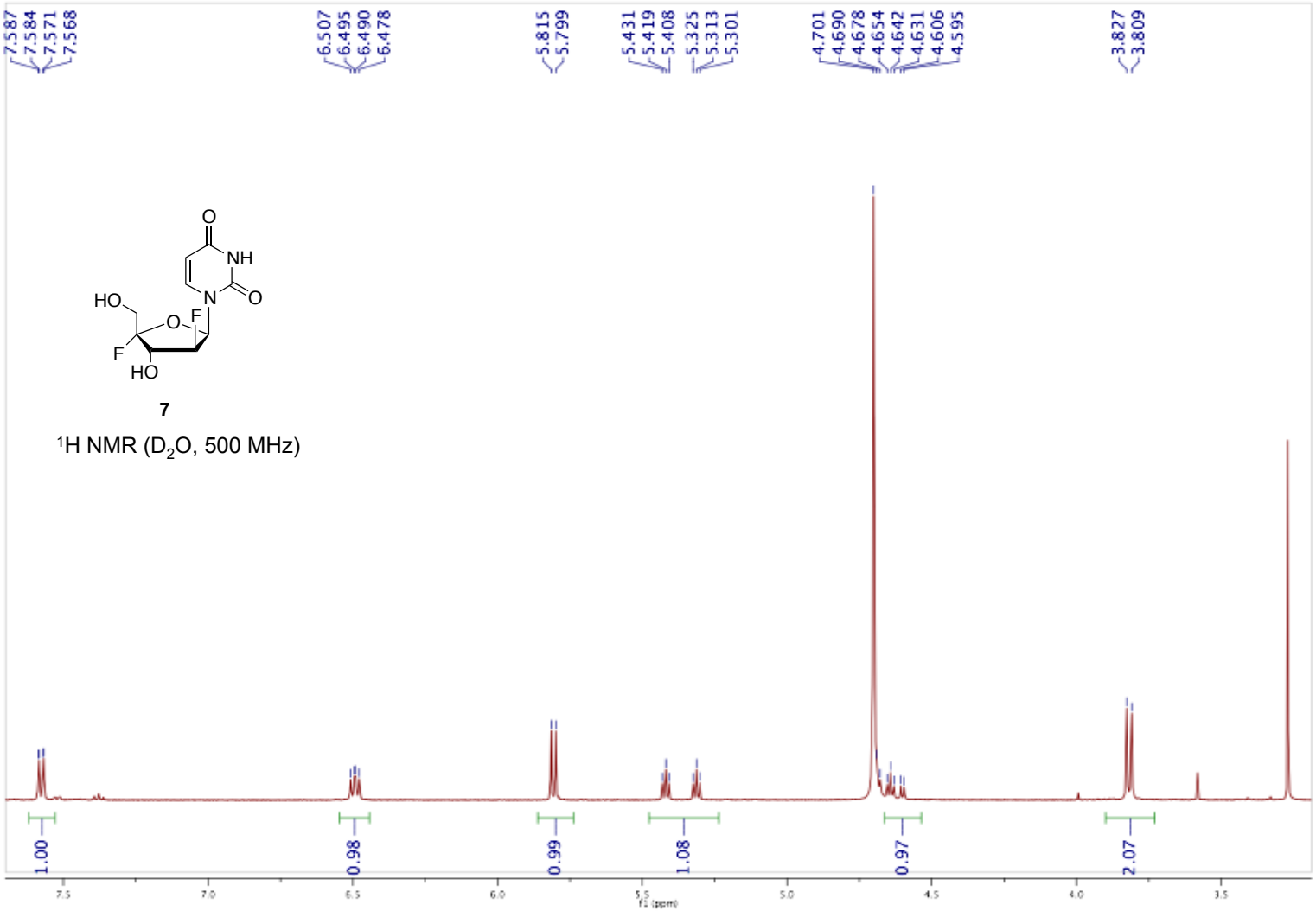
6

^{1}H NMR (MeCN-d_3, 500 MHz)



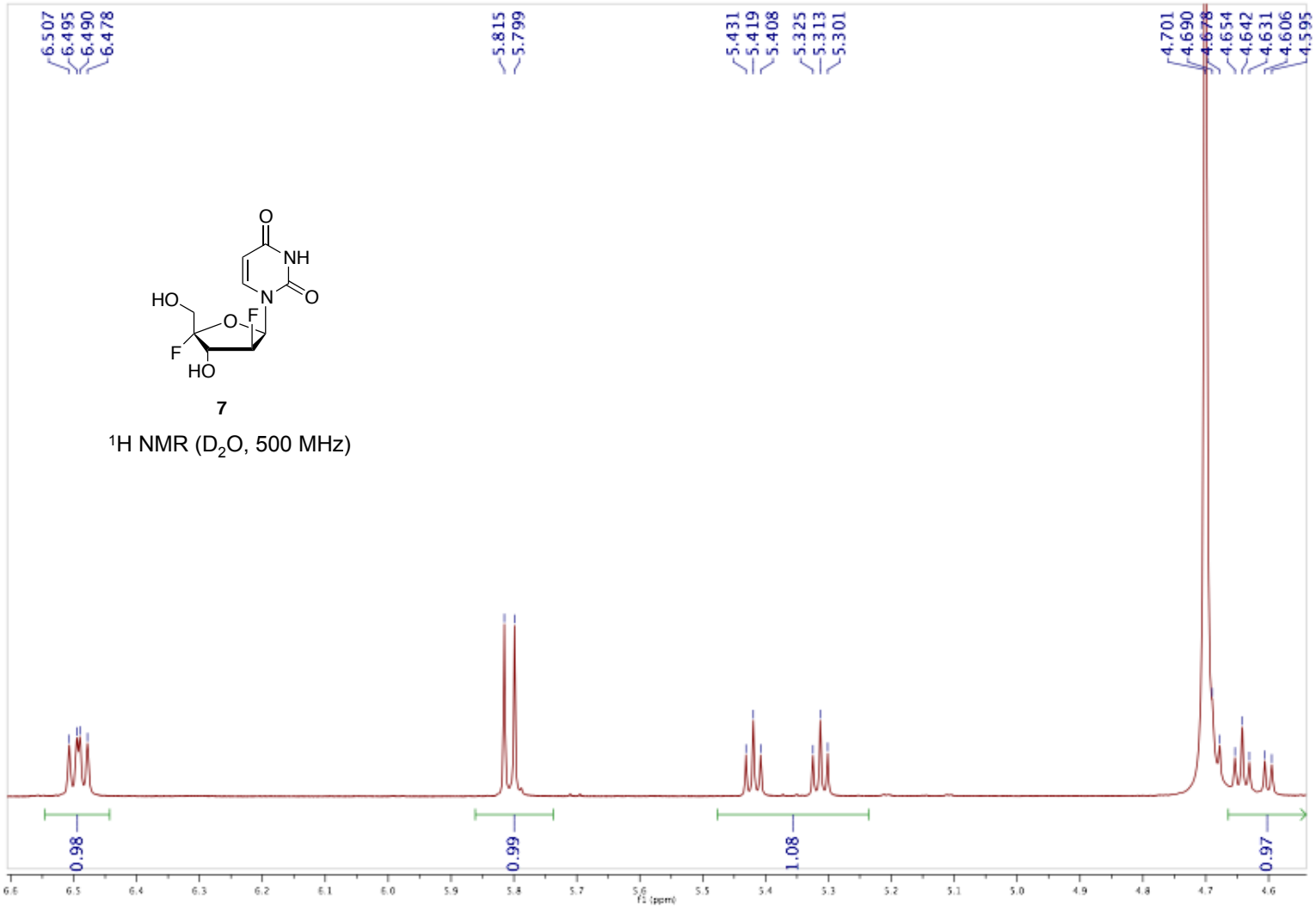
6

^{13}C NMR (MeCN-d_3, 125 MHz)



