Regioselective N1-Ribosylation of Hydantoin: Synthesis and Properties of the First Contracted Uridine Analog

Supplementary Material

Odai Bsoul [a], Yakir Lampel [a], Maayan Rofe [a], Natalie Pariente-Cohen[a], Chen Timsit [a], and Bilha Fischer [a]*

[a] Department of Chemistry, Bar-Ilan University, Ramat-Gan 52900, Israel.

* Corresponding author. Tel: 972-3-5318303

Bar-Ilan University, Ramat-Gan 52900, Israel

Bilha.fischer@biu.ac.il

Keywords: N1 hydantoinyl ribose, regioselective synthesis, Vorbrüggen reaction, conformational analysis, pK_a, base stacking, base pairing

Table of contents

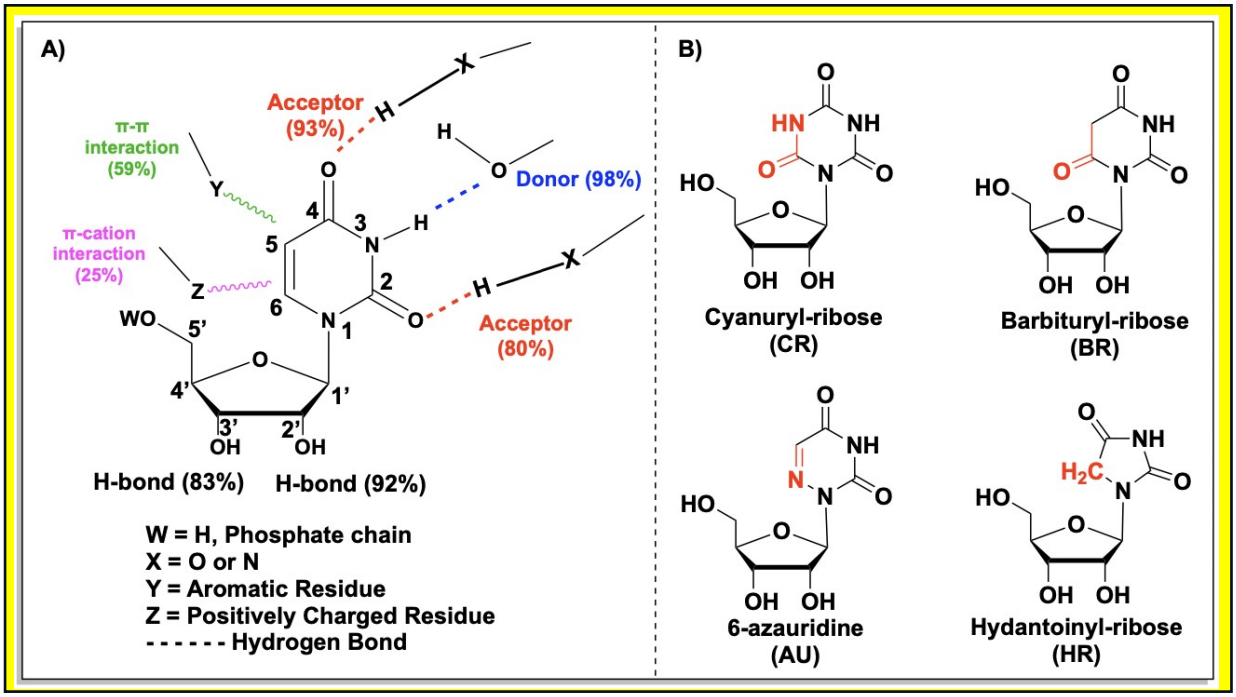
Figures

Tables

Equations

Experimental

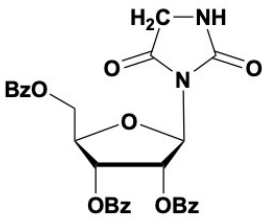
References

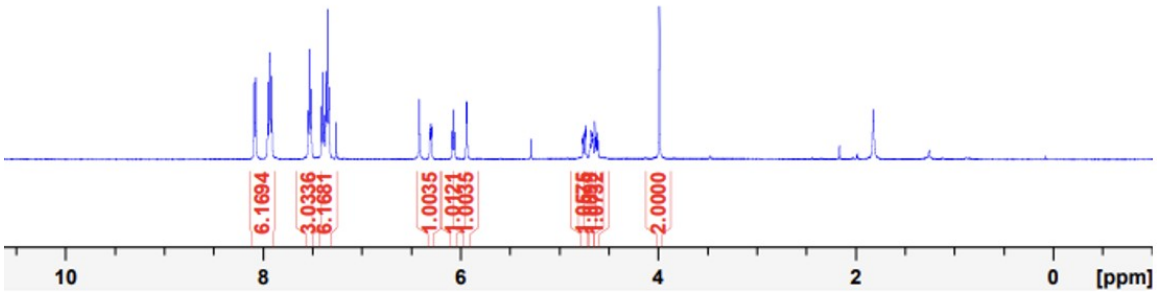


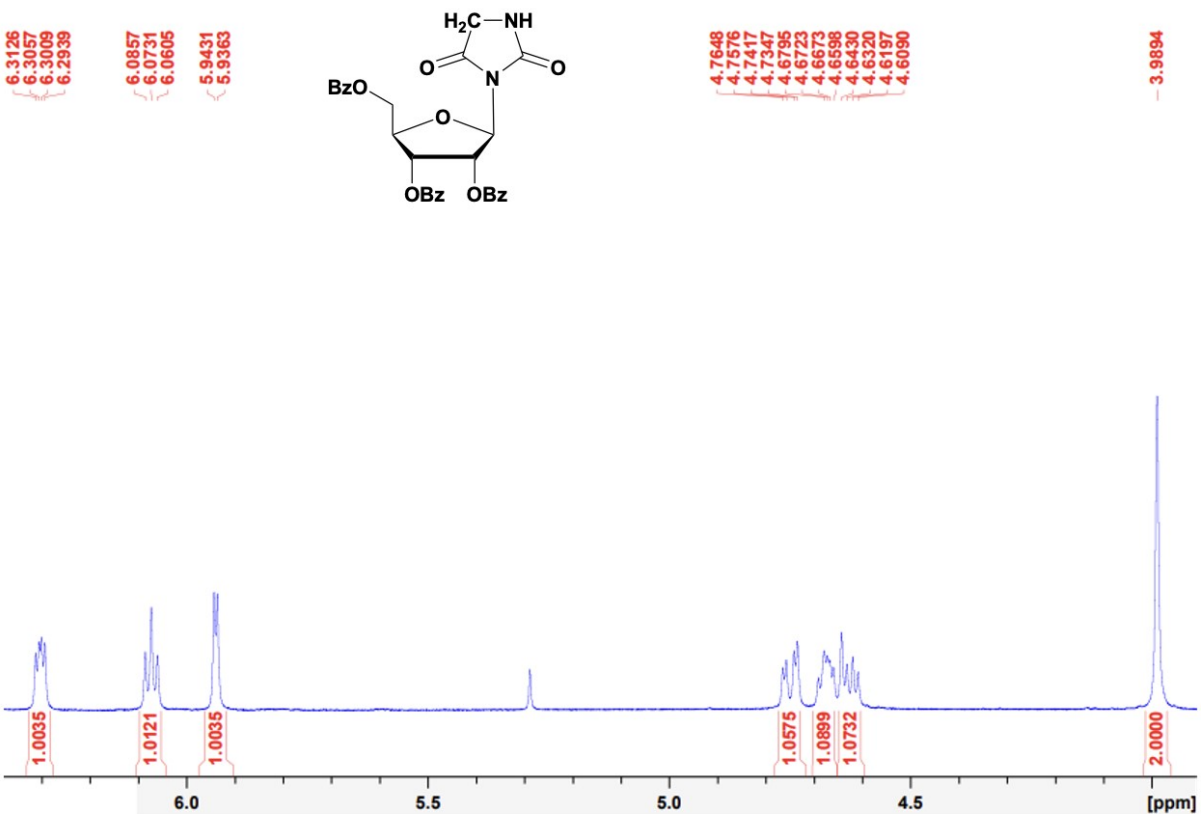
Figures

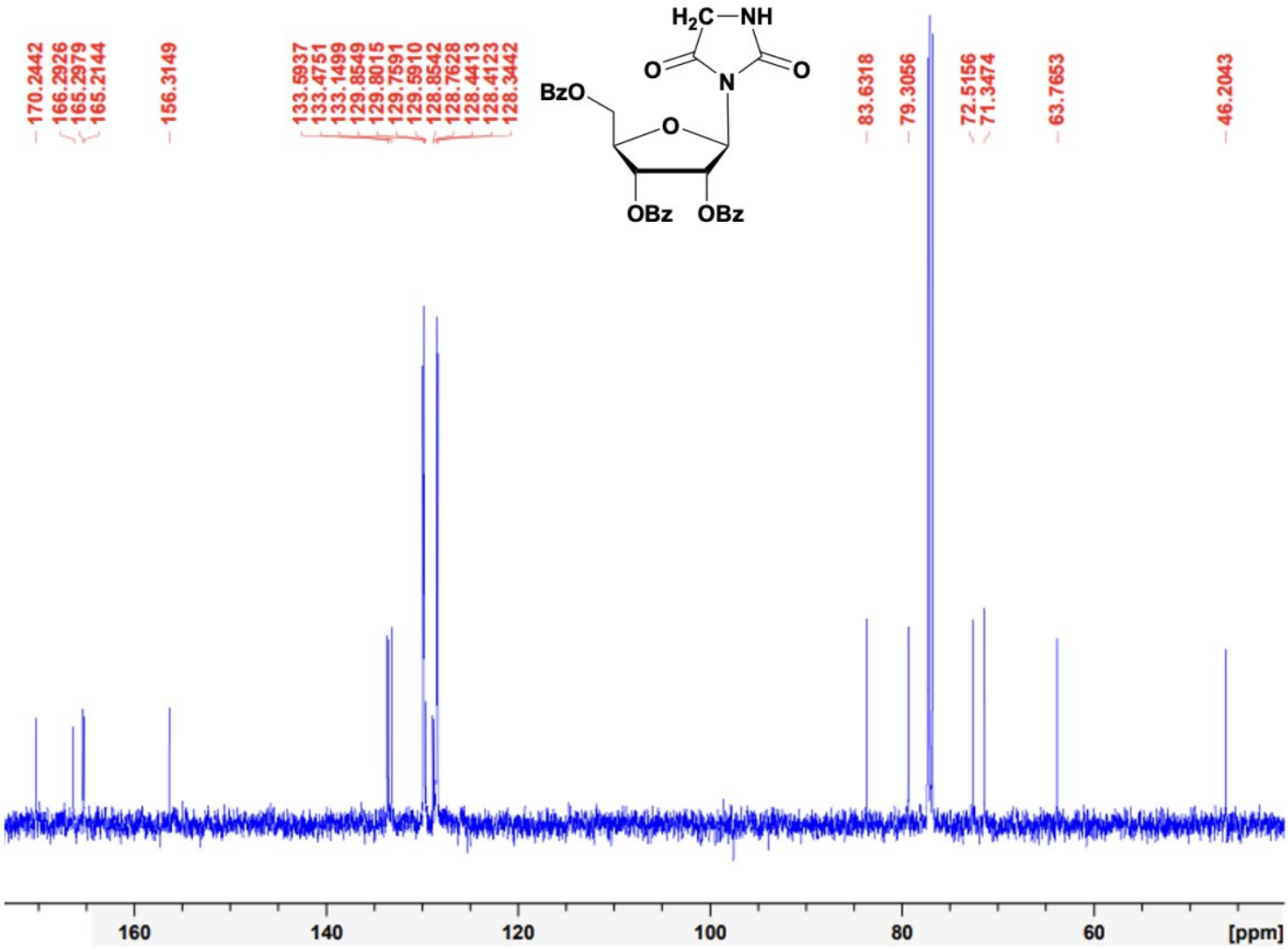
Fig. S1**.** Molecular recognition of the uracil ring in uridine by proteins (A)[1] and expansion of the recognition pattern by uridine mimics (B).[2]

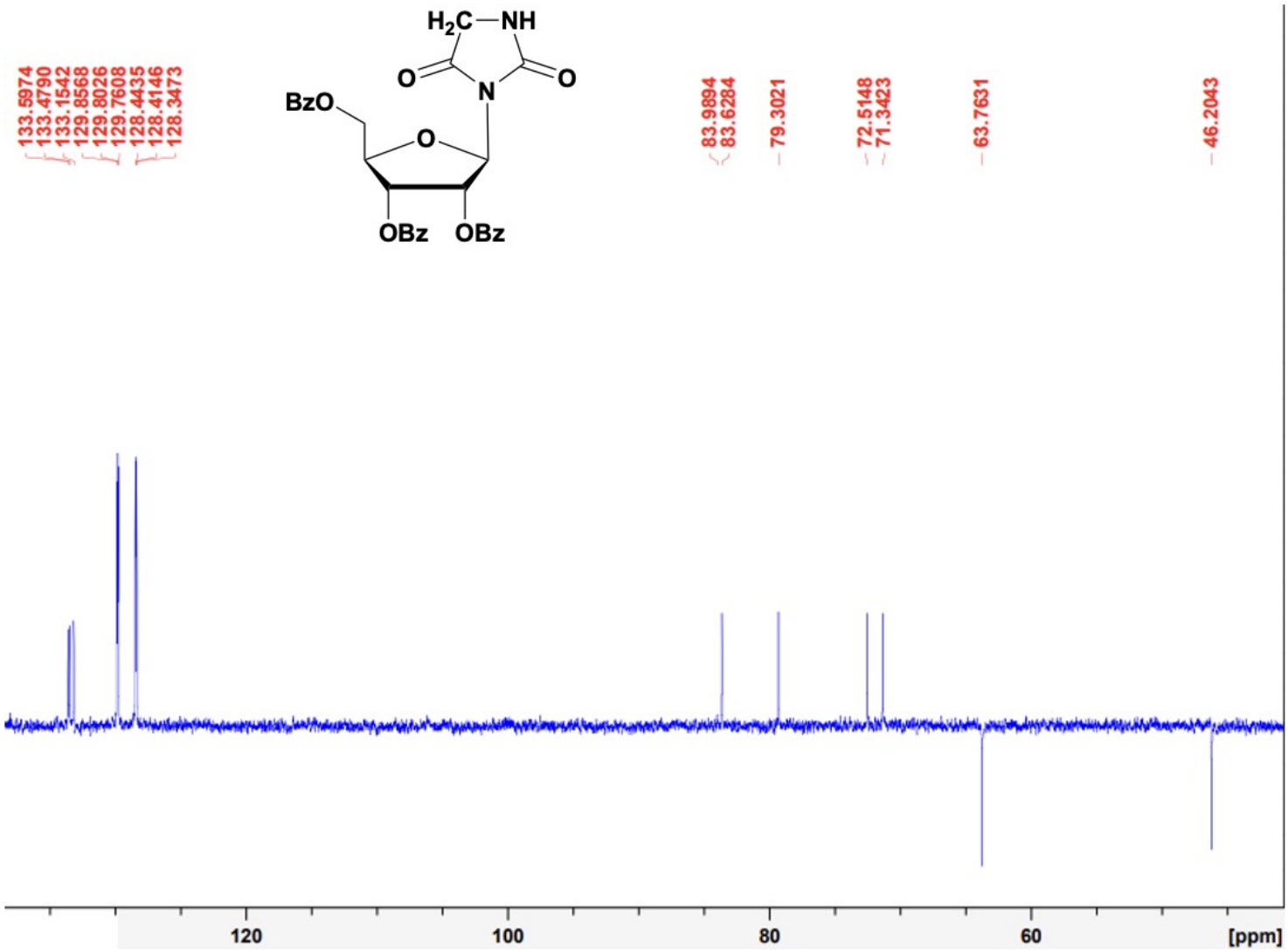
8.0902
8.0751
7.9510
7.9329
7.9146
7.5436
7.5287
7.5141
7.4109
7.3955
7.3802
7.3618
7.3471
7.3323
6.3126
6.3057
6.3009
6.2939
6.0857
6.0731
6.0605
5.9431
5.9363
4.7648
4.7576
4.7417
4.7347
4.6795
4.6723
4.6673
4.6598
4.6430
4.6320
4.6197
4.6090
3.9894






6.3126
6.3057
6.3009
6.2939
6.0857
6.0731
6.0605
5.9431
5.9363
4.7648
4.7576
4.7417
4.7347
4.6795
4.6723
4.6673
4.6598
4.6430
4.6320
4.6197
4.6090
3.9894
H2C—NH
O
O
N
BzO
O
OBz
OBz
1.0035
1.0121
1.0035
1.0575
1.0899
1.0732
2.0000
6.0
5.5
5.0
4.5
[ppm]


170.2442
166.2926
165.2979
165.2144
156.3149
133.5937
133.4751
133.1499
129.8549
129.8015
129.7591
129.5910
128.8542
128.7628
128.4413
128.4123
128.3442
83.6318
79.3056
72.5156
71.3474
63.7653
46.2043
H2C
NH
O
O
N
BzO
O
OBz
OBz
160
140
120
100
80
60
[ppm]


133.5974
133.4790
133.1542
129.8568
129.8026
129.7608
128.4435
128.4146
128.3473
83.9894
83.6284
79.3021
72.5148
71.3423
63.7631
46.2043
H2C
NH
O
O
N
BzO
O
OBz
OBz
120
100
80
60
[ppm]

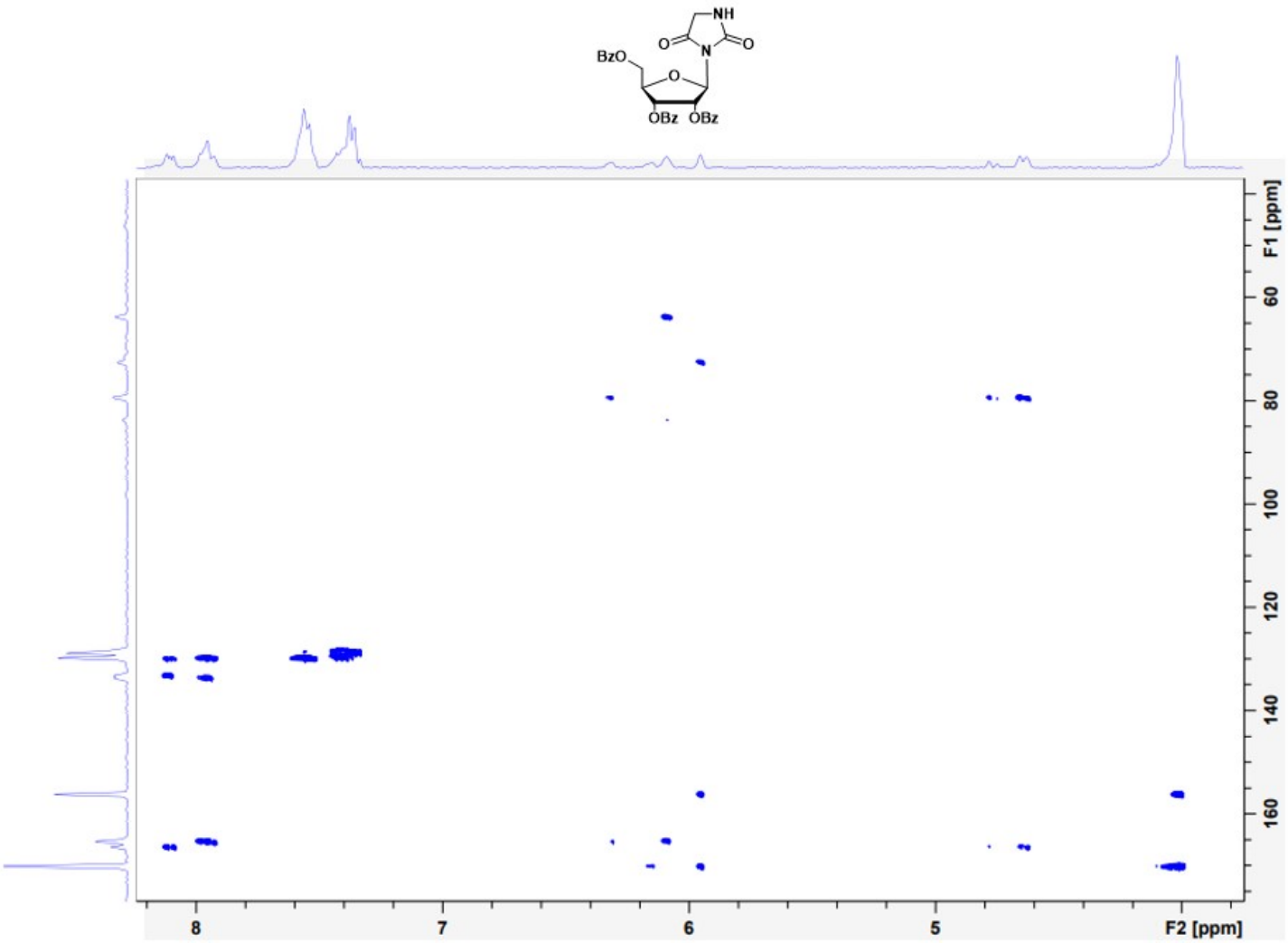
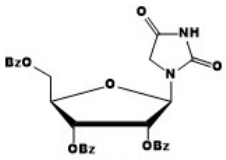
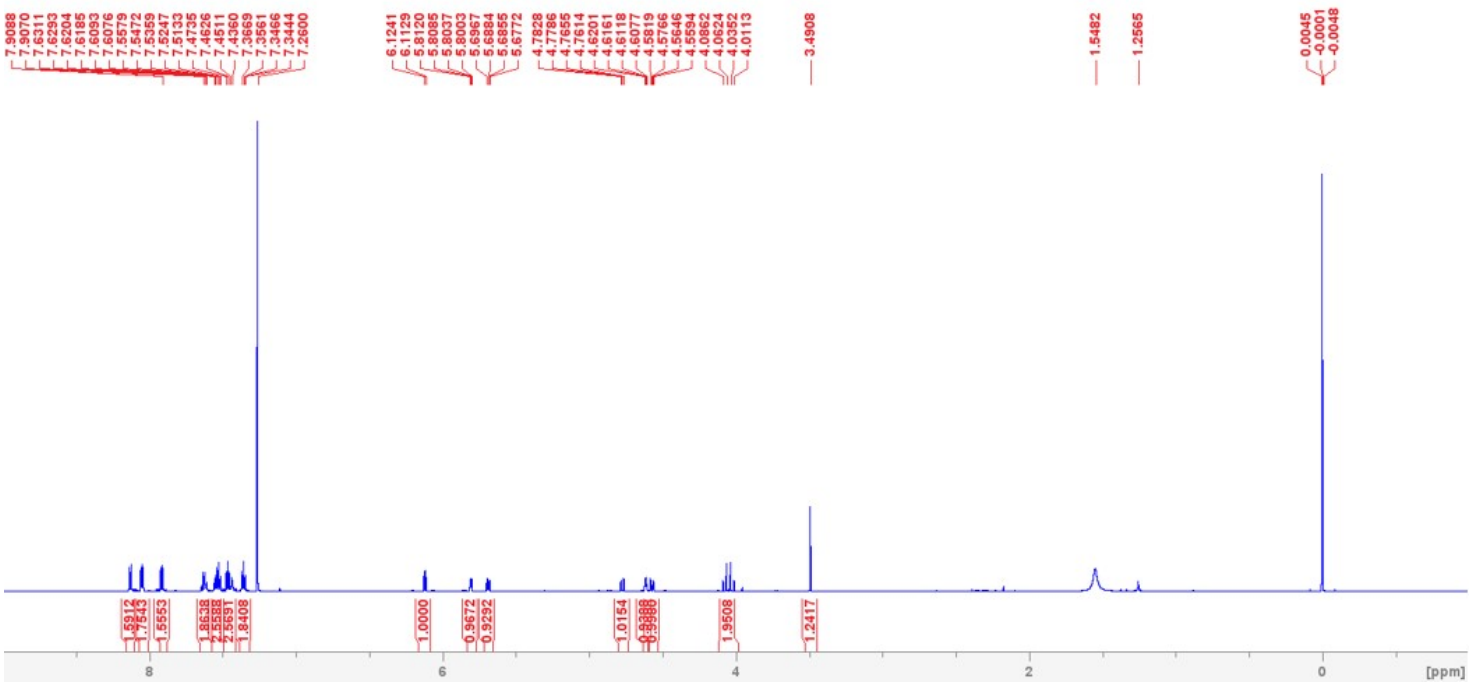
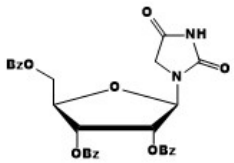
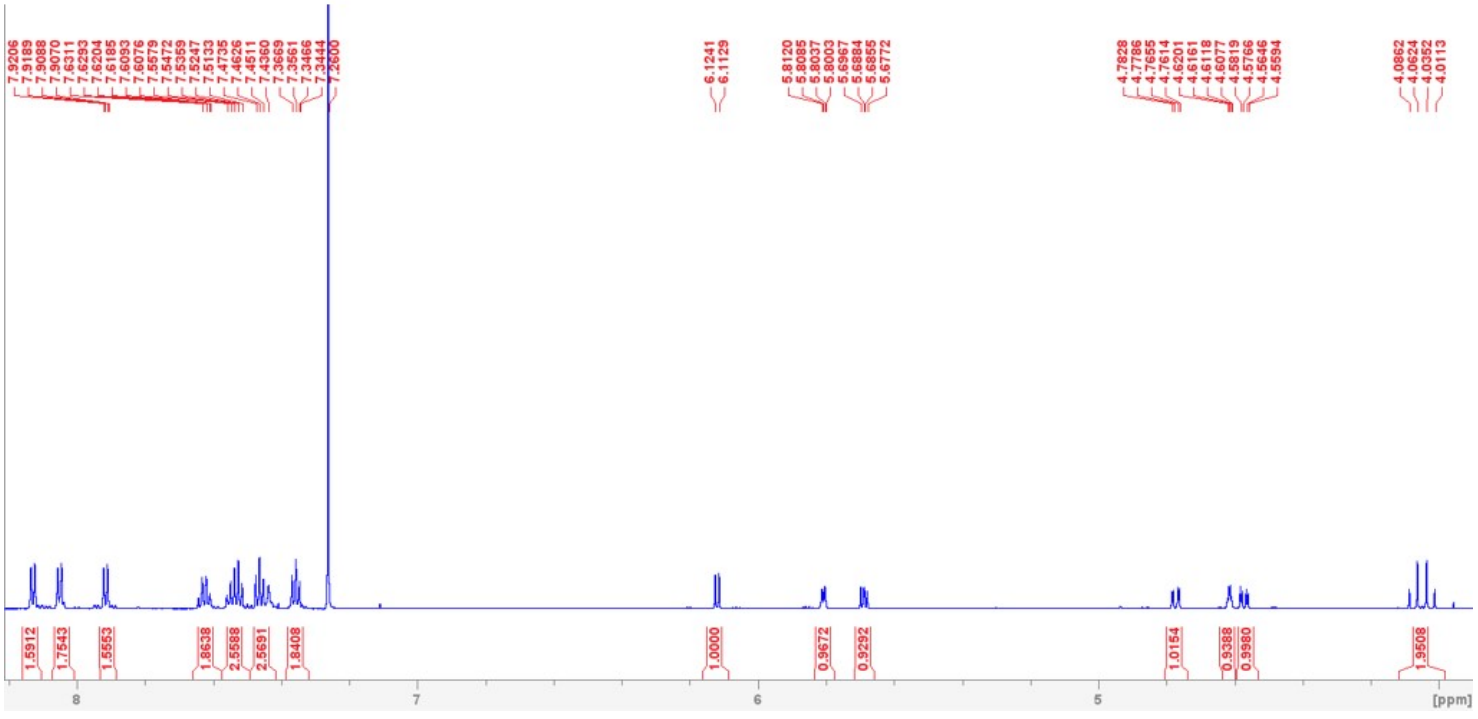


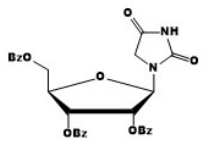
Fig. S2. 700 MHz ^{1}H-NMR/^{13}C-NMR/DEPT/HMBC spectra in $CDCl_3$, 300 K, of (2*R*,3*R*,4*R*,5*R*)-2-((benzoyloxy)methyl)-5-(2,5-dioxoimidazolidin-1-yl)tetrahydrofuran-3,4-diyl dibenzoate (8).

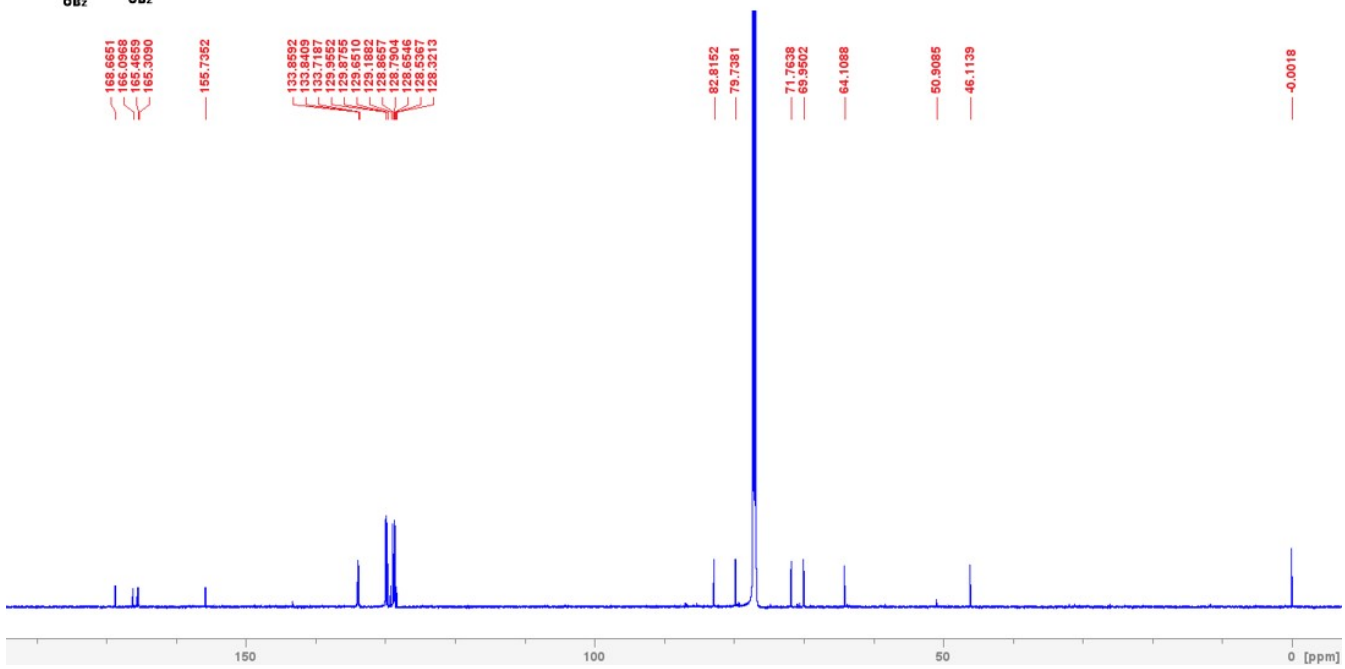

O
NH
BzO
O
N
O
OBz
OBz

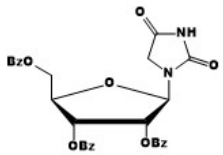

7.9088
7.9070
7.6311
7.6293
7.6204
7.6185
7.6093
7.6076
7.5579
7.5472
7.5359
7.5247
7.5133
7.4735
7.4626
7.4511
7.4360
7.3669
7.3561
7.3466
7.3444
7.2600
6.1241
6.1129
5.8120
5.8085
5.8037
5.8003
5.6967
5.6884
5.6855
5.6772
4.7828
4.7786
4.7655
4.7614
4.6201
4.6161
4.6118
4.6077
4.5819
4.5766
4.5646
4.5594
4.0862
4.0624
4.0352
4.0113
3.4908
1.5482
1.2565
0.0045
-0.0001
-0.0048
1.5912
1.7543
1.5553
1.8638
2.5588
2.5691
1.8408
1.0000
0.9672
0.9292
1.0154
0.9388
0.9980
1.9508
1.2417
8
6
4
2
0
[ppm]


O
NH
BzO
O
N
O
OBz
OBz


7.9206
7.9189
7.9088
7.9070
7.6311
7.6293
7.6204
7.6185
7.6093
7.6076
7.5579
7.5472
7.5359
7.5247
7.5133
7.4735
7.4626
7.4511
7.4360
7.3669
7.3561
7.3466
7.3444
7.2600
6.1241
6.1129
5.8120
5.8085
5.8037
5.8003
5.6967
5.6884
5.6855
5.6772
4.7828
4.7786
4.7655
4.7614
4.6201
4.6161
4.6118
4.6077
4.5819
4.5766
4.5646
4.5594
4.0862
4.0624
4.0352
4.0113
1.5912
1.7543
1.5553
1.8638
2.5588
2.5691
1.8408
1.0000
0.9672
0.9292
1.0154
0.9388
0.9980
1.9508
8
7
6
5
[ppm]


BzO
NH
O
N
OBz
OBz


168.6651
166.0968
165.4659
165.3090
155.7352
133.8592
133.8409
133.7187
129.9552
129.8755
129.6510
129.1882
128.8657
128.7904
128.6546
128.5367
128.3213
82.8152
79.7381
71.7638
69.9502
64.1088
50.9085
46.1139
-0.0018
150
100
50
0 [ppm]


BzO
NH
O
N
OBz
OBz

133.8573
133.8386
133.7164
129.9527
129.8733
129.6491
128.8636
128.6524
128.5754
128.5347
82.8143
79.7367
79.2285
77.2172
71.7627
69.9499
64.1087
46.1146
150
100
50
0 [ppm]