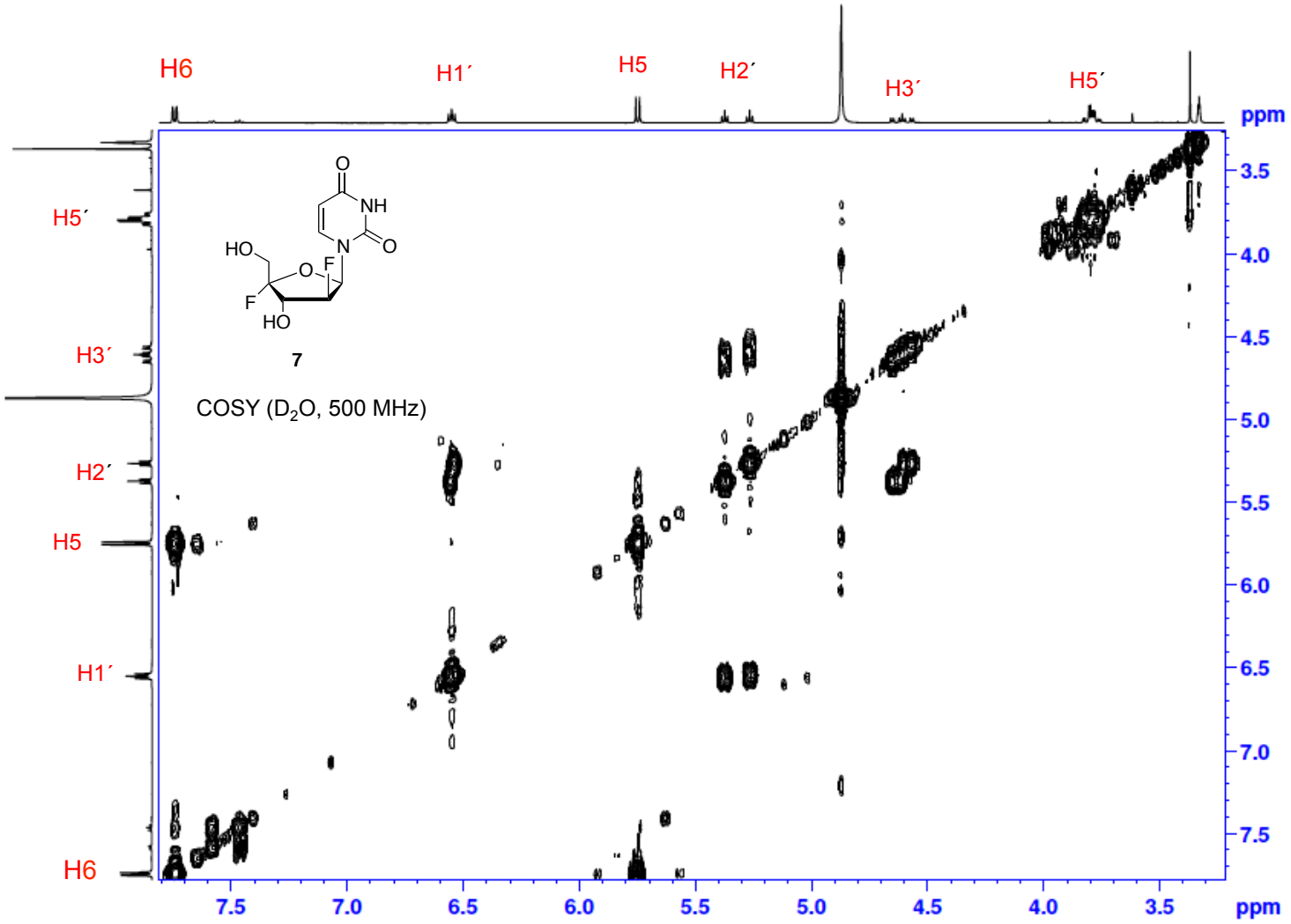
7

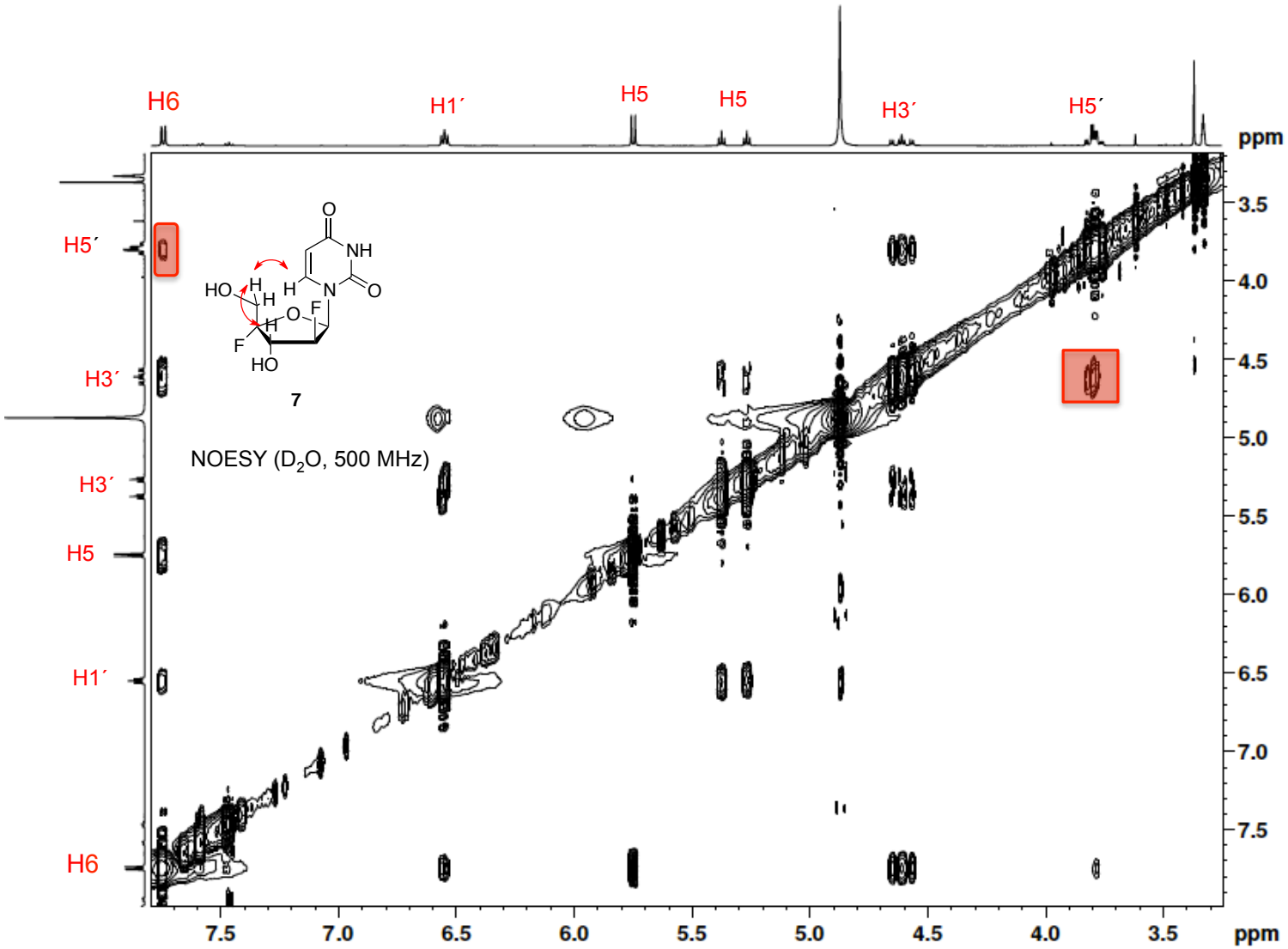
^{1}H NMR (D_2O, 500 MHz)

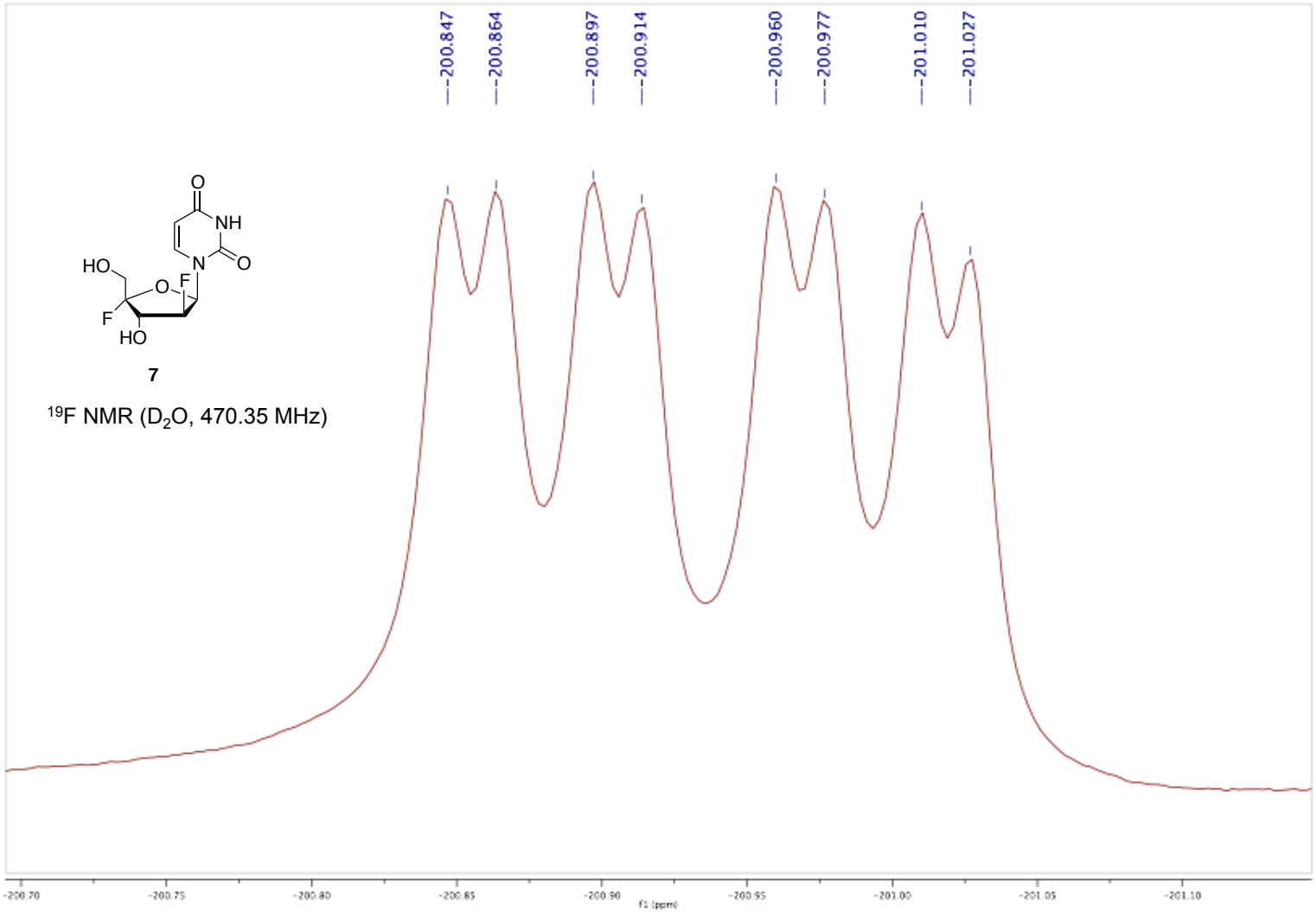


7

^{1}H NMR (D_2O, 500 MHz)

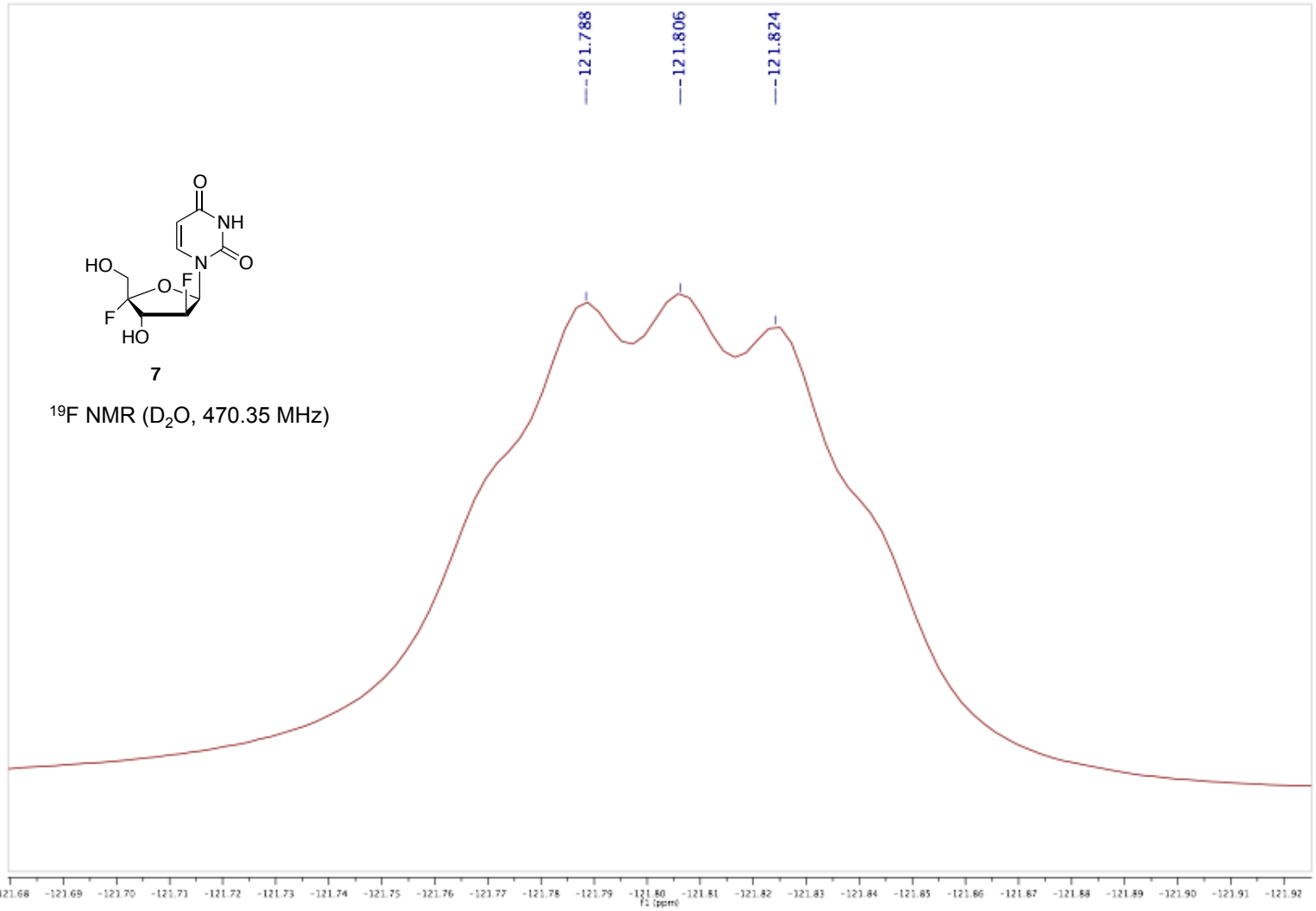

H6
H1´
H5
H2´
H3´
H5´
ppm
3.5
4.0
4.5
5.0
5.5
6.0
6.5
7.0
7.5
O
NH
N
O
HO
O
F
F
HO
7
COSY (D2O, 500 MHz)