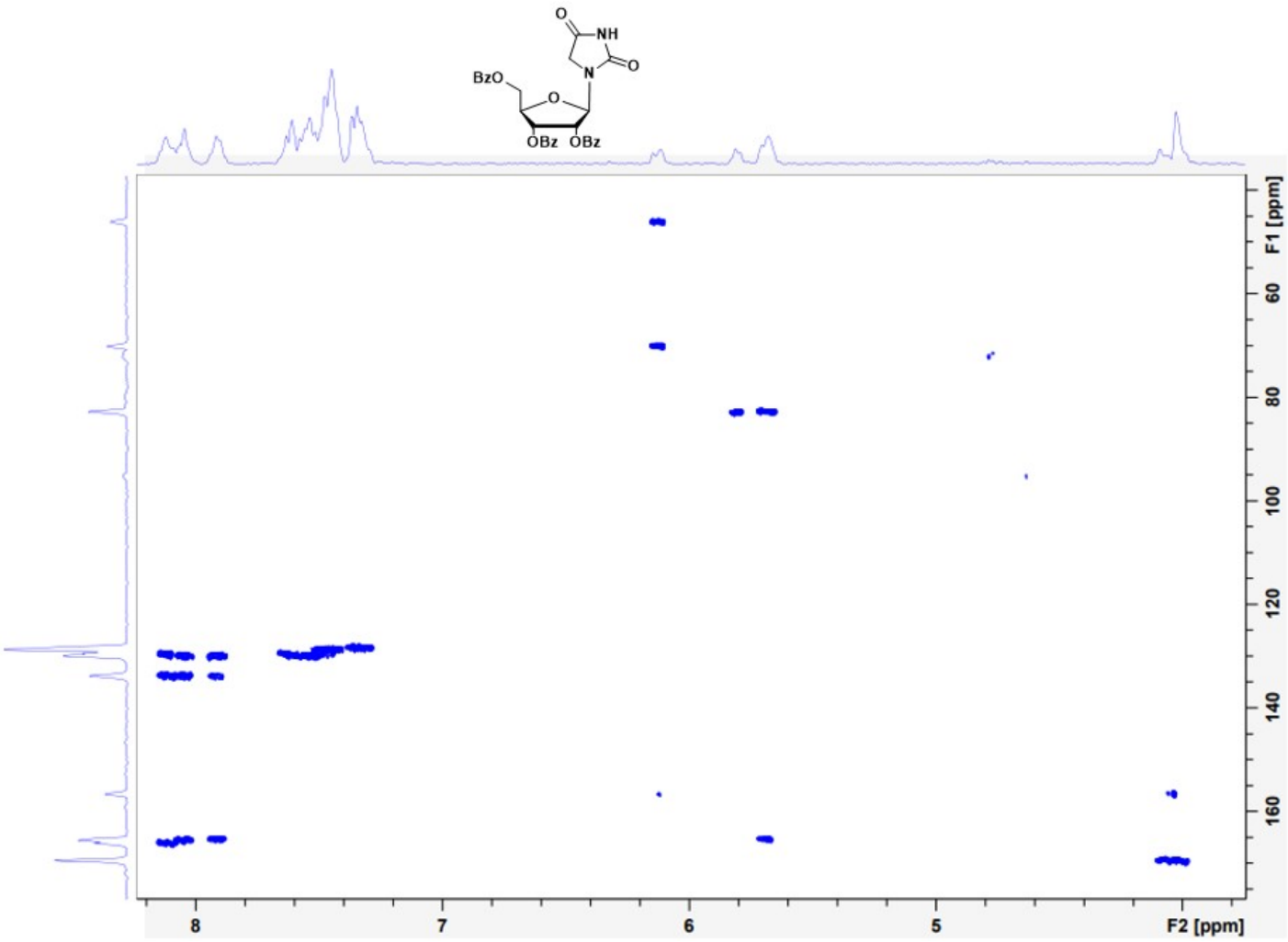
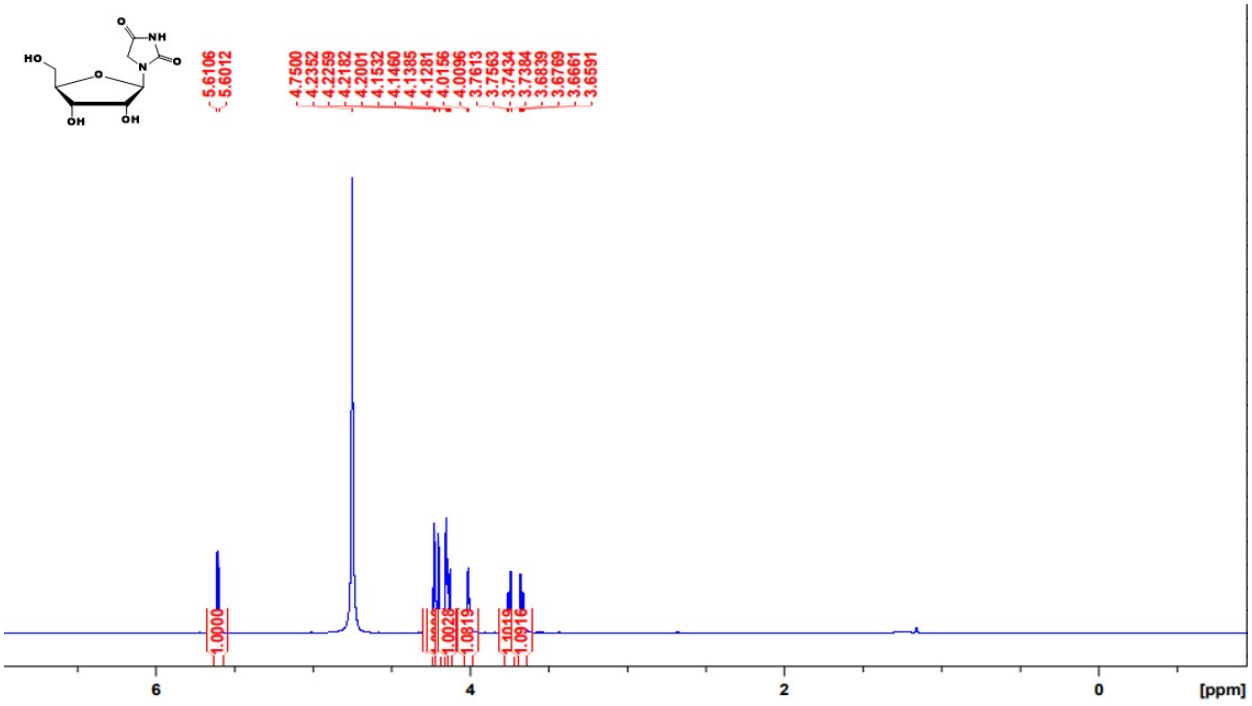
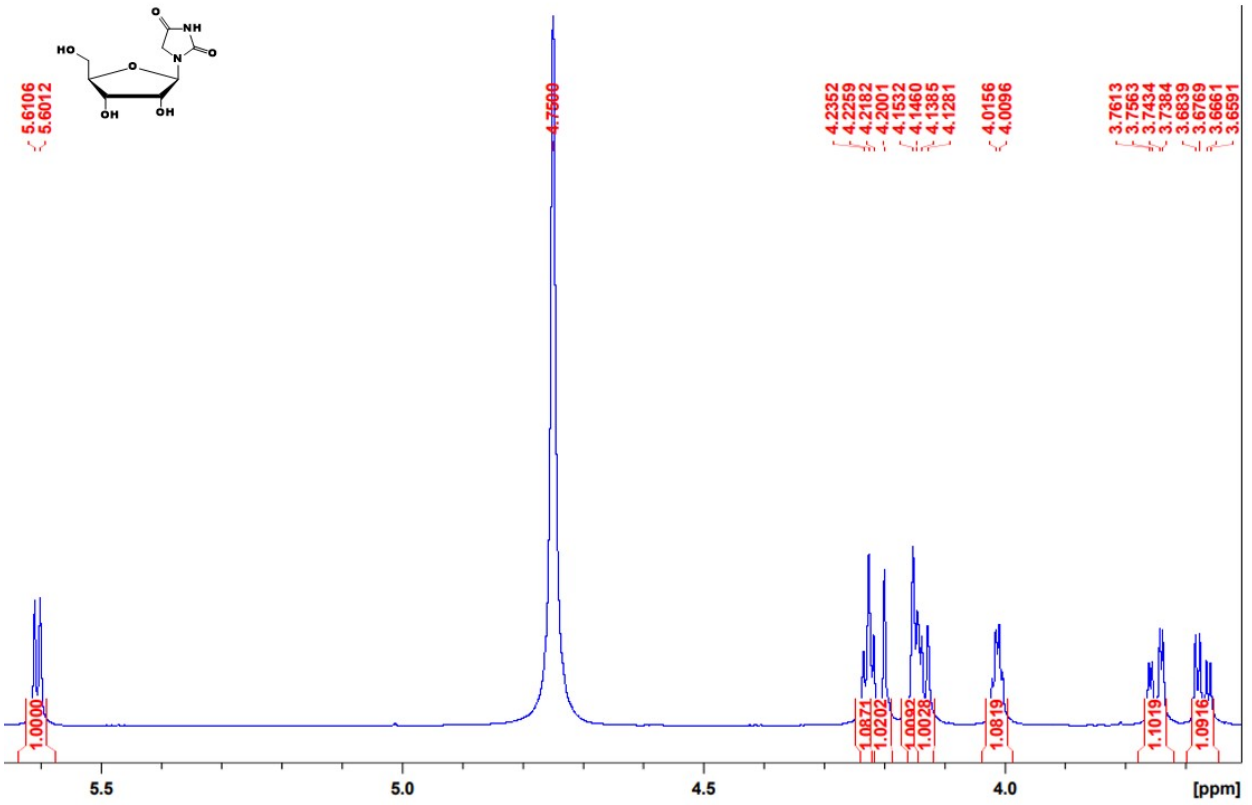


Fig. S3. 700 MHz ^{1}H-NMR/^{13}C-NMR/ DEPT/HMBC spectra in $CDCl_3$, 300 K, of (2R,3R,4R,5R)-2-((Benzoyloxy) methyl)-5-(2,4-dioxoimidazolidin-1yl)tetrahydrofuran-3,4-diyldibenzoate (9).


5.6106
5.6012
4.7500
4.2352
4.2259
4.2182
4.2001
4.1532
4.1460
4.1385
4.1281
4.0156
4.0096
3.7613
3.7563
3.7434
3.7384
3.6839
3.6769
3.6661
3.6591
1.0000
1.0028
1.0819
1.1019
1.0916
6
4
2
0
[ppm]


5.6106
5.6012
4.7500
4.2352
4.2259
4.2182
4.2001
4.1532
4.1460
4.1385
4.1281
4.0156
4.0096
3.7613
3.7563
3.7434
3.7384
3.6839
3.6769
3.6661
3.6591
1.0000
1.0871
1.0202
1.0092
1.0028
1.0819
1.1019
1.0916
5.5
5.0
4.5
4.0
[ppm]

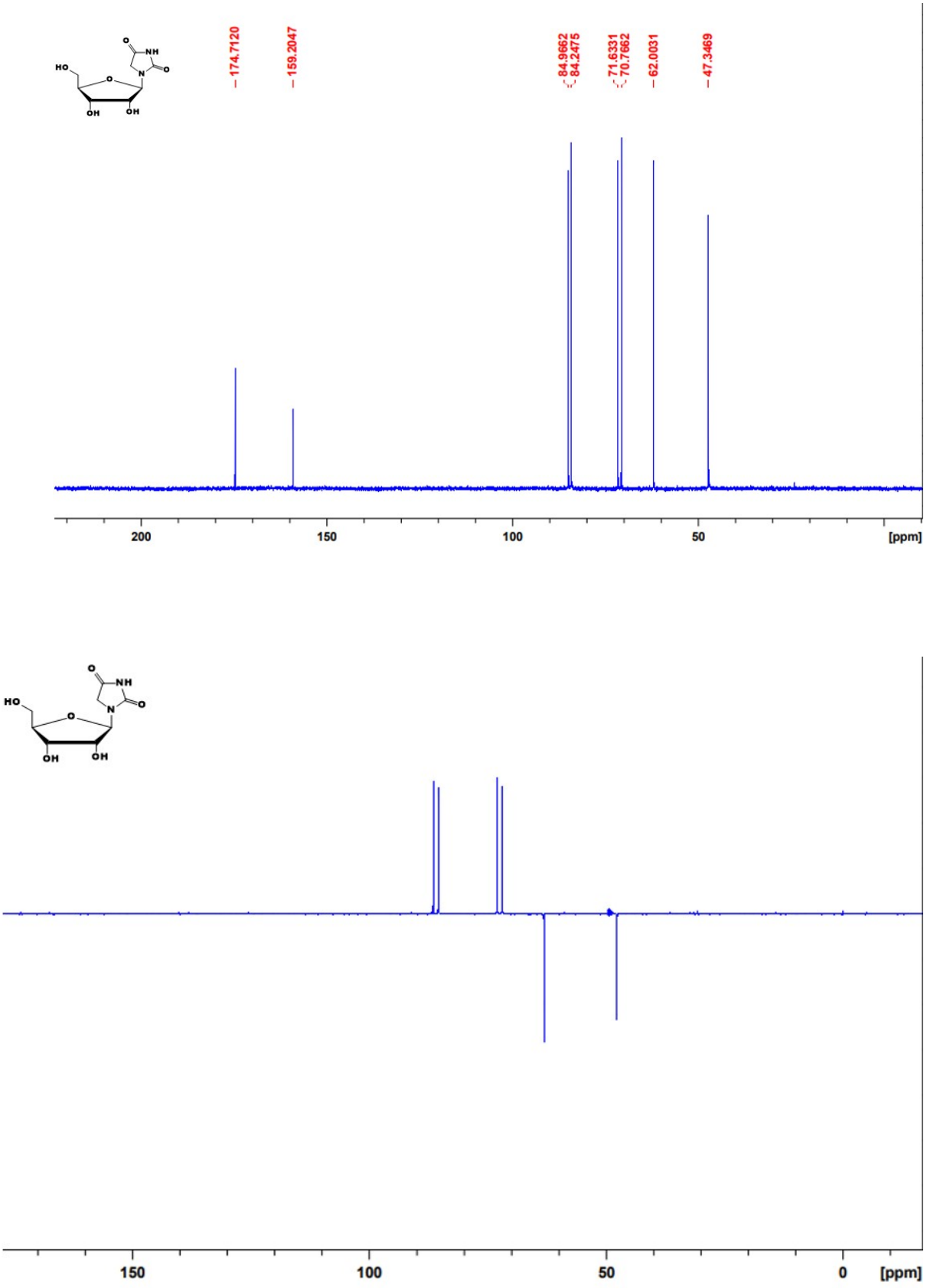


Fig. S4. 700 MHz ^{1}H-NMR/^{13}C-NMR/ DEPT spectra in D_2O, 300 K, of Hydantoinyl Ribose (15).

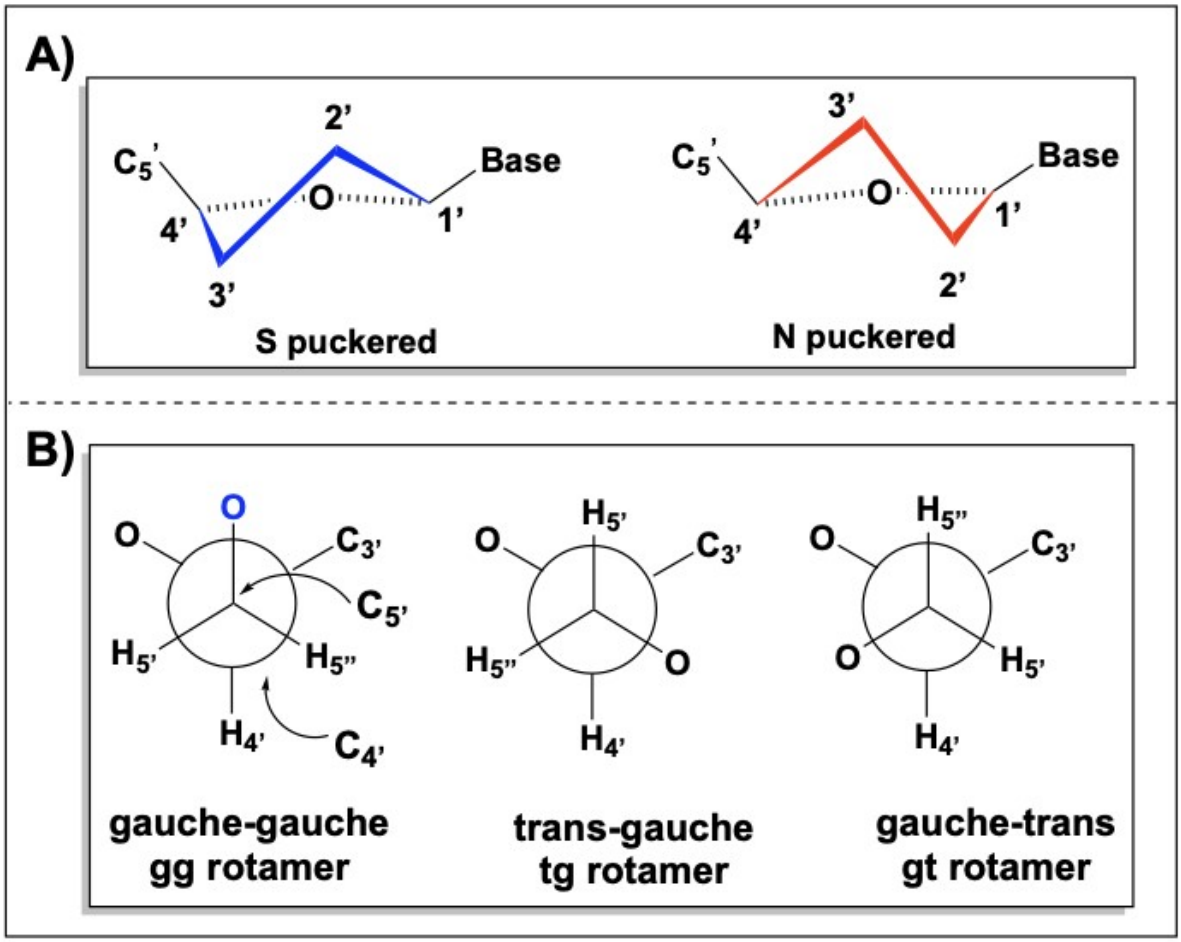


Fig. S5. Solution conformation of ribose in nucleoside is composed of equilibrating *N* and *S* conformers (A). Conformation of exocyclic CH_2OH is composed of three equilibrating rotamers, *gg, gt,* and *tg* (B).

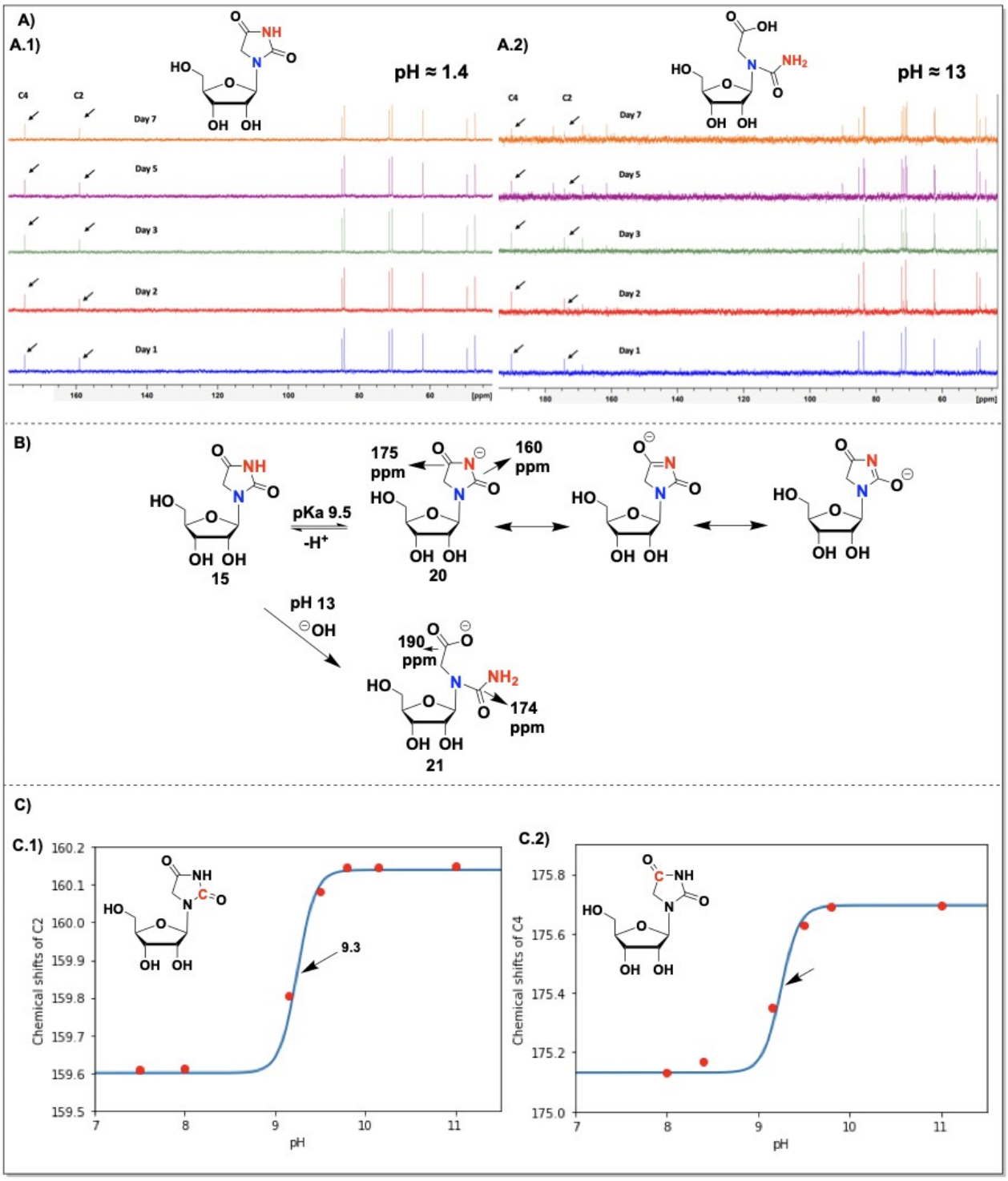
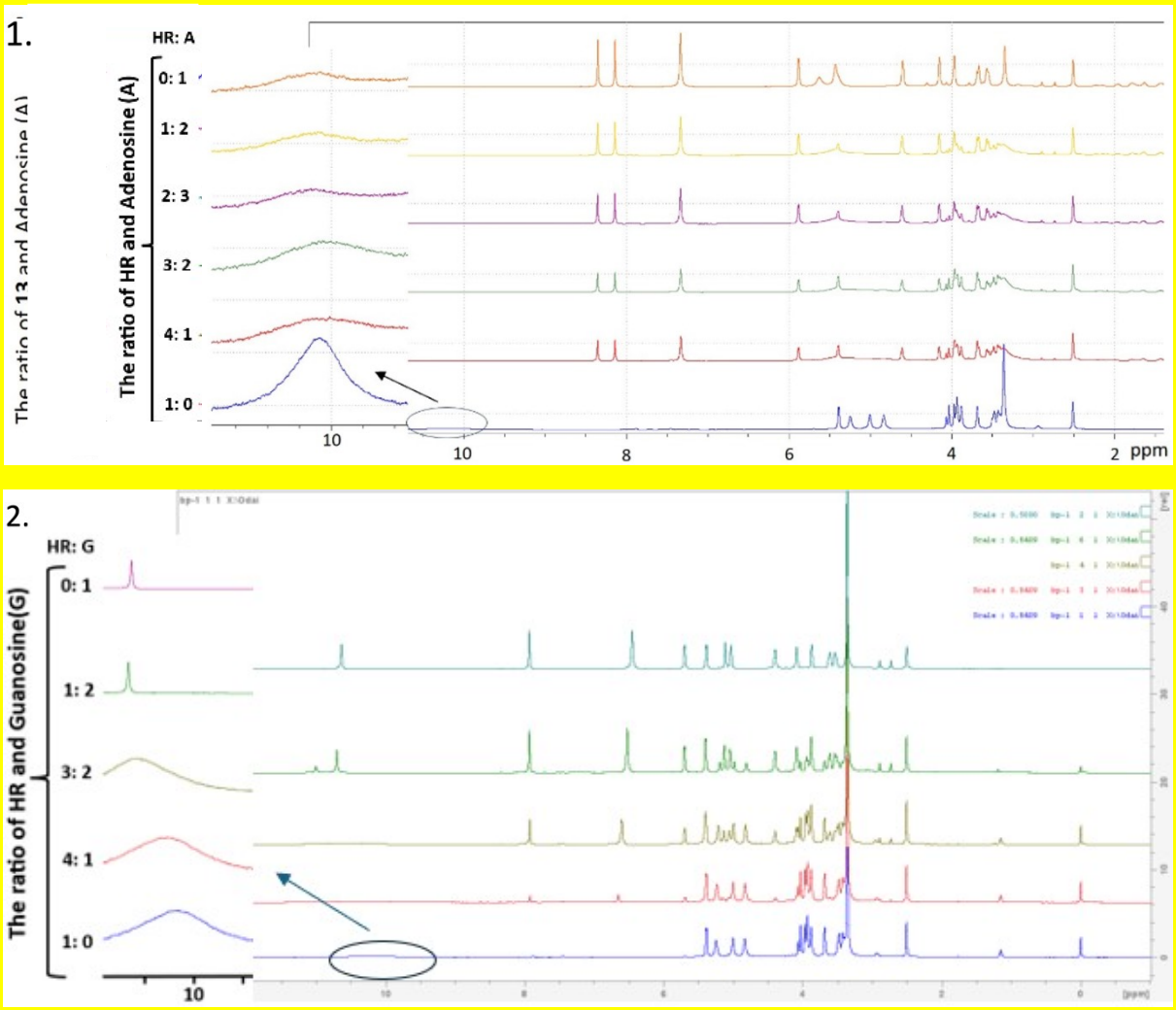


Fig. S6. pH-dependent hydrolytic stability of HR at RT under A.1) pH 1.4 A.2) pH 13, monitored by ^{13}C NMR (A), Ionization and hydrolysis of HR under basic pH (B), Sigmoidal dependance of the chemical shift of C.1) C2 and C.2) C4 on pH (C).

A)



B)

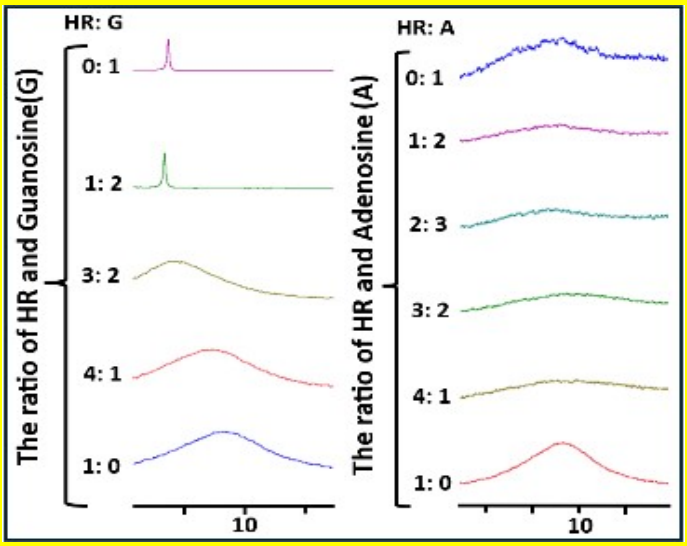
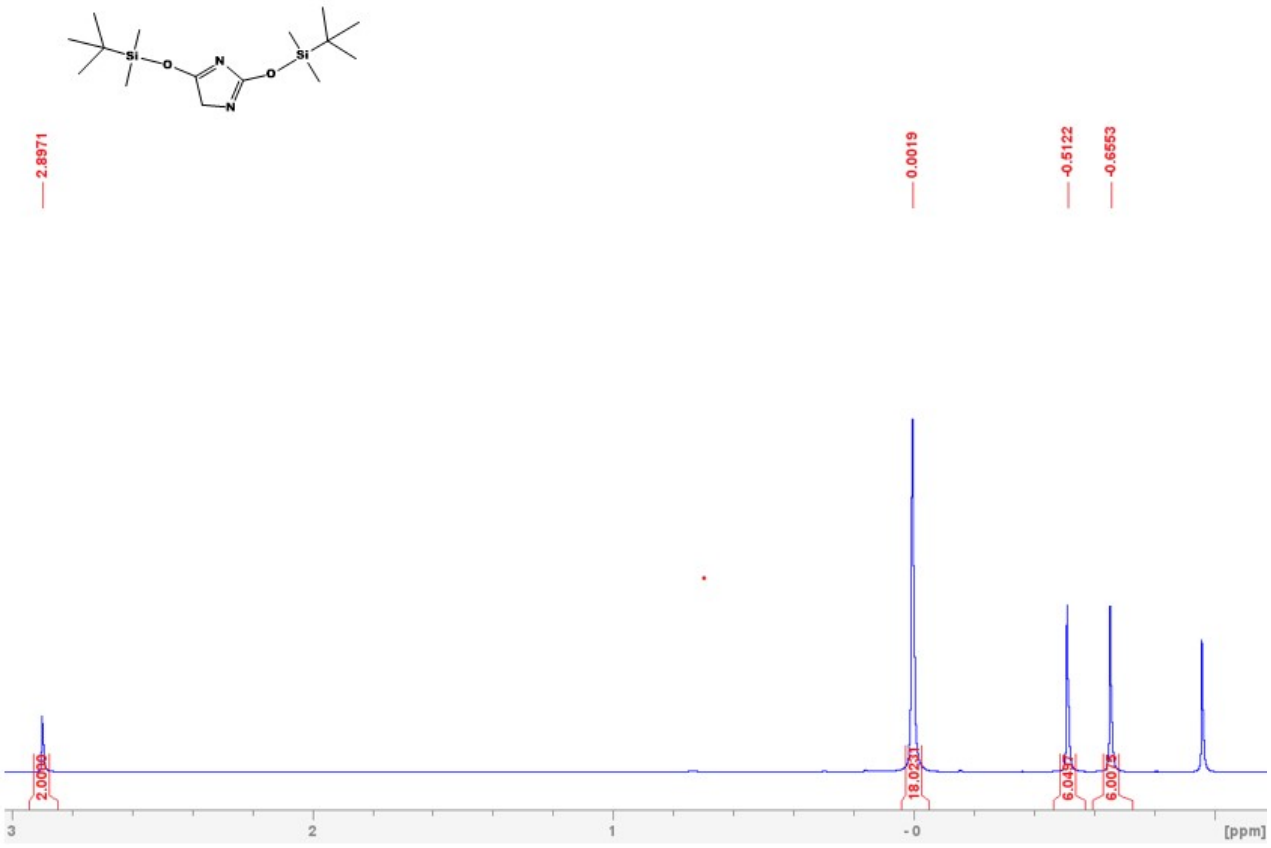


Fig. S7. 0.1 M HR and A.1) 0.1 M adenosine or A.2) 0.1 M guanosine in DMSO-d_6. ^{1}H NMR spectra were measured in dry DMSO-d_6 at 500 MHz. The data were collected at 300 K. ^{1}H-NMR spectra were obtained at a range of HR:A ratios (A), Enlarged view of Fig. S7.A.1/2 (B).


2.8971
0.0019
-0.5122
-0.6553
2.0000
18.0231
6.0432
6.0073
3
2
1
-0
[ppm]

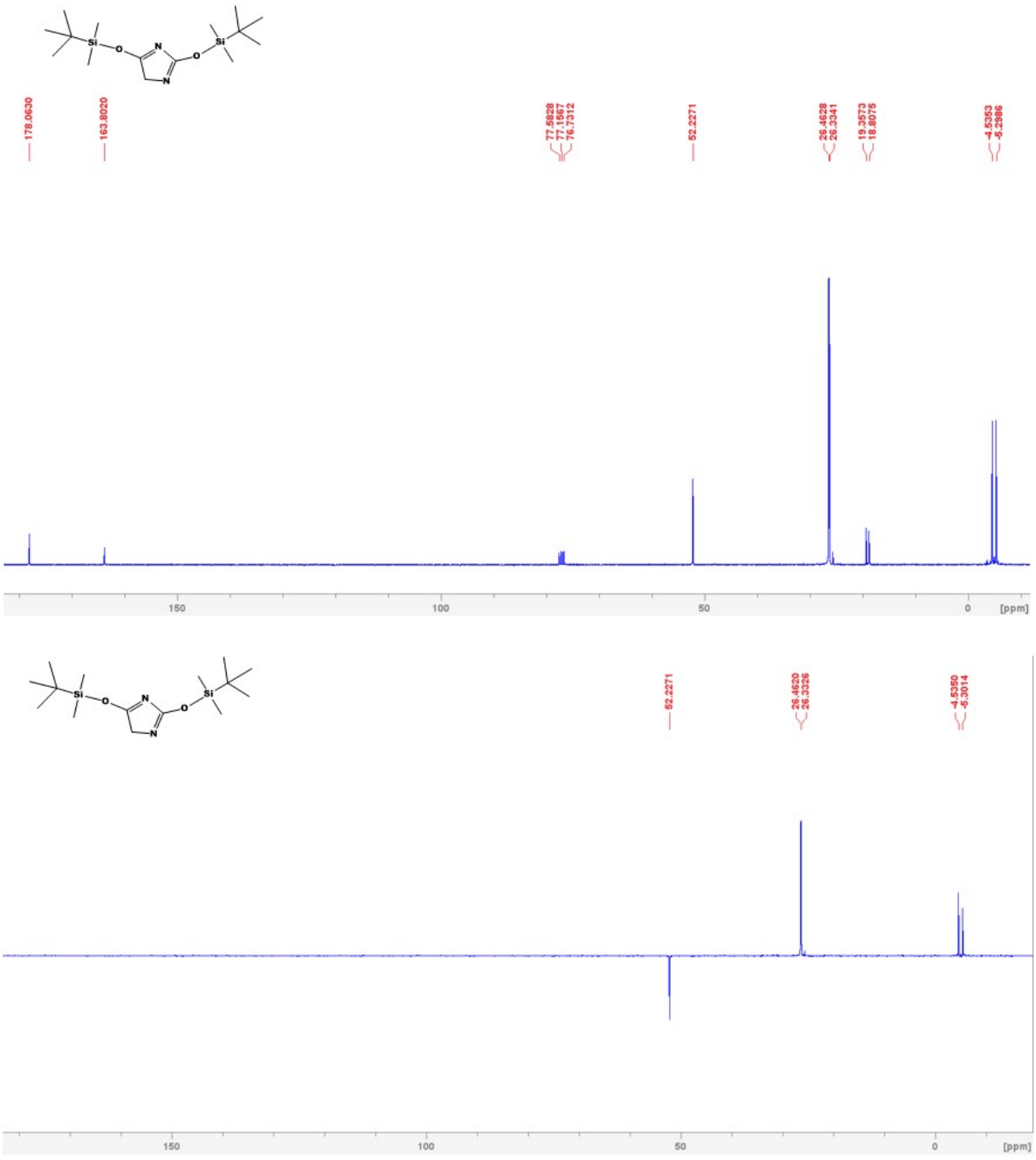
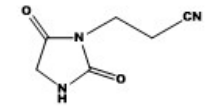
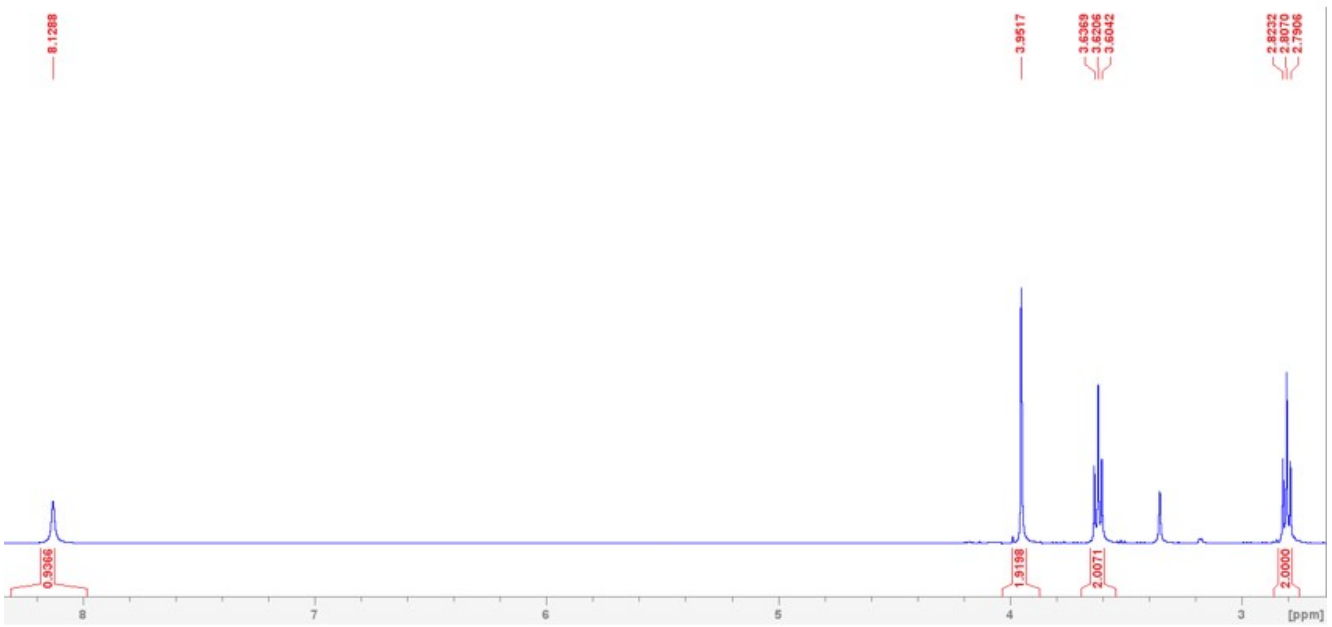
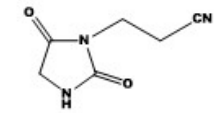
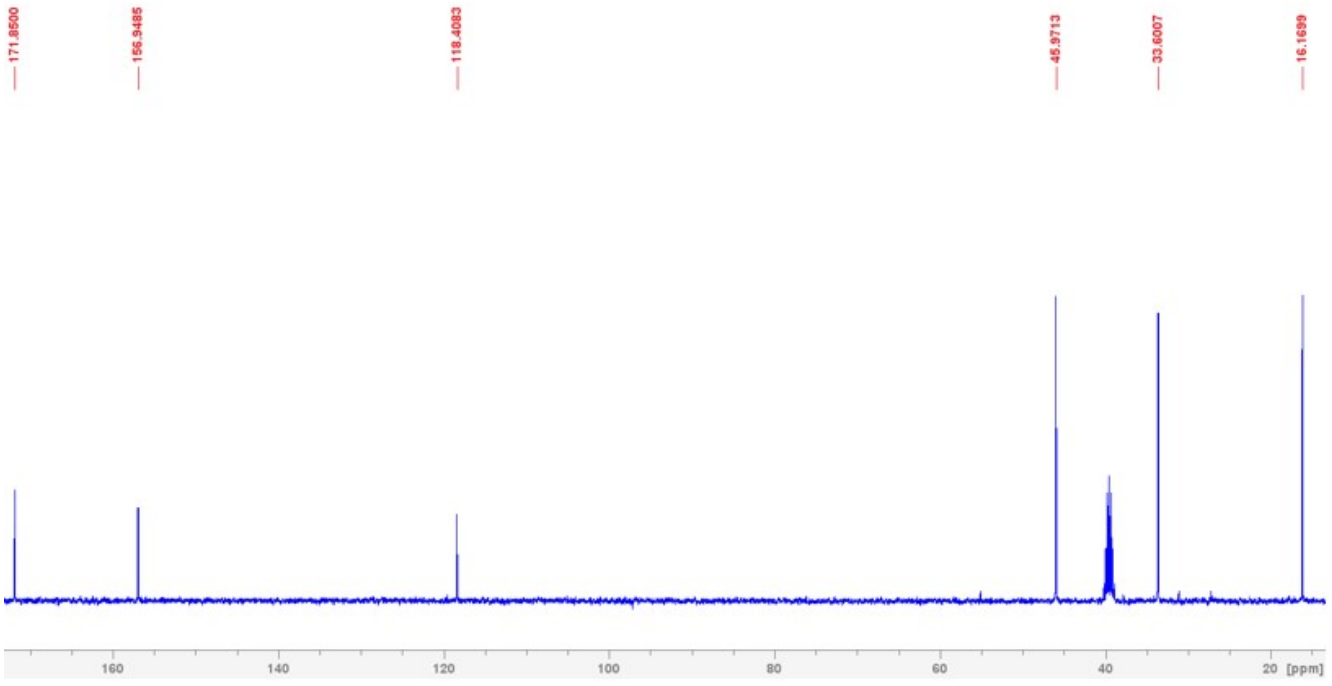


Fig. S8. 400 MHz ^{1}H-NMR/^{13}C-NMR/ DEPT spectra in $CDCl_3$, 300 K, of 2,5-Bis((tert butyldimethylsilyl)oxy)- 4H-imidazole (16).