H6
H1´
H5
H5
H3´
H5´
ppm
H5´
H3´
H3´
H5
H1´
H6
7
NOESY (D2O, 500 MHz)
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
ppm



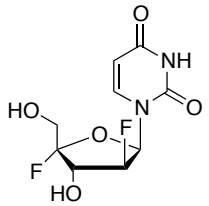
7

^{19}F NMR (D_2O, 470.35 MHz)



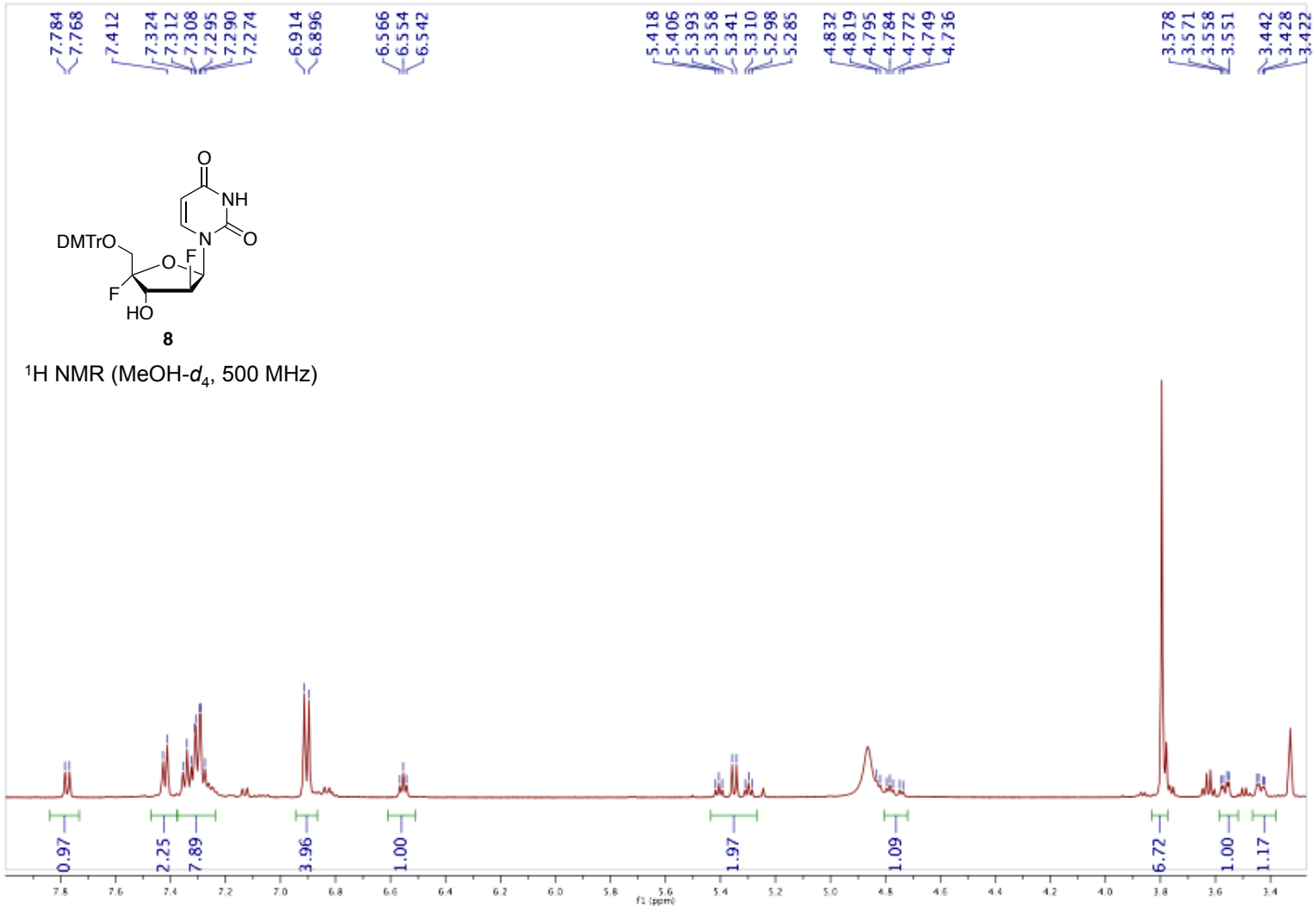
7

^{19}F NMR (D_2O, 470.35 MHz)



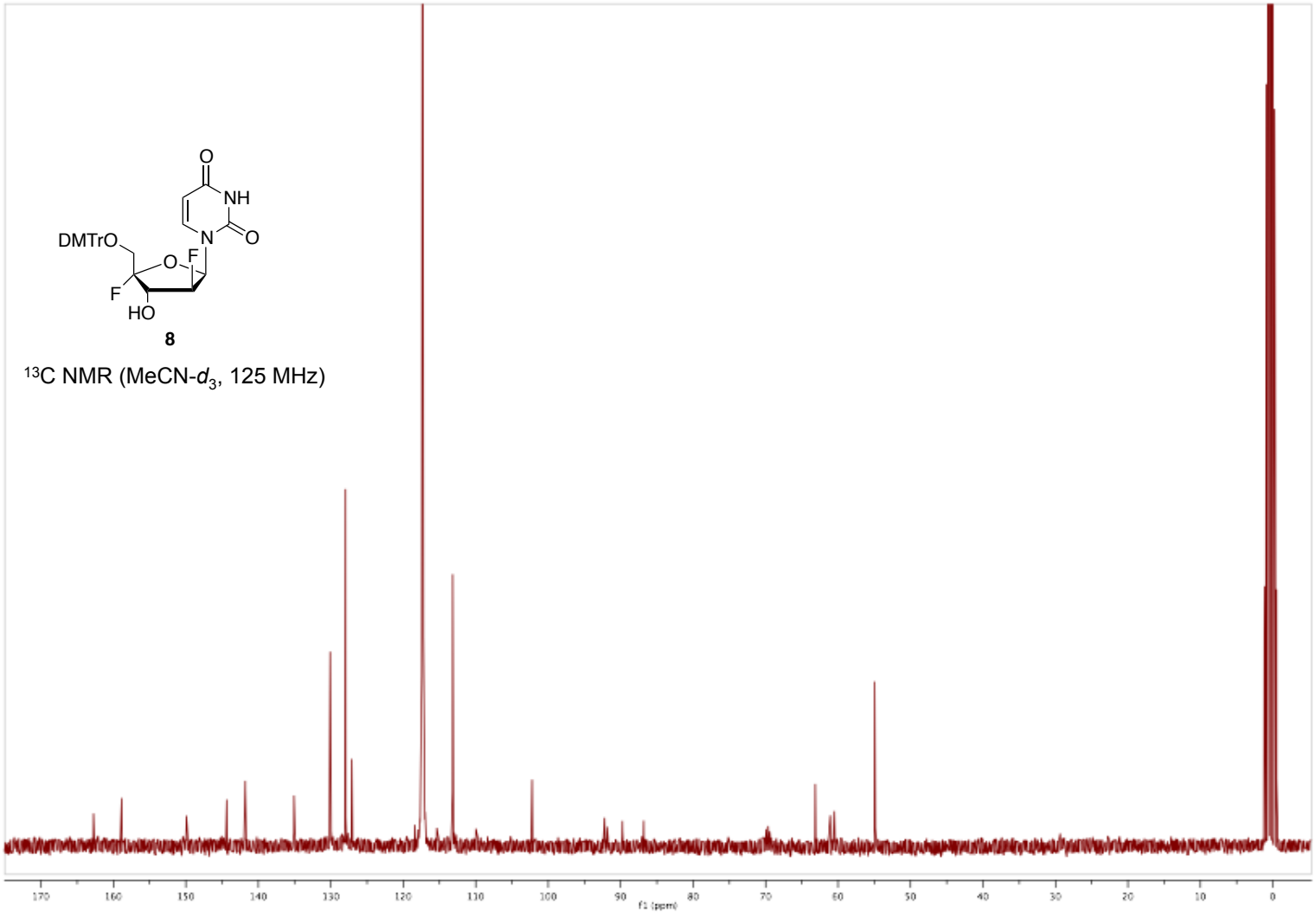
7

^{13}C NMR (MeOH-d_4, 125 MHz)