O
N
CN
O
N
H


8.1288
3.9517
3.6369
3.6206
3.6042
2.8232
2.8070
2.7906
0.9366
1.9158
2.0071
2.0000
8
7
6
5
4
3
[ppm]


O
N
CN
O
N
H


171.8500
156.9485
118.4083
45.9713
33.6007
16.1699
160
140
120
100
80
60
40
20
[ppm]

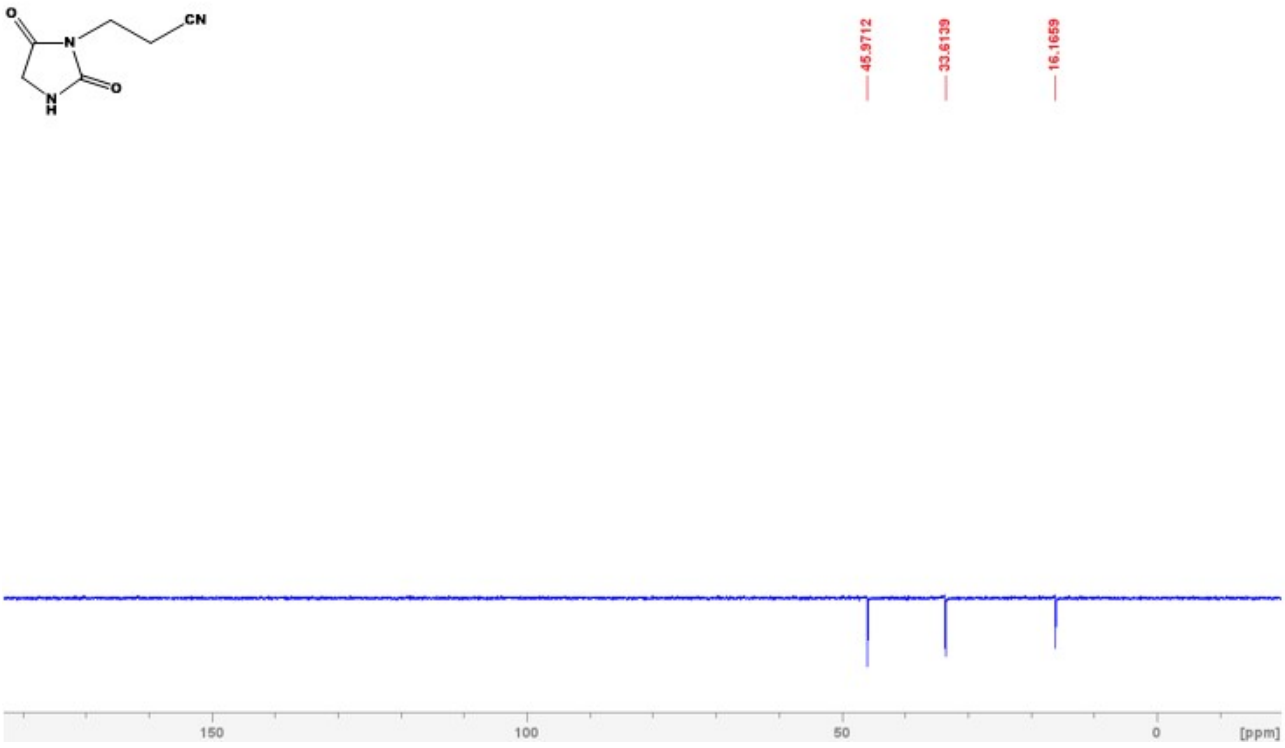
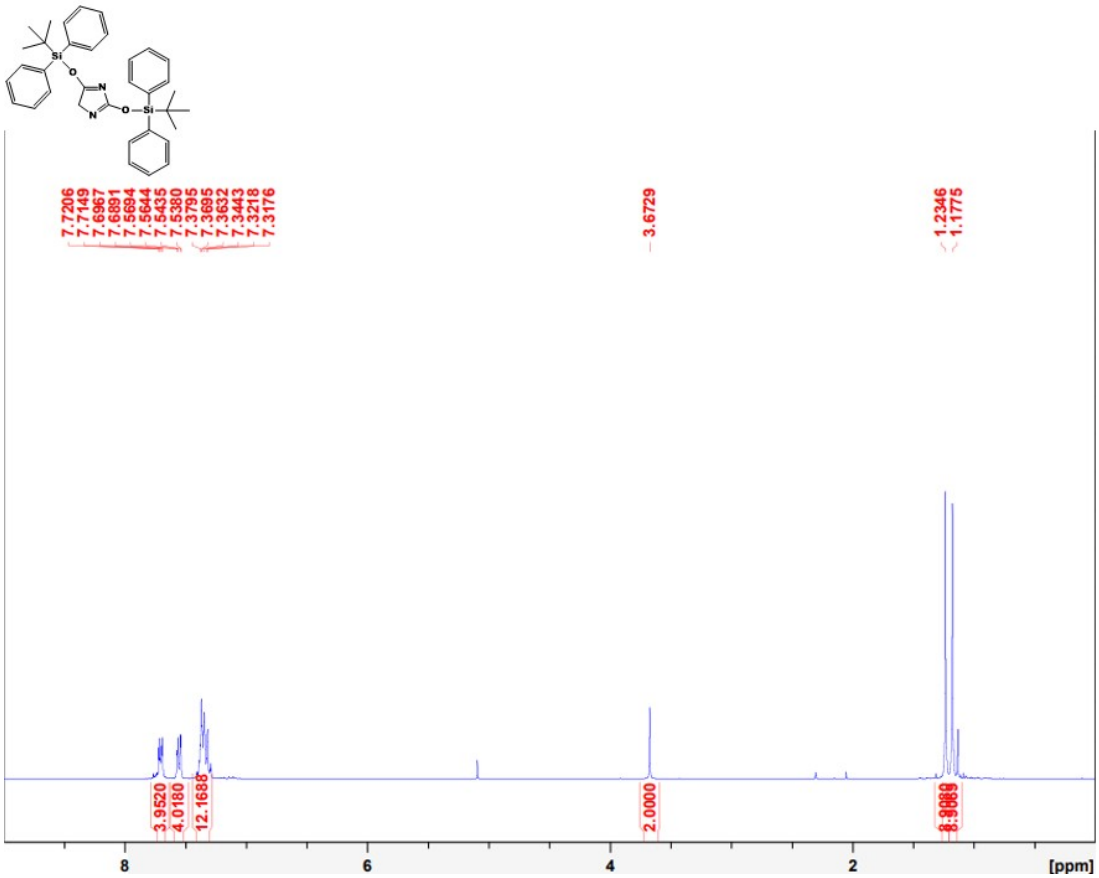
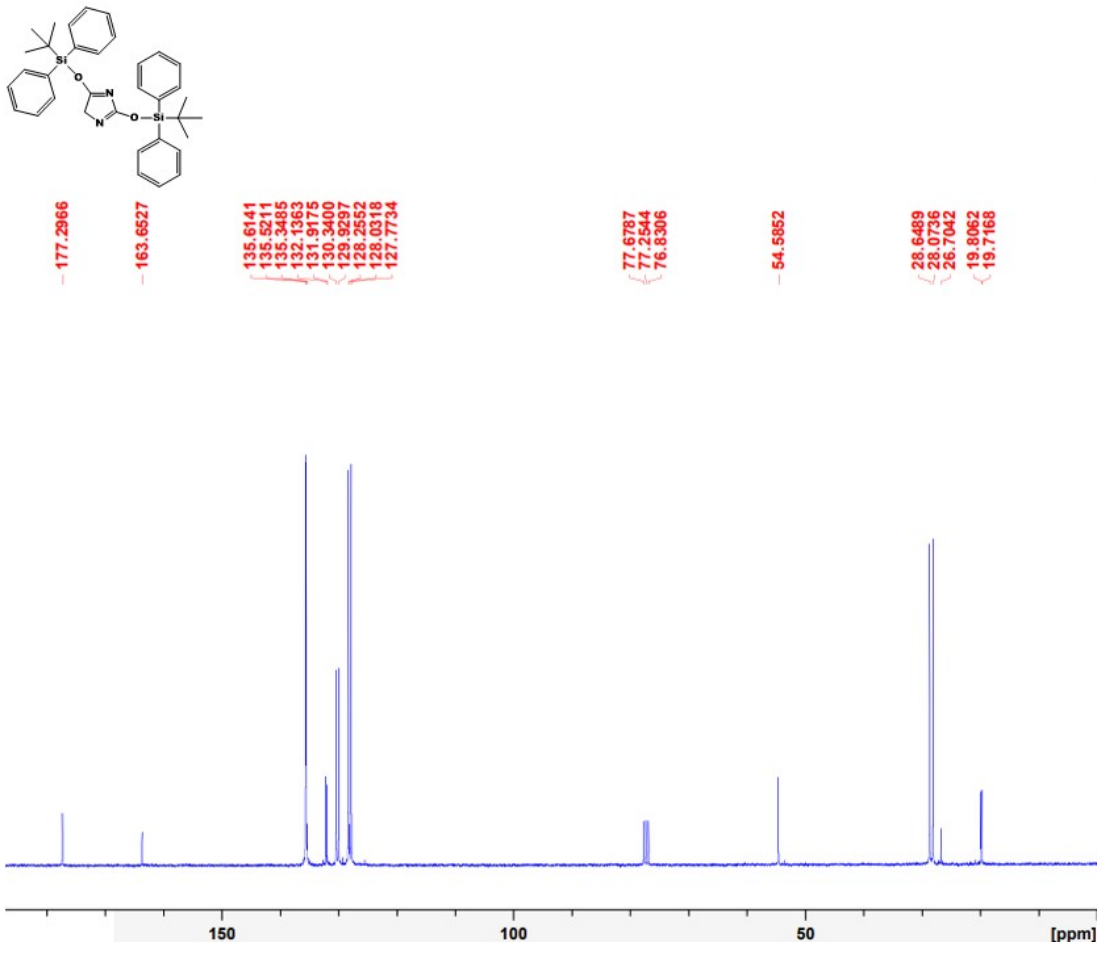


Fig. S9. 400 MHz ^{1}H-NMR/^{13}C-NMR/ DEPT spectra in DMSO, 300 K, of 3-(2,5-Dioxoimidazolidin-1-yl)propanenitrile (12).


7.7206
7.7149
7.6967
7.6891
7.5694
7.5644
7.5435
7.5380
7.3795
7.3695
7.3632
7.3443
7.3218
7.3176
3.6729
1.2346
1.1775
3.9520
4.0180
12.1688
2.0000
8
6
4
2
[ppm]


177.2966
163.6527
135.6141
135.5211
135.3485
132.1363
131.9175
130.3400
129.9297
128.2552
128.0318
127.7734
77.6787
77.2544
76.8306
54.5852
28.6489
28.0736
26.7042
19.8062
19.7168
150
100
50
[ppm]

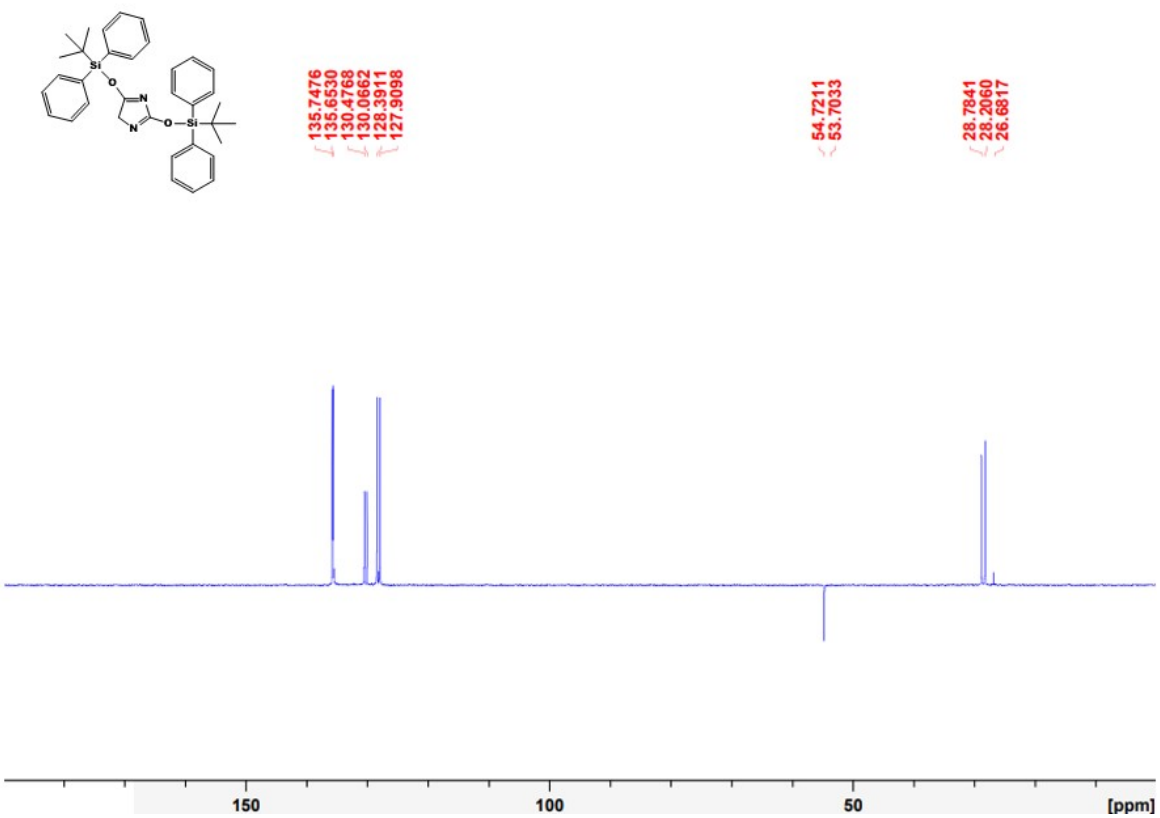
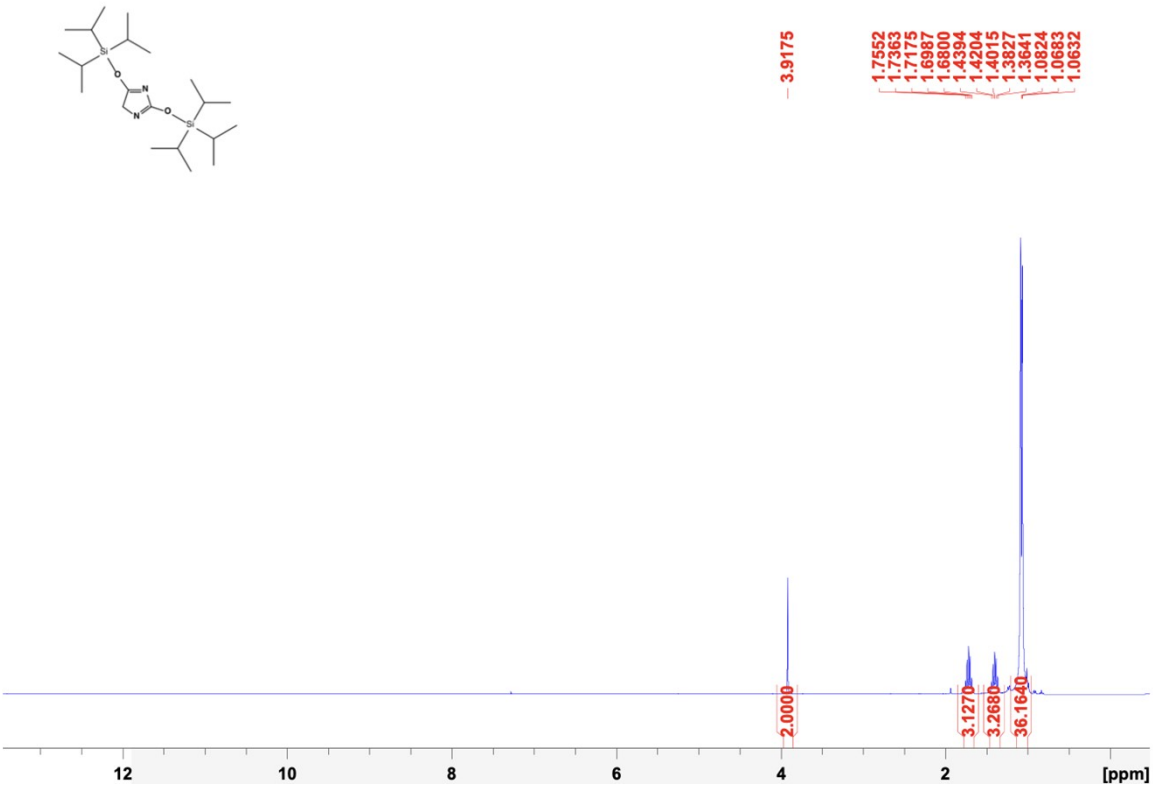
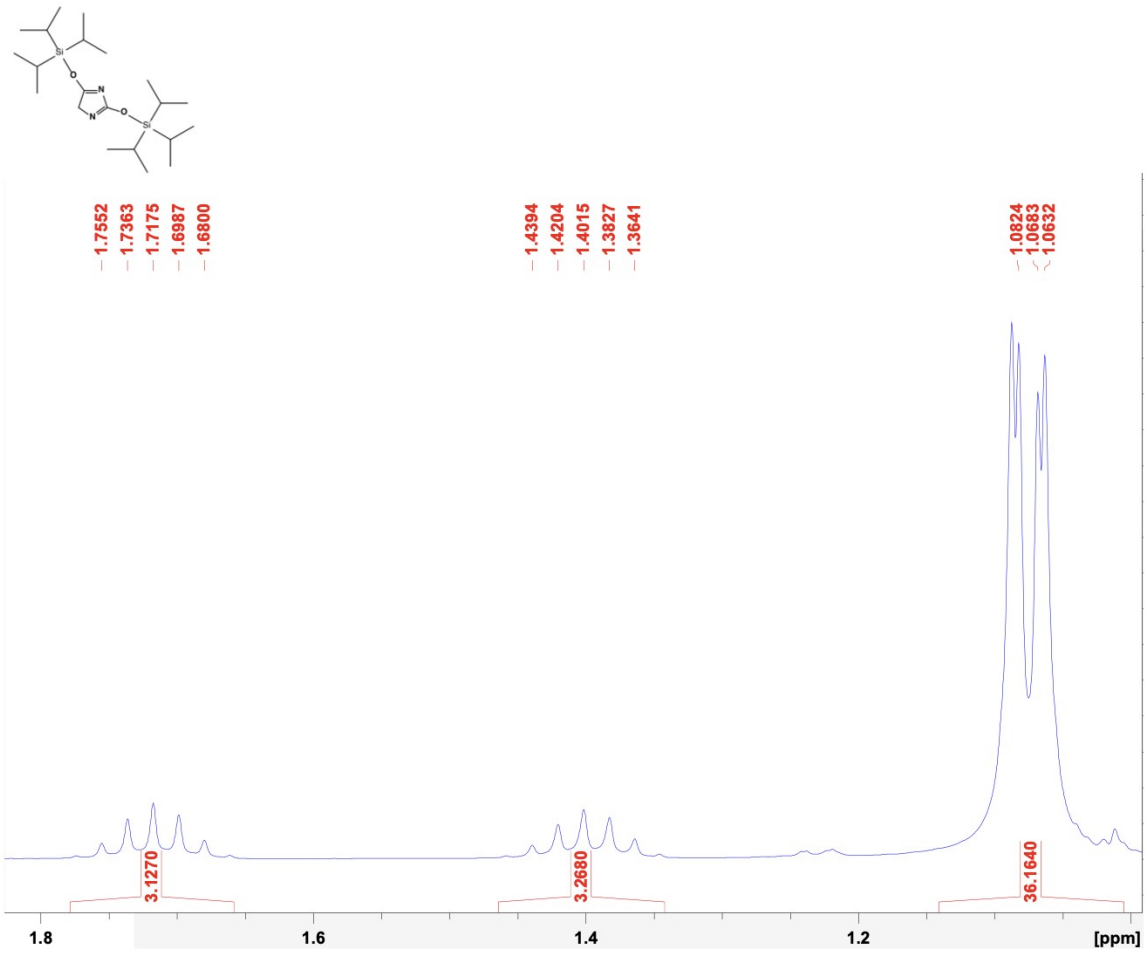