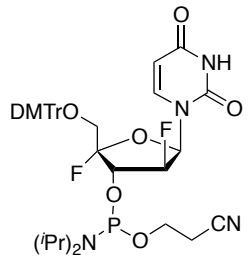
8

^{1}H NMR (MeOH-d_4, 500 MHz)



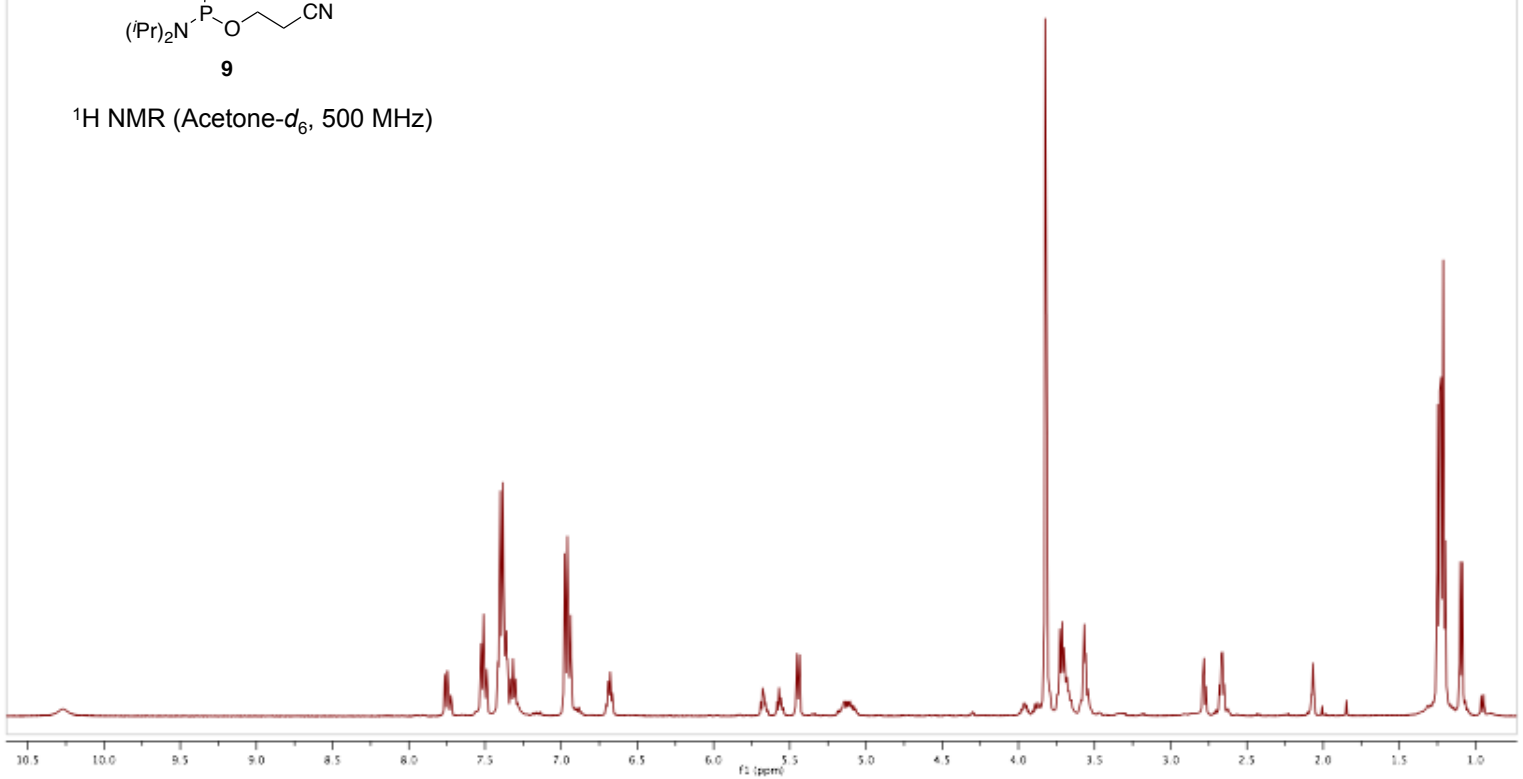
8

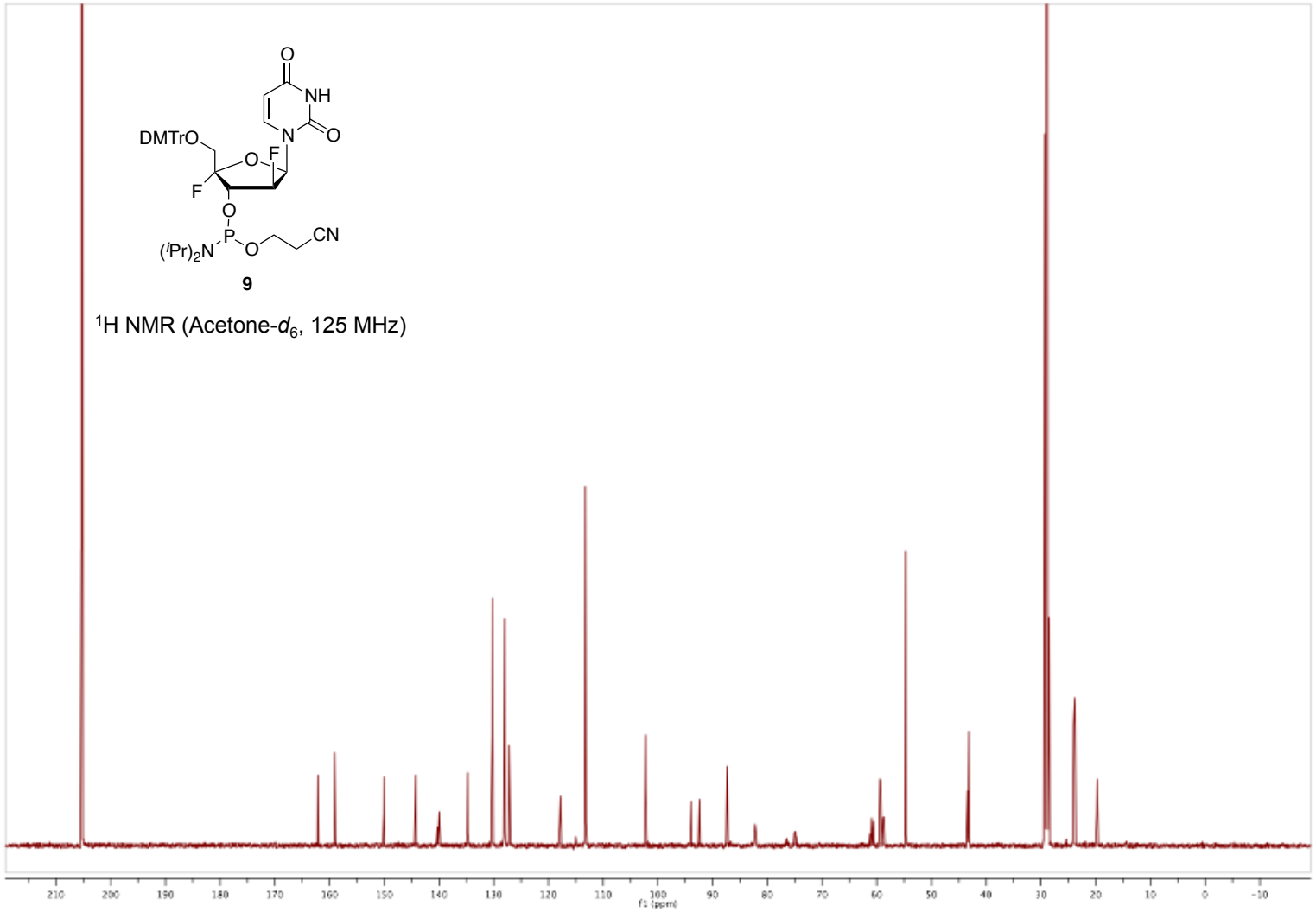
^{13}C NMR (MeCN-d_3, 125 MHz)

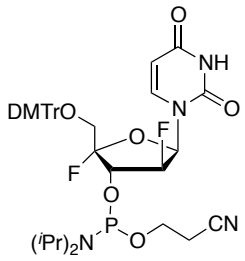


9

^{1}H NMR (Acetone-d_6, 500 MHz)




O
NH
N
O
DMTrO
O
F
F
O
O
P
O
CN
(iPr)2N
9
1H NMR (Acetone-d6, 125 MHz)
210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10
f1 (ppm)



9

^{31}P NMR (acetone-d_6, 81 MHz)

2. Conformational analysis

2.1. $J_{H1'H2''}$ *and* $J_{H2''H3'}$ vs pseudorotational angle calculation

Vicinal coupling constants (J_{HH}) vs sugar phase angle were calculated based on the Karplus equation using an “in house program”:

$$Karplus\ equation\text{:}\ J_{HH} = P_1 cos^2\varphi + P^2 cos\varphi + P_3 + \sum_1 \Delta\chi_i [P_4 + P_5 cos^2(\xi_i \varphi + P^6 |\Delta\chi_i|)] - \Delta J$$

In this equation φ is the torsion angle of H-C-C-H fragment; ξ value is +1 or -1 depends on the orientation of the substituent respect to the germinal proton; Δχ represents the electronegativity difference between the α-substituent and the hydrogen; Δ*J* measures the spin-spin interaction through space and P1, P2, P3, P4, P5, P6 y P7 values are dependent on the number of substituent in α.

Fluorine electronegativity in positions 2’ and 4’ of the sugar was taken into account to calculate $J_{H1'H2''}$ and $J_{H2''H3'}$ in 2´-deoxy-2´-fluoroarauridine and 2´-deoxy-2´,4´-difluoroarabinouridine.

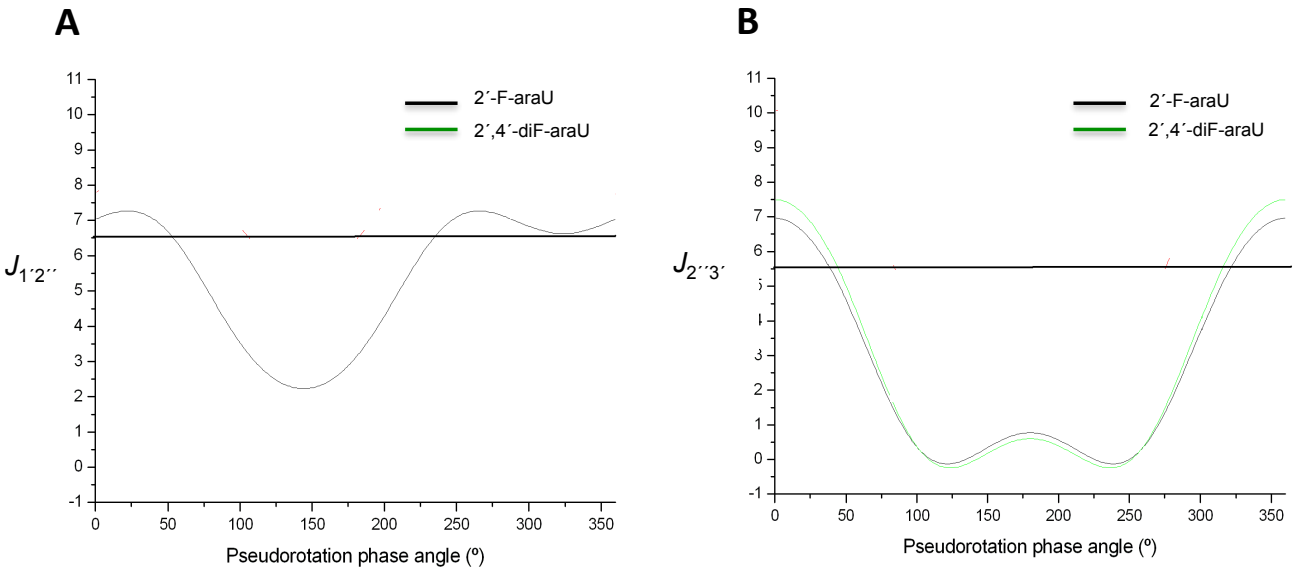


Figure S1. (A) $J_{1'2''}$ vs sugar pseudorotation angle for 2´-F-ara-U (black) and 2´,4´-diF-araU (green). Sugar pucker of both nucleosides overlap. (B) $J_{2''3'}$ vs sugar pseudorotation angle.

2.2. Estimation of the North/South population for 2´-F-araU and 2´,4´-diF-araU.

$^3J_{H1'2''}$ and $^3J_{H2''H3'}$ vs sugar pseudorotation phase angle for 2´-F-araU and 2´,4´-diF-ara-U, calculated from the generalized Karplus equation. Values from idealized North and South conformer can be obtained from averaging J values between 0º-50º and 120º-180º, resulting in $^{North}J_{1'2''}$ =7.1 Hz, $^{South}J_{1'2''}$ =2.5 Hz, $^{North}J_{2''3'}$ =6.6 Hz and $^{South}J_{2''3'}$ = 0 for 2´-F-araU, and $^{North}J_{1'2''}$ =7.1 Hz, $^{South}J_{1'2''}$ =2.5 Hz, $^{North}J_{2''3'}$ =7.2 Hz and 6.1 $^{South}J_{2''3'}$ = 0 for 2´,4´-diF-ara-U.

Relative population of North (P_{North}) and (P_{South}) South conformers can be estimated from

$$^{Exp}J_{1'2''} = P_{North}\ ^{North}J_{1'2''} + P_{South}\ ^{South}J_{1'2''}$$

$$^{Exp}J_{2''3'} = P_{North}\ ^{North}J_{2''3'} + P_{South}\ ^{South}J_{2''3'}$$

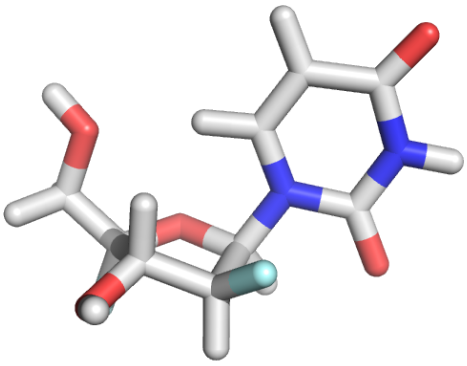
Experimental values of $J_{1'2''}$ and $J_{2''3'}$ are consistent with approximate relative North/South ratios of 80/20 and 40/60, for 2′,4′-diF-araU and 2′-F-araU, respectively.