Fig. S10. 300 MHz ^{1}H-NMR/^{13}C-NMR/ DEPT spectra in $CDCl_3$, at 300 K, of 2,5-bis((*tert*-butyldiphenylsilyl) oxy)-4*H*-imidazole (17).


3.9175
1.7552
1.7363
1.7175
1.6987
1.6800
1.4394
1.4204
1.4015
1.3827
1.3641
1.0824
1.0683
1.0632
2.0000
3.1270
3.2680
36.1640
12
10
8
6
4
2
[ppm]


1.7552
1.7363
1.7175
1.6987
1.6800
1.4394
1.4204
1.4015
1.3827
1.3641
1.0824
1.0683
1.0632
3.1270
3.2680
36.1640
1.8
1.6
1.4
1.2
[ppm]

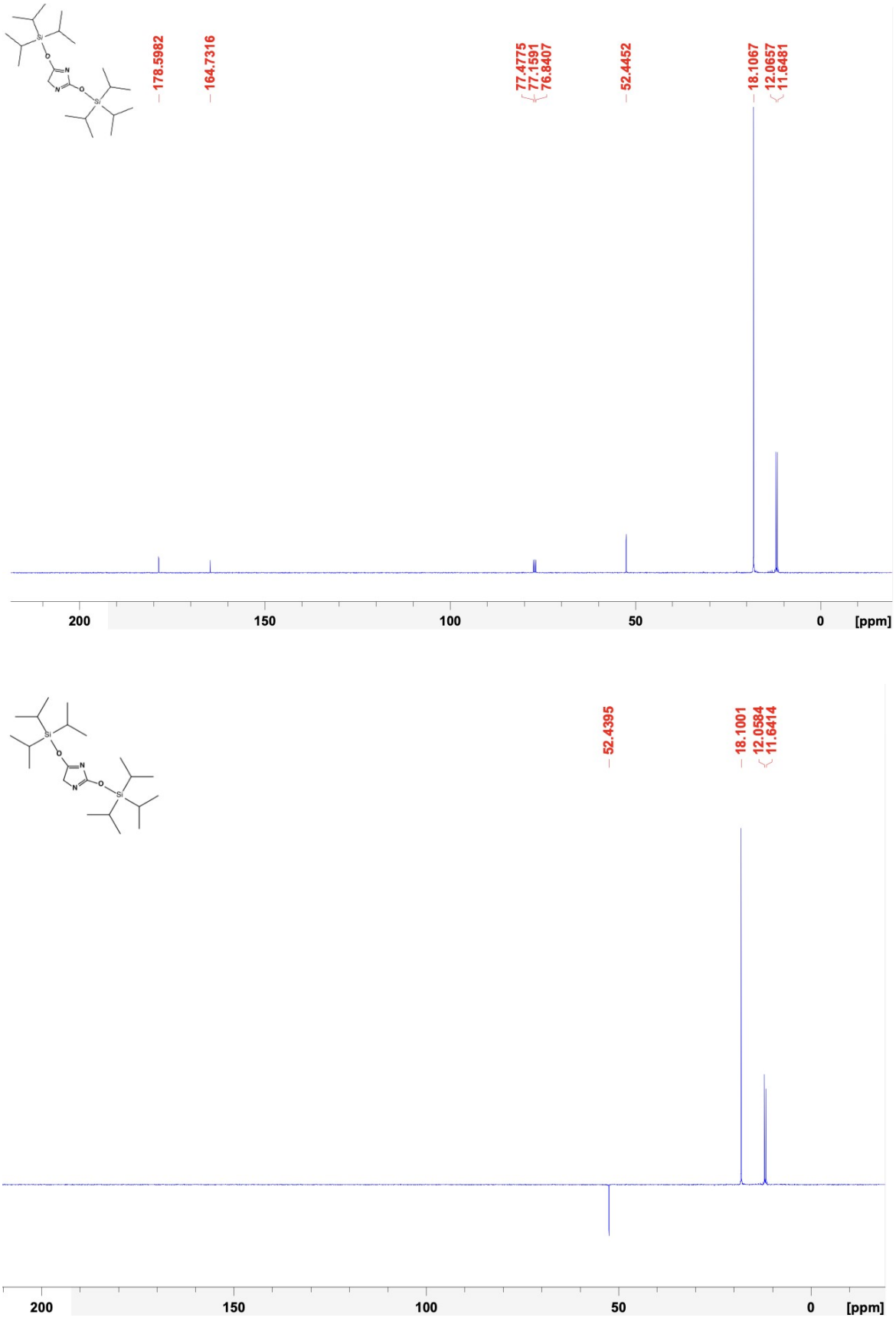
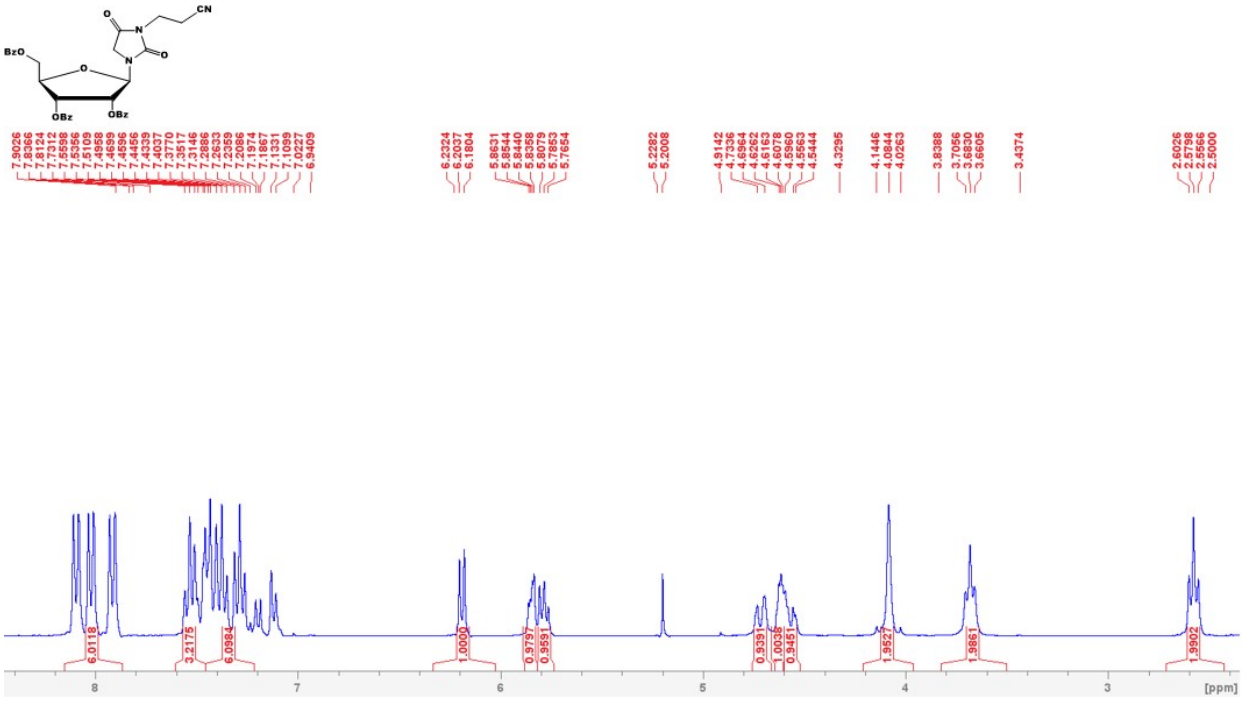
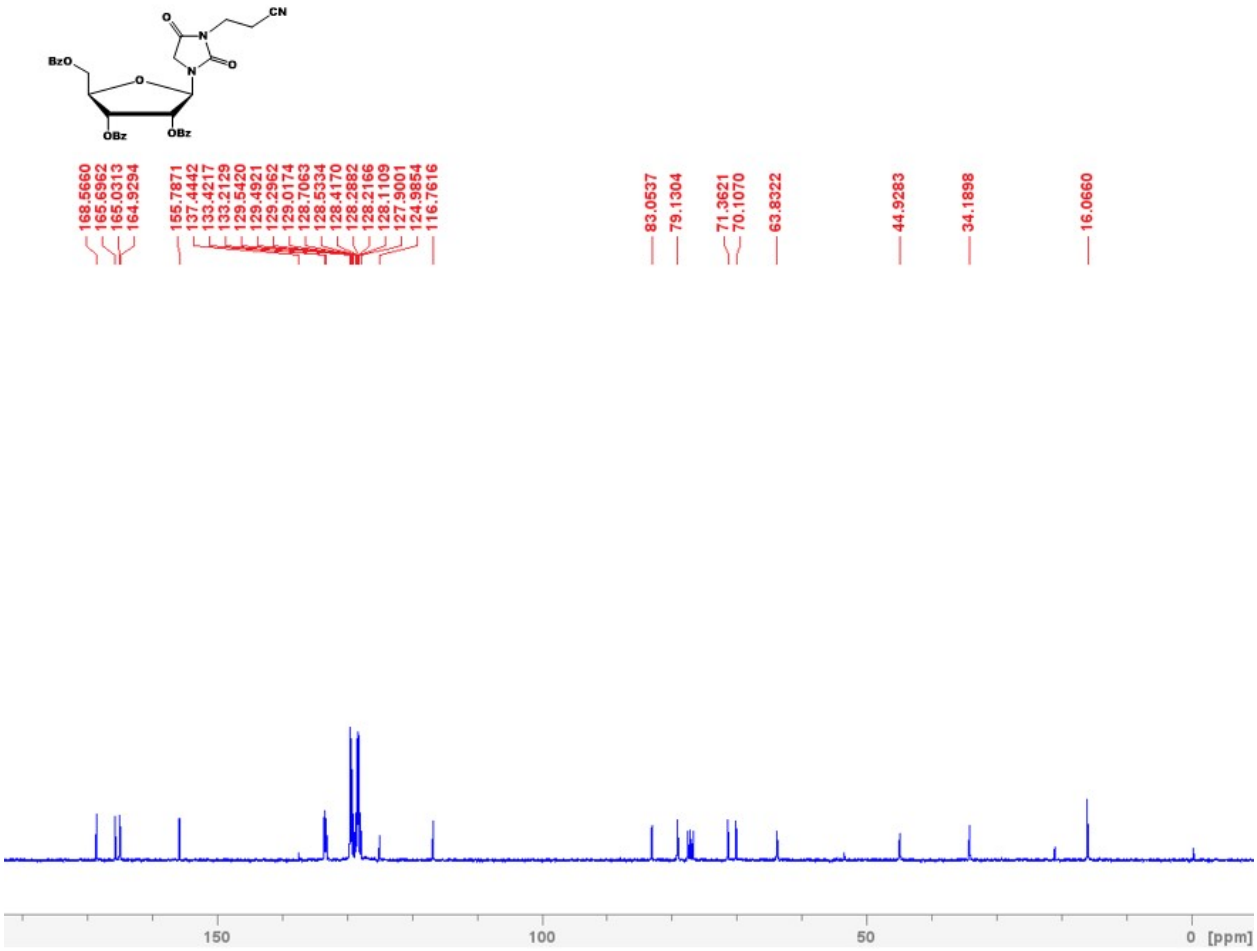


Fig. S11. 400 MHz ^{1}H-NMR/^{13}C-NMR/ DEPT spectra in $CDCl_3$, 300 K, of 2,5-bis((Triisopropylsilyl)oxy) -4*H*-imidazole (18).


BzO
OBz
OBz
CN
O
N
O
N
O
7.9026
7.8366
7.8134
7.7312
7.5598
7.5356
7.5109
7.4898
7.4688
7.4596
7.4456
7.4339
7.4037
7.3770
7.3617
7.3146
7.2886
7.2633
7.2359
7.2086
7.1874
7.1667
7.1331
7.1099
7.0227
6.9409
6.2324
6.2037
6.1804
5.8631
5.8544
5.8440
5.8348
5.8079
5.7853
5.7654
5.2282
5.2008
4.9142
4.7336
4.6964
4.6852
4.6163
4.6078
4.5960
4.5563
4.5444
4.3295
4.1446
4.0844
4.0263
3.8388
3.7056
3.6830
3.6605
3.4374
2.6026
2.5788
2.5566
2.5000
6.0118
3.2175
6.0984
1.0000
0.9797
0.9551
0.9391
1.0036
0.9451
1.9527
1.9861
1.9902
8
7
6
5
4
3
[ppm]


BzO
OBz
OBz
CN
O
N
O
N
O
168.5660
165.6962
165.0313
164.9294
155.7871
137.4442
133.4217
133.2129
129.5420
129.4921
129.2962
129.0174
128.7063
128.5334
128.4170
128.2882
128.2166
128.1109
127.9001
124.3854
116.7616
83.0537
79.1304
71.3621
70.1070
63.8322
44.9283
34.1898
16.0660
150
100
50
0
[ppm]

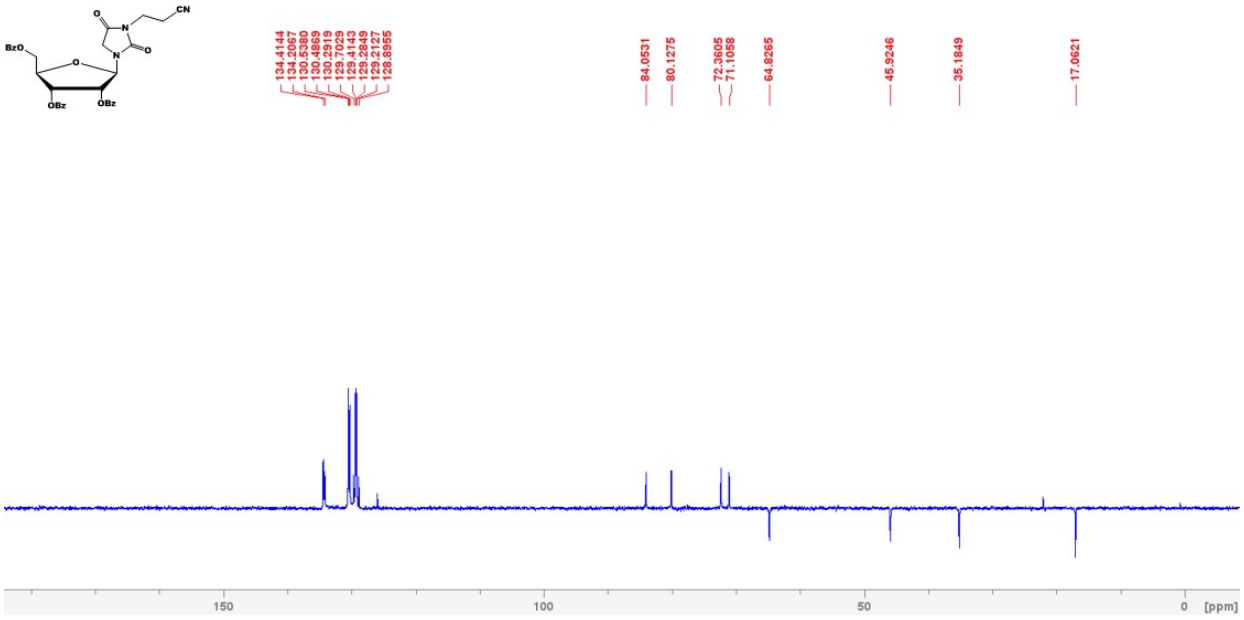


Fig. S12. 300 MHz ^{1}H-NMR/^{13}C-NMR/ DEPT spectra in $CDCl_3$, 300 K, of (2R,3R,4R,5R)-2-((Benzoyloxy)methyl)-5-(3-(2-cyanoethyl)-2,4dioxoimidazolidin-1-yl)tetrahydrofuran-3,4-diyl dibenzoate (14).