2.3. Quantum mechanical calculations

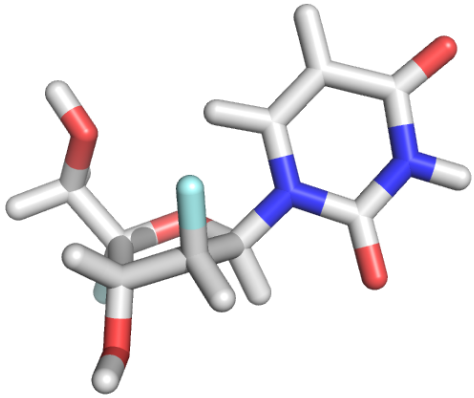
To explore the impact of the 4′-fluorine substitutions on the sugar puckering preference, we have calculated the pseudorotation energetic profile of 2′,4′-diF-araU, 2′-F-araU, and 2′,4′-diF-rU by means of constrained energy optimizations, also at the M062x/6-31+G(d,p) level. Initial coordinates for the constrained energy optimizations were obtained based on canonical parameters and Nucleic Acid Builder language.[1] In these optimizations the beta, gamma, and epsilon dihedral angles were fixed at values compatible with those found in NMR-refined ensembles of duplex DNA (beta=172 deg, gamma=58.3 deg, epsilon=192), while the dihedral angles v1 and v2 of the sugar where restrained to values leading to a unique pseudorotation angle. All quantum mechanical calculations have been carried out with Gaussian 09 rev B 01, using default threshold values for the optimization procedure. After optimization, the lowest energy value was recorded and an associated with the value of the sugar pucker pseudorotation angle for the optimized structure. The pseudorotation angle for the sugar puckering in each optimized structure was computed using Curves+.[2]

[1] T. Macke, D. A. Case, in *Molecular Modeling of Nucleic Acids* (Eds.: N. Leontes, J. J. SantaLucia), American Chemical Society, Washington, DC, **1998**, pp. 379-393.
[2]R. Lavery, M. Moakher, J. H. Maddocks, D. Petkeviciute, K. Zakrzewska, *Nucleic Acids Res.* **2009**, *37*, 5917–5929.

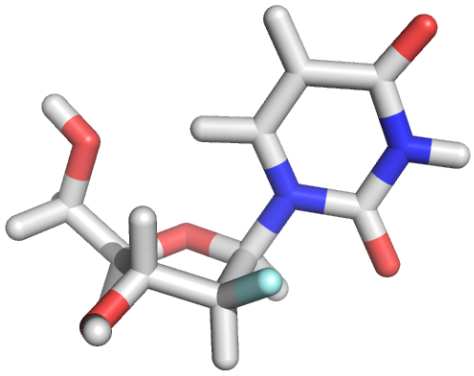
2.3.1. Minimized structures for the North and South conformers of 2´,4´-diF-araU 2´-F-araU and 2´,4´-diF-rU.



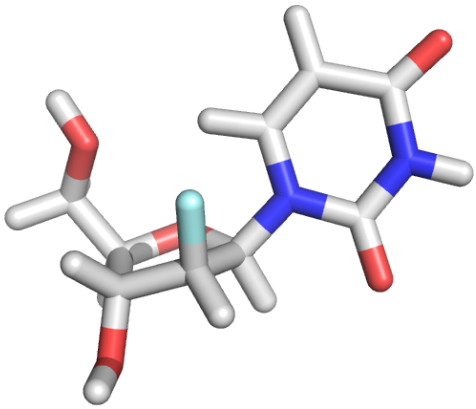
North 2´,4´-diF-araU



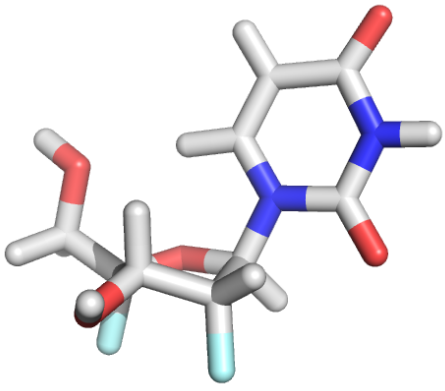
South 2´,4´-diF-araU



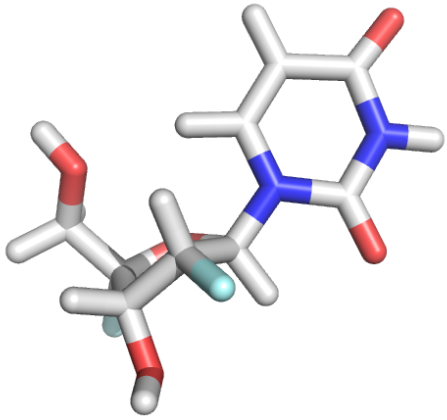
North 2´-F-araU



South 2´-F-araU



North 2´,4´-diF-rU



South 2´,4´-diF-rU

2.3.2. X,Y,Z coordinates for the minimized structures for the North and South conformers of 2´,4´-diF-araU, 2´-F-araU and 2´,4´-diF-rU.

	X	Y	Z
H	-2.25170	3.38846	0.03825
O	-1.97056	2.47275	0.13033
C	-2.92401	1.61603	-0.46832
H	-3.12921	1.88966	-1.50959
H	-3.86149	1.58215	0.09879
C	-2.33087	0.22442	-0.46650
F	-3.26855	-0.61166	-1.04366
O	-1.17116	0.19252	-1.22887
C	-0.27461	-0.80512	-0.77349
H	-0.11369	-1.55577	-1.54561
N	1.01538	-0.18572	-0.50070
C	1.11115	1.09691	0.00154
H	0.17504	1.64737	0.03607
C	2.28545	1.63631	0.37942
H	2.35270	2.64748	0.75459
C	3.51301	0.85678	0.27758
O	4.62880	1.22683	0.58391
N	3.31107	-0.43483	-0.23350
H	4.13008	-1.02828	-0.30366
C	2.12259	-1.02621	-0.59601
O	2.03792	-2.17873	-0.97550
C	-1.94721	-0.34648	0.89645
H	-1.43196	0.45023	1.45361
C	-0.91904	-1.39804	0.50645
F	0.01660	-1.57716	1.50406
H	-1.39411	-2.36223	0.30675
O	-3.08218	-0.80787	1.56953
H	-2.85463	-0.98750	2.48876

North 2´,4´-diF-araU

	X	Y	Z
H	-2.14016	3.27701	0.35105
O	-1.98418	2.32798	0.38619
C	-2.79274	1.67612	-0.57043
H	-2.65077	2.08155	-1.57922
H	-3.85774	1.71470	-0.30917
C	-2.35523	0.22337	-0.60123
F	-3.14425	-0.39728	-1.53697
O	-1.02871	0.15521	-1.01244
C	-0.28728	-0.76930	-0.23113
H	-0.28557	-1.76697	-0.67623
N	1.08074	-0.30305	-0.17865
C	1.35125	1.04299	-0.00266
H	0.46792	1.66884	0.07494
C	2.60756	1.52154	0.05123
H	2.81063	2.57467	0.18240
C	3.73911	0.61148	-0.06582
O	4.91552	0.91316	-0.03577
N	3.36071	-0.73293	-0.22386
H	4.11264	-1.40961	-0.29006
C	2.09060	-1.25851	-0.26936
O	1.86535	-2.44997	-0.37147
C	-2.45416	-0.53647	0.75635
H	-2.93634	0.09434	1.51254
C	-0.98896	-0.77585	1.13021
F	-0.51925	0.26939	1.90079
H	-0.83161	-1.70779	1.67962
O	-3.13085	-1.74633	0.53702
H	-3.36567	-2.13648	1.38615

South 2´,4´-diF-araU

	X	Y	Z
H	-2.45008	3.28859	0.06152
O	-2.16646	2.37000	0.10389
C	-3.10834	1.54553	-0.56150
H	-3.30200	1.91266	-1.57673
H	-4.05653	1.48908	-0.01056
C	-2.53672	0.14921	-0.64779
H	-3.27376	-0.49513	-1.15099
O	-1.31646	0.16632	-1.38355
C	-0.42261	-0.81151	-0.90995
H	-0.18731	-1.53558	-1.69043
N	0.84054	-0.17788	-0.53849
C	0.87774	1.09133	-0.00047
H	-0.07463	1.61383	-0.00337
C	2.01649	1.64755	0.45661
H	2.03785	2.64943	0.86103
C	3.26546	0.89975	0.40358
O	4.35461	1.28559	0.78154
N	3.12286	-0.38133	-0.15116
H	3.95823	-0.95410	-0.19057
C	1.96986	-0.98908	-0.59495
O	1.93703	-2.13289	-1.00979
C	-2.19406	-0.47920	0.69900
H	-1.74908	0.27988	1.35891
C	-1.11452	-1.47590	0.30845
F	-0.22537	-1.69902	1.33948
H	-1.55799	-2.43833	0.03308
O	-3.35692	-1.03931	1.25598
H	-3.19443	-1.26837	2.17791