Tables

Table S1. Crystal data and structure refinement for HR.

Identification code	HR
Empirical formula	$C_8H_{12}N_2O_6$
Formula weight	232.20
Temperature/K	?
Crystal system	orthorhombic
Space group	$P2_12_12_1$
a/Å	9.08590(10)
b/Å	9.16110(10)
c/Å	11.4717(2)
α/°	90
β/°	90
γ/°	90
Volume/Å^3	954.87(2)
Z	4
ρ_{calc}g/cm^3	1.615
μ/mm^{-1}	1.214
F(000)	488.0
Crystal size/mm^3	0.35 × 0.2 × 0.1
Radiation	? (λ = 1.54184)
2Θ range for data collection/°	12.364 to 149.344
Index ranges	-9 ≤ h ≤ 10, -11 ≤ k ≤ 11, -14 ≤ l ≤ 14
Reflections collected	10131
Independent reflections	1910 [R_{int} = 0.0181, R_{sigma} = 0.0115]
Data/restraints/parameters	1910/0/148
Goodness-of-fit on F^2	1.061
Final R indexes [I>=2σ (I)]	R_1 = 0.0243, wR_2 = 0.0629
Final R indexes [all data]	R_1 = 0.0245, wR_2 = 0.0631
Largest diff. peak/hole / e Å^{-3}	0.17/-0.21

Flack parameter	-0.06(5)

Table S2. Fractional Atomic Coordinates (×10^4) and Equivalent Isotropic Displacement Parameters (Å^2×10^3) for HR. U_{eq} is defined as 1/3 of the trace of the orthogonalised U_{IJ} tensor.

Atom	*x*	*y*	*z*	U(eq)
O1	7434.1(13)	5586.7(12)	5367.3(10)	15.5(3)
O2	6221.4(15)	2733.6(13)	4288.8(11)	22.4(3)
O3	6148.8(15)	4158.2(14)	7771.4(10)	22.1(3)
O4	3994.9(14)	5958.3(13)	6872.8(11)	19.5(3)
O5	6967.0(13)	9537.1(13)	6101.1(11)	16.5(3)
O6	4824.0(14)	9296.9(13)	2507.7(11)	19.2(3)
N1	5905.2(16)	7598.2(15)	5105.1(12)	14.5(3)
N2	5886.8(16)	9780.4(15)	4277.6(12)	14.7(3)
C1	7372.5(18)	4129.3(17)	5865.0(14)	15.0(3)
C2	7565(2)	3032.3(18)	4891.8(16)	19.8(4)
C3	5916.0(19)	4018.9(18)	6549.3(14)	15.9(3)
C4	5067.3(18)	5357.9(18)	6118.2(14)	14.7(3)
C5	6303.6(18)	6444.9(18)	5886.1(15)	14.5(3)
C6	6322.7(17)	9015.9(18)	5261.2(14)	13.5(3)
C7	5292.3(18)	8897.4(18)	3453.4(15)	14.6(3)
C8	5321.5(19)	7362.4(18)	3934.3(15)	15.5(3)

Table S3. Anisotropic Displacement Parameters (Å^2×10^3) for HR. The Anisotropic displacement factor exponent takes the form: $-2\pi^2[h^2a^{*2}U_{11}+2hka^*b^*U_{12}+\ldots]$.

Atom	U_{11}	U_{22}	U_{33}	U_{23}	U_{13}	U_{12}
O1	15.9(6)	10.5(5)	20.1(6)	2.2(5)	3.5(5)	0.8(5)
O2	32.9(7)	13.7(6)	20.6(6)	1.3(5)	-7.8(6)	-3.0(5)
O3	26.6(7)	24.2(6)	15.5(6)	5.8(5)	-0.1(5)	1.5(6)
O4	18.9(6)	19.5(6)	20.1(6)	4.9(5)	5.0(5)	3.0(5)
O5	17.2(6)	15.2(6)	17.1(6)	-2.7(5)	-1.0(5)	-1.6(5)

O6	24.0(6)	17.4(6)	16.1(5)	1.3(5)	-2.9(5)	1.1(5)
N1	18.8(7)	10.8(6)	13.8(6)	-0.3(5)	-1.7(6)	-1.5(6)
N2	16.8(7)	10.7(6)	16.7(7)	0.7(5)	-1.4(6)	-0.9(5)
C1	16.4(8)	10.6(7)	18.1(8)	2.2(6)	-2.3(7)	0.3(6)
C2	23.0(9)	13.9(7)	22.7(9)	-0.9(7)	-0.6(7)	2.6(7)
C3	18.0(8)	13.1(8)	16.6(8)	1.3(6)	-1.2(7)	-1.2(7)
C4	15.5(8)	12.7(7)	15.8(7)	1.3(6)	-0.2(6)	-0.3(6)
C5	16.3(8)	11.5(7)	15.7(8)	-0.1(6)	0.2(6)	0.8(6)
C6	12.0(7)	12.5(7)	16.1(8)	-0.9(6)	2.7(6)	0.5(6)
C7	13.5(7)	14.2(8)	16.2(8)	0.0(6)	1.6(6)	0.9(6)
C8	18.3(8)	12.7(7)	15.4(7)	-0.3(6)	-2.3(7)	-0.1(7)

Table S4. Bond Lengths for HR.

Atom	Atom	Length/Å		Atom	Atom	Length/Å
O1	C1	1.4531(18)		N1	C8	1.460(2)
O1	C5	1.424(2)		N2	C6	1.386(2)
O2	C2	1.430(2)		N2	C7	1.357(2)
O3	C3	1.424(2)		C1	C2	1.512(2)
O4	C4	1.415(2)		C1	C3	1.542(2)
O5	C6	1.224(2)		C3	C4	1.531(2)
O6	C7	1.221(2)		C4	C5	1.525(2)
N1	C5	1.432(2)		C7	C8	1.511(2)
N1	C6	1.365(2)				

Table S5. Bond Angles for HR.

Atom	Atom	Atom	Angle/˚		Atom	Atom	Atom	Angle/˚
C5	O1	C1	108.38(12)		O4	C4	C5	111.12(13)
C5	N1	C8	123.94(14)		C5	C4	C3	102.04(13)
C6	N1	C5	123.36(14)		O1	C5	N1	109.16(13)
C6	N1	C8	111.24(14)		O1	C5	C4	104.11(13)
C7	N2	C6	112.33(14)		N1	C5	C4	113.89(14)
O1	C1	C2	108.43(13)		O5	C6	N1	127.39(15)
O1	C1	C3	107.06(12)		O5	C6	N2	125.47(15)
C2	C1	C3	115.56(15)		N1	C6	N2	107.14(14)
O2	C2	C1	112.68(15)		O6	C7	N2	125.39(16)
O3	C3	C1	111.59(14)		O6	C7	C8	127.54(15)
O3	C3	C4	108.73(13)		N2	C7	C8	107.07(14)
C4	C3	C1	102.43(13)		N1	C8	C7	101.81(13)
O4	C4	C3	117.44(13)					

Table S6. Torsion Angles for HR.

A	B	C	D	Angle/˚		A	B	C	D	Angle/˚
O1	C1	C2	O2	-83.23(17)		C3	C4	C5	N1	159.25(13)
O1	C1	C3	O3	-102.40(15)		C5	O1	C1	C2	137.08(14)
O1	C1	C3	C4	13.76(16)		C5	O1	C1	C3	11.77(16)
O3	C3	C4	O4	-35.6(2)		C5	N1	C6	O5	-6.5(3)
O3	C3	C4	C5	86.09(15)		C5	N1	C6	N2	173.24(14)
O4	C4	C5	O1	166.44(13)		C5	N1	C8	C7	-172.72(15)
O4	C4	C5	N1	-74.77(17)		C6	N1	C5	O1	-104.63(17)
O6	C7	C8	N1	-176.35(17)		C6	N1	C5	C4	139.51(16)
N2	C7	C8	N1	3.40(17)		C6	N1	C8	C7	-6.13(18)
C1	O1	C5	N1	-154.73(13)		C6	N2	C7	O6	-179.94(16)
C1	O1	C5	C4	-32.76(16)		C6	N2	C7	C8	0.30(18)

C1	C3	C4	O4	-153.83(14)		C7	N2	C6	O5	175.48(16)
C1	C3	C4	C5	-32.11(15)		C7	N2	C6	N1	-4.23(19)
C2	C1	C3	O3	136.72(15)		C8	N1	C5	O1	60.4(2)
C2	C1	C3	C4	-107.12(15)		C8	N1	C5	C4	-55.5(2)
C3	C1	C2	O2	36.9(2)		C8	N1	C6	O5	-173.13(16)
C3	C4	C5	O1	40.47(15)		C8	N1	C6	N2	6.56(19)

Table S7. Hydrogen Atom Coordinates (Å×104) and Isotropic Displacement Parameters (Å2×103) for HR.

Atom	*x*	*y*	*z*	U(eq)
H2	6121.42	3327.34	3736.43	34
H3	6183.93	3324.89	8075.6	33
H4	3920.36	5434.65	7470.94	29
H2A	5984.54	10730.74	4196.89	18
H1	8208.28	4016.78	6425.03	18
H2B	8298.55	3410.23	4329.46	24
H2C	7954.56	2110.94	5220.79	24
H3A	5380.31	3095.34	6359.62	19
H4A	4585.52	5109.59	5359.9	18
H5	6660.87	6864.35	6639.52	17
H8A	4322.46	6932.28	3959.49	19
H8B	5974.94	6722.6	3469.66	19

H’	Split (Hz)	Chemical shift (ppm)		Coupling constants (Hz)		
	HR	**U**	**HR**		**U**	**HR**
1'	d 6.6	5.90	5.61	**1'2'**	4.4	6.6
2'	t 5.4	4.34	4.23	**2'3'**	5.3	5.4
3'	dd 5.4, 3.5	4.22	4.14	**3'4'**	5.5	3.5
4'	q 3.6, 4.9	4.13	4.01	**4'5'**	3.0	4.9
5'	dd 12.5, 3.58	3.91	3.67	**4'5''**	4.4	3.6
5''	dd 12.5, 3.58	3.80	3.75	**5'5''**		12.5
5	d 17.35		4.14,4.21			

Table S8. Chemical shifts and coupling constants of HR, 15, and uridine, U, in D_2O solution.[1]

Table S9. Chemical shifts (ppm) for monomeric (δ_0) and self-stacked (δ_∞) HR and uridine.

	H5			H6			H1’		
compound	**δ_0**	**δ_∞**	**Δδ**	**δ_0**	**δ_∞**	**Δδ**	**δ_0**	**δ_∞**	**Δδ**
U	5.90	5.85	0.05	7.87	7.84	0.03	5.91	5.88	0.03
HR	4.22	4.22	0.00	-	-	-	5.61	5.60	0.01

Equations

Eqn. S1. Equations used for the calculation of *S/N* conformers' population.[2]

1. $J_{1'2'} = 9.3(1 - X_N) = 9.3X_S$
2. $J_{3'4'} = 9.3X_N$

Eqn. S2. Equations used for the calculation of populations of conformers *gg, tg,* and *gt*. (3)-(5).[2]

3. $\rho_{gg}=(J_t+J_g)-(J_{4'5'}+J_{4'5''})/(J_t-J_g)$
4. $\rho_{tg}= (J_{4'5''}-J_g)/(J_t-J_g)$
5. $\rho_{gt}=(J_{4'5''}-J_g)/(J_t-J_g)$

The coupling constants for pure rotamers were estimated as J_g 2.04 Hz (*gauche* rotamer) and J_t 11.72 (*trans* rotamer) for the C4'-C5' bond.[3]

Eqn. S3. Equation used for the calculation of the glycosidic angle of HR.[4]

$$6. \quad {}^3J_{C2/4\text{-}H1'} = 4.7\cos^2(\chi - 60°) + 2.3cos(\chi - 60°) + 0.1$$

Experimental

General

Commercial reagents and solvents were used without further purification unless otherwise noted. All moisture-sensitive reactions were carried out in flame-dried reaction flasks with rubber septa, and the reagents were introduced with a syringe. All reactants in moisture-sensitive reactions were dried overnight in a vacuum oven. Progress of reactions was monitored by TLC on precoated Merck silica gel plates (60F-254). Visualization was accomplished by UV light. Compounds were characterized by nuclear magnetic resonance using Bruker, DPX-300 and DMX-500 spectrometers. ^{1}H NMR spectra were measured at 300, 400, 500 and 700 MHz. Chemical shifts are expressed in ppm, downfield from Me_4Si (TMS), used as internal standard. Liquid chromatography (LC) was performed using CombiFlash NextGen 300+ (Nebraska, USA) with silica gel columns. Purification of HR was achieved on a CombiFlash RF+ Teledyne ISCO system (Nebraska, USA) using a C18, reverse phase column. High-resolution mass spectra were recorded on an AutoSpec-ESI mass spectrometer.

3-(2,5-Dioxoimidazolidin-1-yl)propanenitrile (12)

Hydantoin (15.0 g, 150 mmol) was dissolved in deionized water (150 mL) under nitrogen atmosphere. Diisopropylethylamine (DIPEA; 2.61 mL, 15 mmol) was then added, followed by acrylonitrile (8.19 mL, 125 mmol). The reaction mixture was stirred at 80°C under a nitrogen atmosphere for 16 h. Progress was monitored by TLC using chloroform/ethanol (9:1) eluent system and $KMnO_4$ staining. Upon completion of the reaction, the solvent was evaporated under reduced pressure, affording a sticky white crude product. This residue was purified by silica gel column chromatography, employing a gradient eluent system of chloroform and ethanol (10:0 v/v to 8.5:1.5 v/v). Pure 3-(2,5-dioxoimidazolidin-1-yl) propanenitrile, **12**, was isolated as a white powder in 85% yield (10.16 g; >95% purity according to ^{1}H NMR). ^{1}H NMR (400 MHz, DMSO)**:** δ 8.13 (s, NH1), 3.95 (s, 1H, H5), 3.62 (t, J = 6.5 Hz, 2H, N3(**CH2**CH2CN)), 2.8 (t, J = 6.5, 2H, CH2**CH2**CN) ppm. ^{13}C NMR (400 MHz, DMSO): δ 171.85 (C4), 156.94 (C2), 118.4 (CN), 45.97 (C5), 33.6 (N3(**CH2**CH2CN)), 16.17 (CH2**CH2**CN) ppm. HRMS (m/z) negative: calcd. For $C_6H_7N_3O_2$ 153.0538; found 153.0539.

(2R,3R,4R,5R)-2-((Benzoyloxy)methyl)-5-(3-(2-cyanoethyl)-2,4dioxoimidazolidin-1-yl)tetrahydrofuran-3,4-diyl dibenzoate (14)

3-(2,5-dioxoimidazolidin-1-yl)propanenitrile, **12**, (1.0 g, 6.53 mmol) was dissolved in dry MeCN (25 mL). HMDS (2.74 mL, 13.06 mmol) and TMSCl (0.20 mL) were added, and the mixture was heated under reflux for 1 h under N_2 atmosphere. Then, MeCN and HMDS were evaporated under reduced pressure. The residue was dissolved in

dry MeCN (30 mL). β-D-Ribofuranose 1-acetate 2,3,5-tribenzoate, **7,** (2.63 g, 5.22 mmol) was added, followed by TMSOTf (0.95 mL, 5.22 mmol) and stirred under N_2 atmosphere overnight at RT. The reaction mixture was diluted with DCM (60 mL) and washed with saturated $NaHCO_3$ solution (90 mL). The organic phase was treated with brine and dried over $MgSO_4$. TLC (toluene:ethyl acetate 2 : 1) showed complete consumption of the β-D-ribofuranose, **7**. After evaporation of the solvent under reduced pressure, a white-brownish residue was obtained. The crude product was purified by silica column chromatography using toluene: ethyl acetate (8: 2) eluent mixture to give pure product **14** as a white powder in 70% yield (2.18 g) and > 95% purity by ^{1}H NMR. ^{1}H NMR (300 MHz, $CDCl_3$)**:** δ 8.13-7.91(3d, 6H, O-Bz), 7.6-7.45(m, 3H, P-Bz), 7.4-7.2(m, 6H, m-Bz), 6.2 (d, J =7 Hz, 1H, H1'), 5.85 (dd, J =5.61, 2.51 Hz, 1H, H3'), 5.78 (t, J =6.2 Hz, 1H, H2'), 4.72 (dd, J =11.2, 2.03 Hz, 1H, H5'), 4.61 (q, J = 2.6 Hz, 1H, H4'), 4.57 (dd, J =11.73, 3.74 Hz, 1H, H5''), 4.08 (t, J =17.5 Hz, 2H, H5), 3.68 (t, J = 6.75 Hz, 2H, N3(**CH2**CH2CN)), 2.57 (t, J = 6.9, 2H, CH2**CH2**CN) ppm. ^{13}C NMR (300 MHz, $CDCl_3$): δ 168.56 (C4), 165.7, 165.03, 165 (CO), 155.78 (C2), 133.42-128 (C-Ar), 116.76 (CN), 83.05 (C1'), 79.13 (C4'), 71.36 (C3'), 70.1 (C2'), 63.83 (C5'), 44.92 (C5), 34.18 (N3(**CH2**CH2CN)), 16.06 (CH2**CH2**CN) ppm. HRMS (m/z) positive: calcd. For $C_{32}H_{27}N_3O_9$ 579.1747; found 579.1763.

(2R,3R,4R,5R)-2-((Benzoyloxy)methyl)-5-(2,4-dioxoimidazolidin-1 yl) tetrahydrofuran-3,4-diyl dibenzoate (9)

A mixture of tert-butanol (10 mL) and THF (10 mL) was stirred under an argon atmosphere at 40°C in the presence of 3Å molecular sieves. Concurrently, (2R,3R,4R,5R)-2-((benzoyloxy)methyl)-5-(3-(2-cyanoethyl)-2,4-dioxoimidazolidin-1-yl)tetrahydrofuran-3,4-diyl dibenzoate, **14**, (0.500 g, 0.837 mmol) was dissolved in anhydrous THF (5 mL). This solution was then added dropwise to the preheated reaction mixture, followed by (0.63 mL, 1.05 mmol) 20% w/v solution of potassium tert-butoxide in THF. The reaction was stirred for 1 h at 40°C under an argon atmosphere, and reaction progress was monitored by TLC using toluene:ethyl acetate (2:1) as the eluent system and UV visualization. After completion, the reaction mixture was acidified with a few drops of formic acid and the solid was filtered. The filtrate was concentrated under reduced pressure, and the resulting crude product was purified by silica gel column chromatography employing a gradient eluent of toluene:ethyl acetate (5:0 v/v to 4:1 v/v). Pure product **9** was isolated as a white powder in 31% yield (0.32 g) and >95% purity by ^{1}H NMR. ^{1}H NMR (700 MHz, $CDCl_3$)**:** δ 8.12-7.91(3d, 6H, O-Bz), 7.62-7.55(m, 3H, P-Bz), 7.53-7.35(m, 6H, m-Bz), 6.12 (d, J =7.81 Hz, 1H, H1'), 5.81 (dd, J =5.84, 2.5 Hz, 1H, H3'), 5.68 (dd, J =5.88, 7.88 Hz, 1H, H2'), 4.77 (dd, J =12.3, 3.01 Hz, 1H, H5'), 4.61 (q, J =2.86 Hz, 1H, H4'), 4.57 (dd, J =12.3, 3.5 Hz, 1H, H5''), 4.07 (d, J =16.8 Hz, 1H, H5), 4.01 (d, J =16.8 Hz, 1H, H5) ppm. ^{13}C NMR (700 MHz, $CDCl_3$): δ 168.66 (C4), 166.1, 165.46, 165,31 (CO), 155.74 (C2), 133.96-128.32 (C-Ar), 82.84 (C1'), 79.74 (C4'), 71.76 (C3'), 69.95 (C2'), 64.1 (C5'), 46.12 (C5) ppm. HRMS (m/z) positive: calcd. For $C_{29}H_{24}N_2O_9$ 544.1466; found 544.1540.

1-((2R,3R,4S,5R)-3,4-Dihydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)imidazolidine-2,4-dione (15)

(2R,3R,4R,5R)-2-((Benzoyloxy)methyl)-5-(2,4-dioxoimidazolidin-1-yl)tetrahydrofuran-3,4-diyl dibenzoate, **9**, (0.45 g, 0.83 mmol) was suspended in 7 N methanolic ammonia solution (60 mL) at room temperature. After 72 h, the reaction was complete as indicated by TLC using H_2O/n-butanol/acetic acid (5:4:1) as the eluent system. The solvent was evaporated under reduced pressure, and the resulting crude product was dissolved in a biphasic

mixture of water and diethyl ether (45 mL of each). The layers were separated, and the aqueous phase was washed with diethyl ether (45 mL x 2). Following evaporation of the aqueous phase under reduced pressure, the crude product was freeze dried. The dry crude product was then purified by medium pressure LC Reverse Phase C18 employing a gradient of water:acetonitrile (9.9:0.1 v/v to 9.8:0.2 v/v). Pure product **15** was isolated as a white powder in 90% yield (0.174 g), and >95% purity by ^{1}H NMR. ^{1}H NMR (700 MHz, D_2O)**:** δ 5.61 (d, J = 6.6 Hz, 1H), 4.23 (t, J = 5.4 Hz, 1H), 4.21 (d, J = 17.35 Hz, 1H), 4.14 (m, J = 17.35, 5.4, 3.5 Hz, 2H), 4.01 (q, J = 4.9, 3.6 Hz, 1H), 3.75 (dd, J = 12.5, 3.58 Hz, 1H), 3.67 (dd, J = 12.5, 3.58 Hz, 1H) ppm. ^{13}C NMR (700 MHz, D_2O): δ 174.7 (C4), 159.2 (C2), 84.9 (C1'), 84.2 (C4'), 71.6 (C2'), 70.7 (C3'), 62 (C5'), 47.3 (C5) ppm. HRMS (m/z) negative: calcd. For $C_8H_{12}N_2O_6$ 232.0694; found 232.0620.

2,5-Bis((tert-butyldimethylsilyl)oxy)-4H-imidazole (16)

Hydantoin (1.00 g, 10 mmol, 1 eq) and t-butyldimethylsilyl chloride (TBDMS-Cl) (3.61 g, 24 mmol, 2.4 eq) were dissolved in anhydrous acetonitrile (50 mL) under an argon atmosphere. Triethylamine (TEA; 7.0 mL, 50 mmol, 5 eq) was then injected, and the reaction mixture was stirred at room temperature for 3 days. Reaction progress was monitored by TLC using DCM: Hexane (7:3) as the eluent system. After completion, the reaction mixture was concentrated under reduced pressure. The resulting crude product was dissolved in DCM (90 mL) and washed with dilute HCl solution (pH 3) (30 mLx3). The organic layer was subsequently treated with brine and dried over $MgSO_4$. The oily yellowish crude product was purified by silica gel column chromatography employing DCM:Hexane (70:30 v/v). Pure bis-silylated hydantoin, **16**, was isolated as a white powder in 75% yield (2.47 g). and >95% purity by ^{1}H NMR. ^{1}H NMR (400 MHz, $CDCl_3$)**:** δ 3.85 (s, 2H), 0.96 (s, 18H), 0.44 (s, 6H), 0.3 (s, 6H) ppm. ^{13}C NMR (400 MHz, $CDCl_3$)**:** δ 178.06(C4), 163.80(C2), 52.22 (C5), 26.4(SiC**$\underline{(CH_3)_3}$**), 19.1(Si**$\underline{C}$**$(CH_3)_3$), -4.9(**$\underline{(CH_3)_2}$**$SiC(CH_3)_3$) ppm. HRMS (m/z) positive: calcd. for $C_{15}H_{32}N_2O_2Si_2$ 328.2005; found 328.2080.

(2R,3R,4R,5R)-2-((Benzoyloxy)methyl)-5-(2,4-dioxoimidazolidin-1yl) tetrahydrofuran-3,4-diyl dibenzoate (9)

2,5-Bis((tert-butyldimethylsilyl)oxy)-4H-imidazole, **16**, (1.0 g, 3.04 mmol) and β-D-ribofuranose 1-acetate 2,3,5-tribenzoate, **7**, (1.23 g, 2.43 mmol) was dissolved in dry MeCN (30 mL). TMSOTf (0.49 mL, 2.68 mmol) was added, and the mixture was stirred under Ar atmosphere at $0^{\circ}C$ to $25^{\circ}C$ for overnight. The reaction mixture was diluted with DCM (120 mL) and washed with saturated $NaHCO_3$ solution (180 mL). The organic phase was treated with brine and dried over $MgSO_4$. TLC (toluene: ethyl acetate 2: 1) showed complete consumption of the β-D-ribofuranose, **7**. The solvent was evaporated under reduced pressure to obtain a brownish residue. The crude product was purified by silica gel column chromatography employing toluene: ethyl acetate (50:50 v/v) eluent. Pure product **9** was isolated as a white powder in 34% yield (0.29 g) and >95% purity by ^{1}H NMR. ^{1}H NMR (700 MHz, $CDCl_3$)**:** δ 8.12-7.91(3d, J = 8.37 Hz, 6H, O-Bz), 7.62-7.55(m, J = 8.0 Hz, 3H, P-Bz), 7.53-7.35(m, J = 7.81 Hz, 6H, m-Bz), 6.12 (d, J = 7.88 Hz, 1H, H1'), 5.81 (dd, J =5.85, 2.47 Hz, 1H, H3'), 5.68 (dd, J = 8.12, 5.88 Hz, 1H, H2'), 4.77 (dd, J = 12.11, 2.92 Hz, 1H, H5'), 4.61 (q, J = 3.12, 2.89 Hz, 1H, H4'), 4.57 (dd, J = 12.11, 3.70 Hz, 1H, H5''), 4.07/4.01 (dd, J = 38.20, 16.76 Hz, 2H) ppm. ^{13}C NMR (700 MHz, $CDCl_3$): δ 168.66 (C4), 166.1, 165.46, 165,31 (CO),

155.74 (C2), 133.96-128.32 (C-Ar), 82.84 (C1'), 79.74 (C4'), 71.76 (C3'), 69.95 (C2'), 64.1 (C5'), 46.12 (C5) ppm. HRMS (m/z) positive: calcd. For $C_{29}H_{24}N_2O_9$ 544.1466; found 544.154.

2,5-Bis((*tert*-butyldiphenylsilyl)oxy)-4*H*-imidazole (17)

Hydantoin (1.00 g, 10 mmol, 1 eq) and t-butyldiphenylsilyl chloride (TBDPS-Cl) (6.25 mL, 24 mmol, 2.4 eq) were dissolved in anhydrous acetonitrile (37 mL) under an argon atmosphere. Triethylamine (TEA; 7.0 mL, 50 mmol, 5 eq) was then injected, and the reaction mixture was stirred at room temperature for 3 days. Reaction progress was monitored by TLC using DCM: Hexane (7:3) as the eluent system. After completion, the reaction mixture was concentrated under reduced pressure. The resulting crude product was dissolved in DCM (90 mL) and washed with dilute HCl solution (pH 3) (30 mLx3). The organic layer was subsequently treated with brine and dried over $MgSO_4$. The oily yellowish crude product was purified by silica gel column chromatography employing DCM: Hexane (70:30 v/v) eluent. Pure N-silylated hydantoin, **17**, was isolated as a white powder in 70% yield (4.05 g). and >95% purity by ^{1}H NMR. ^{1}H NMR (300 MHz, $CDCl_3$)**:** δ 7.7-7.3(H-Ar, 20H), 3.67 (s, 2H), 1.2 (s, 18H) ppm. ^{13}C NMR (300 MHz, $CDCl_3$)**:** δ 177.3(C4), 163.65(C2), 136-128(C-Ar), 54.6(C5), 28.4(SiC**(CH$_3$)$_3$**), 19.76(Si**C**$(CH_3)_3$) ppm. HRMS (m/z) positive: Calcd. for $C_{35}H_{40}N_2O_2Si_2$ 576.2628; found 576.2645.

2,5-Bis((*Triisopropyl silyl*)oxy)-4*H*-imidazole (18)

Hydantoin (0.5 g, 5 mmol, 1 eq) and triisopropylsilyl chloride (TIPS-Cl) (2.56 mL, 12 mmol, 2.4 eq) were dissolved in anhydrous acetonitrile (19 mL) under an argon atmosphere. Triethylamine (TEA, 3.5 mL, 25 mmol, 5 eq) was then injected, and the reaction mixture was stirred at room temperature for 3 days. Reaction progress was monitored by TLC using DCM: Hexane (7:3) as the eluent system. After completion, the reaction mixture was concentrated under reduced pressure. The resulting crude product was dissolved in DCM (60 mL) and washed with dilute HCl solution (pH 3) (20 mLx3). The organic layer was subsequently treated with brine and dried over $MgSO_4$. The oily yellowish crude product was purified by silica gel column chromatography employing an isocratic eluent of DCM: Hexane (70:30 v/v). Pure bis-silylated hydantoin, **18**, was isolated as a white powder in 40% yield (2 g) and >95% purity as determined by ^{1}H NMR. ^{1}H NMR (400 MHz, $CDCl_3$)**:** δ 3.91 (s, 2H), 1.71(C2-O-Si-(C**H**$(CH_3)_2)_3$) (heptet, J = 7.62 Hz, 3H), 1.4(C4-O-Si-(C**H**$(CH_3)_2)_3$) (heptet, J = 7.62 Hz, 3H), 1.08(C2/4-O-Si-(CH**(CH$_3$)$_2$**$)_3$) (dd, J = 7.63 , 2.12 Hz, 36H) ppm. ^{13}C NMR (400 MHz, $CDCl_3$)**:** δ 178.60(C4), 164.73(C2), 52.44(C5), 18.10(Si(**C**$H(CH_3)_2)_3$), 12.06(C2-O-Si-(CH**(CH$_3$)$_2$**$)_3$), 11.64(C4-O-Si-(CH**(CH$_3$)$_2$**$)_3$) ppm. HRMS (m/z) Positive: calcd. For $C_{21}H_{44}N_2O_2Si_2$ 412.2941; found 412.2945.

(2*R*,3*R*,4*R*,5*R*)-2-((benzoyloxy)methyl)-5-(2,5-dioxoimidazolidin-1-yl)tetrahydrofuran-3,4-diyl dibenzoate (8)

Solutions of chloroacetyl chloride (0.1 mL, 1.27 mmol) in DCM, and 0.5 M NaOH (8.2 mL, 4 equiv) in distilled water were added dropwise (during half an hour) from two dropping funnels to compound **10** (570 mg, 1.02 mmol) in DCM (10 mL) in a three necked round bottomed flask. Two phases (aqueous and organic solutions) were formed. The aqueous layer was discarded, and the organic layer was washed with water, dried and concentrated. Pure product **8** was isolated as a white powder in 87% yield (0.181 g) and >95% purity by ^{1}H NMR.

^{1}H NMR (500 MHz, CDCl3): δ 8.10-7.91(3d, J = 7.85 Hz, 6H, O-Bz), 7.52(m, J = 7.37 Hz, 3H, P-Bz), 7.41-7.33(m, J = 7.5 Hz, 6H, m-Bz), 6.30 (dd, J = 6.1, 3.5 Hz, 1H, H-3'), 6.07 (dd, J = 8.12, 5.88 Hz, 1H, H2'), 5.94 (d, J = 3.58 Hz, 1H, H1'), 4.74 (dd, J = 11.47, 3.45 Hz, 1H, H5'), 4.66 (q, J = 5.76, 4.24 Hz, 1H, H4'), 4.62 (dd, J = 11.47, 5.50 Hz, 1H, H5''), 3.98 (s, 2H) ppm. ^{13}C NMR (500 MHz, $CDCl_3$): δ 170.24 (C4), 166.29, 165.49, 165,21 (CO), 156.31 (C2), 133.59-128.34 (C-Ar), 83.63 (C1'), 79.30 (C4'), 72.51 (C3'), 71.34 (C2'), 63.76 (C5'), 46.20 (C5) ppm. HRMS (m/z) positive: Calcd. for $C_{29}H_{24}N_2O_9$ 544.1466; found 544.154.

pK_a measurements

Dilute HCl and NaOH solutions were added to 0.1 M HR in D_2O to reach the following pH values: 7.5, 8, 8.4, 9.15, 9.5, 9.8, 10.15, 11, 11.5, 12.8, 13, 13.5. Apparent pH values were measured with a Hanna instruments pH meter equipped with an electrode. pH is estimated from the pH meter measurement, (apparent reading from pH meter) - 0.41.^{13}C-NMR spectra were measured in D_2O at 500 MHz. The data were collected at 300 K. ^{13}C-NMR chemical shift of the bases' carbonyls were monitored as a function of pH. The chemical shifts of C2 and C4 were plotted vs. pH and a sigmoid curve was obtained. The pK_a value was obtained from the inflection point, which was determined by the second derivative of the fitted sigmoid function using Python.

Monitoring base pairing of HR with A/G

HR (74.26 mg), adenosine (53.42 mg), and guanosine (57.84 mg) in volumetric flasks were stored under vacuum overnight to remove absorbed water. 0.1 M HR and 0.1 M adenosine solutions in dry DMSO-d_6 were prepared. ^{1}H-NMR spectra were measured at 500 MHz. The data were collected at 300 K. ^{1}H-NMR spectra were obtained across a range of ratios between 0.1 M HR and 0.1 M adenosine. The final volume of the solution in the tube was 1 mL. This protocol was repeated for guanosine.

Monitoring base stacking of HR

HR (111.4 mg) in a volumetric flask was dried under vacuum overnight. 0.003, 0.025, 0.04, 0.05, 0.25, and 0.4 M solutions were prepared. $NaNO_3$ was added to increase the ionic strength to 0.1 M. ^{1}H-NMR spectra of these solutions were measured in D_2O at 500 MHz. The data were collected at 300 K.

References:

1. T. Ginsburg-Shmuel, M. Haas, M. Schumann, G. Reiser, O. Kalid, N. Stern and B. Fischer, *J Med Chem*, 2010, 53(4), 1673-1685.

2. H. Salameh, M. Afri, H. E. Gottlieb and B. Fischer, *ACS Omega*, 2020, 5(48), 31314–31322.

3. Hruska, F. E.; Grey, A. A.; Smith, I. C., Nuclear magnetic resonance study of the molecular conformation of. beta.-pseudouridine in aqueous solution. *J. Am. Chem. Soc.* **1970,** *92* (13), 4088-4094.

4. Donohue, J.; Trueblood, K. N., Base pairing in DNA. *J. Mol. biol* **1960,** *2* (6), 363-371.

5. Stern, N.; Major, D. T.; Gottlieb, H. E.; Weizman, D.; Fischer, B., What is the conformation of physiologically-active dinucleoside polyphosphates in solution? Conformational analysis of free dinucleoside polyphosphates by NMR and molecular dynamics simulations. Org. biolmol. Chem. **2010,** *8* (20), 4637-4652.

6. Ippel, J. H.; Wijmenga, S. S.; de Jong, R.; Heus, H. A.; Hilbers, C. W. Heteronuclear Scalar Coupling in the Bases and Sugar Rings of Nucleic Acids: Their Determination and Application in Assignment and Conformational Analysis. *Magn. Reson. Chem*. 1996, *34*, S156-S176.