North 2´-F-araU

	X	Y	Z
H	-2.14016	3.27701	0.35105
O	-1.98418	2.32798	0.38619
C	-2.79274	1.67612	-0.57043
H	-2.65077	2.08155	-1.57922
H	-3.85774	1.71470	-0.30917
C	-2.35523	0.22337	-0.60123
F	-3.14425	-0.39728	-1.53697
O	-1.02871	0.15521	-1.01244
C	-0.28728	-0.76930	-0.23113
H	-0.28557	-1.76697	-0.67623
N	1.08074	-0.30305	-0.17865
C	1.35125	1.04299	-0.00266
H	0.46792	1.66884	0.07494
C	2.60756	1.52154	0.05123
H	2.81063	2.57467	0.18240
C	3.73911	0.61148	-0.06582
O	4.91552	0.91316	-0.03577
N	3.36071	-0.73293	-0.22386
H	4.11264	-1.40961	-0.29006
C	2.09060	-1.25851	-0.26936
O	1.86535	-2.44997	-0.37147
C	-2.45416	-0.53647	0.75635
H	-2.93634	0.09434	1.51254
C	-0.98896	-0.77585	1.13021
F	-0.51925	0.26939	1.90079
H	-0.83161	-1.70779	1.67962
O	-3.13085	-1.74633	0.53702
H	-3.36567	-2.13648	1.38615

South 2´-F-araU

	X	Y	Z
H	-1.73525	3.59371	0.14791
O	-1.54943	2.65248	0.22451
C	-2.61452	1.91083	-0.33974
H	-2.83167	2.21979	-1.36868
H	-3.52601	1.97492	0.26576
C	-2.19383	0.45667	-0.37842
F	-3.24429	-0.22882	-0.94087
O	-1.06258	0.29234	-1.17663
C	-0.25017	-0.76709	-0.71212
H	-0.12283	-1.52655	-1.48511
N	1.07682	-0.22686	-0.37458
C	1.31340	1.11857	-0.17603
H	0.44326	1.75767	-0.28187
C	2.53681	1.59584	0.12748
H	2.71010	2.65160	0.27969
C	3.66953	0.68833	0.24939
O	4.81592	0.99208	0.51241
N	3.32678	-0.65736	0.03254
H	4.07493	-1.33659	0.11622
C	2.08622	-1.17695	-0.25609
O	1.86910	-2.36733	-0.39142
C	-1.82442	-0.17931	0.96379
H	-1.17975	0.54590	1.48440
C	-0.95700	-1.35856	0.51612
H	-0.25095	-1.73086	1.26223
F	-1.75668	-2.40345	0.11752
O	-2.97835	-0.48151	1.68413
H	-2.75167	-0.70414	2.59353

North 2´,4´-diF-rU

	X	Y	Z
H	-1.79111	3.28794	0.90430
O	-1.73873	2.33640	0.76764
C	-2.59946	1.95790	-0.29355
H	-2.38760	2.52165	-1.20931
H	-3.65577	2.07492	-0.02442
C	-2.32240	0.49349	-0.58932
F	-3.15417	0.14992	-1.61952
O	-0.99570	0.33557	-1.00132
C	-0.31551	-0.63267	-0.20500
H	-0.33932	-1.62695	-0.65646
N	1.06863	-0.23559	-0.09396
C	1.37829	1.04733	0.31842
H	0.51274	1.66492	0.54220
C	2.64160	1.49858	0.41001
H	2.86946	2.50475	0.73190
C	3.75051	0.62101	0.05077
O	4.93130	0.90185	0.09498
N	3.33856	-0.64740	-0.38254
H	4.07182	-1.29101	-0.65828
C	2.05584	-1.13927	-0.50041
O	1.81269	-2.25560	-0.90736
C	-2.53012	-0.46296	0.60802
H	-3.20556	-0.02756	1.35321
C	-1.09523	-0.61110	1.11050
H	-0.79661	0.23328	1.73880
F	-0.92403	-1.77534	1.82167
O	-2.99650	-1.68535	0.10834
H	-3.31028	-2.22911	0.83939

South 2´,4´-diF-rU

3. Thermal denaturation studies

Sequences:

Control 5′ d(GCG TTT TTT GCT) 3′
3′ r(CGC AAA AAA CGA) 5′

Modified duplex 5′ d(GCG TT***U*** TTT GCT) 3′
3′ r(CGC AAA AAA CGA) 5

Sequences:

Control 5′ d(CUA UAG UAU AC) 3′
3′ r(GAU AUC AUA UG) 5′

Modified duplex 5′ d(C***U***A ***U***AG ***U***A***U*** AC) 3′
3′ r(GAU AUC AUA UG) 5′

U=2´-F-araU or 2´,4´-diF-araU

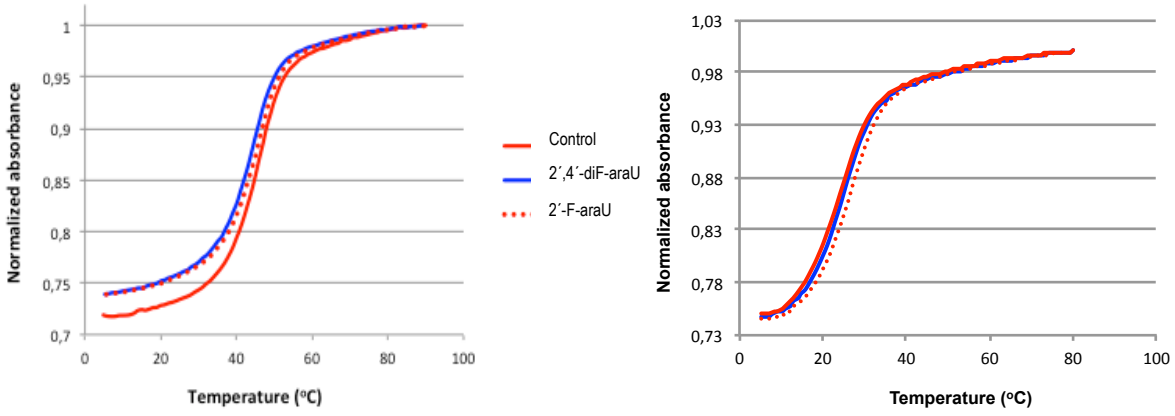


Figure S2. Sequences of studied duplexes and the corresponding thermal denaturation curves.

4. ESI-MS of synthesized oligonucleotides

Table S1. ESI-MS of DNA and RNA synthesized strands.

Oligomer	Modification	Sequence (5´-3´)	MS Calculated	Experimental
A1	DNA	d(GCGTTTTTTGCT)	3633.4[a]	3633.9
A2	2´,4´-diF-araU	d(GCGTTUTTTGCT)	3635.6083[b]	3653.5000
A3	2´-F-araU	d(GCGTTUTTTGCT)	3655.4[a]	3655.9
-	RNA	r(AGCAAAAAACGC)	3846.6322[b]	3846.5313
B1	DNA	d(CUAUAGUAUAC)	3257.6072[b]	3257.5000
B2	2´,4´-diF-araU	d(CUAUAGUAUAC)	3401.5320[b]	3401.4063
B3	2´-F-araU	d(CUAUAGUAUAC)	3329.5696[b]	3329.4531
-	RNA	r(GUAUACUAUAG)	3473.4912[b]	3473.4063

[a]Molecular weight. [b]Exact MS