

## Electronic Supplementary Information

### **A general method for *N*-glycosylation of nucleobases promoted by (*p*-Tol)<sub>2</sub>SO/Tf<sub>2</sub>O with thioglycoside as donor**

Guang-jian Liu, Xiao-tai Zhang and Guo-wen Xing\*

Department of Chemistry, Beijing Normal University, Beijing 100875, China.

Email: [gwxing@bnu.edu.cn](mailto:gwxing@bnu.edu.cn)

#### *Contents*

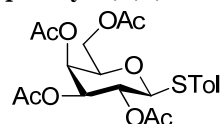
<b>1. General information</b>	<b>S1</b>
<b>2. Characterization data for glycosyl donors</b>	<b>S1–S3</b>
<b>3. Preparation of acceptors</b>	<b>S4</b>
<b>4. General procedure for <i>N</i>-glycosylations of nucleobases</b>	<b>S5</b>
<b>5. NMR data for coupling products</b>	<b>S6–S17</b>
<b>6. References</b>	<b>S18</b>
<b>7. NMR spectra for compounds</b>	<b>S19–S53</b>

## 1. General information

All the starting materials were purchased as reagent grade and used without further purification. Reactions were monitored with analytical thin-layer chromatography (TLC) on silica gel F<sub>254</sub> glass plates and visualized under UV light (254nm) or by staining with acidic ceric ammonium molybdate or EtOH-H<sub>2</sub>SO<sub>4</sub> (3%). All the *N*-glycosylation reactions were performed in flame-dried glass vessel under an inert argon atmosphere. Purification of products was performed by flash column chromatography on silica gel (200-300 mesh). <sup>1</sup>H NMR(400MHz) and <sup>13</sup>C NMR(100MHz) spectra were recorded on a Bruker Avanced III spectrometer. Chemical shifts for <sup>1</sup>H were reported in δ-values (ppm) with tetramethylsilane as an internal standard. The δ values for <sup>13</sup>C were calibrated with deuterated solvents (CDCl<sub>3</sub> δ = 77.16 ppm). Mass spectra were recorded using electrospray ionization (ESI).

## 2. Characterization data for glycosyl donors

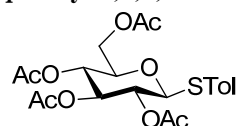
### *p*-Tolyl 2,3,4,6-tetra-*O*-acetyl-1-thio-β-D-galactopyranoside (1a)



<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.41 (d, *J* = 8.1 Hz, 2H, ArH), 7.13 (d, *J* = 7.9 Hz, 2H, ArH), 5.41 (d, *J* = 2.5 Hz, 1H, H-4), 5.22 (t, *J* = 10.0 Hz, 1H, H-2), 5.04 (dd, *J* = 9.9, 3.3 Hz, 1H, H-3), 4.65 (d, *J* = 10.0 Hz, 1H, H-1), 4.19 (dd, *J* = 11.3, 6.9 Hz, 1H, H-6), 4.11 (dd, *J* = 11.3, 6.3 Hz, 1H, H-6), 3.91 (t, *J* = 6.6 Hz, 1H, H-5), 2.35 (s, 3H, STol CH<sub>3</sub>), 2.12 (s, 3H, Ac), 2.10 (s, 3H, Ac), 2.04 (s, 3H, Ac), 1.97 (s, 3H, Ac).

The spectroscopic data coincide with the previous report<sup>[1]</sup>

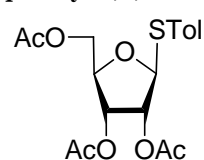
### *p*-Tolyl 2,3,4,6-tetra-*O*-acetyl-1-thio-β-D-glucopyranoside (1b)



<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.39 (d, *J* = 8.1 Hz, 2H, ArH), 7.12 (d, *J* = 7.9 Hz, 2H, ArH), 5.21 (t, *J* = 9.4 Hz, 1H, H-3), 5.02 (t, *J* = 9.8 Hz, 1H, H-4), 4.93 (t, *J* = 9.7 Hz, 1H, H-2), 4.63 (d, *J* = 10.1 Hz, 1H, H-1), 4.20 (m, 2H, H-6), 3.70 (ddd, *J* = 10.1, 4.8, 2.7 Hz, 1H, H-5), 2.35 (s, 3H, STol CH<sub>3</sub>), 2.09 (s, 3H, Ac), 2.08 (s, 3H, Ac), 2.01 (s, 3H, Ac), 1.98 (s, 3H, Ac).

The spectroscopic data coincide with the previous report<sup>[2]</sup>

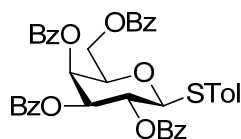
### *p*-Tolyl 2,3,5-tri-*O*-acetyl-1-thio-β-D-ribofuranoside (1c)



<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.41 (d, *J* = 8.1 Hz, 2H, ArH), 7.14 (d, *J* = 7.9 Hz, 2H, ArH), 5.27 – 5.17 (m, 3H), 4.27 – 4.21 (m, 2H), 4.10 – 4.04 (m, 1H), 2.34 (s, 3H), 2.10 (s, 3H), 2.07 (s, 3H), 2.04 (s, 3H).

The spectroscopic data coincide with the previous report<sup>[3]</sup>

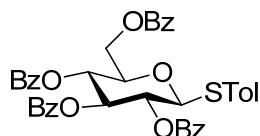
***p*-Tolyl 2,3,4,6-tetra-*O*-benzoyl-1-thio- $\beta$ -D-galactopyranoside (1d)**



$^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.04 – 8.01 (m, 2H), 8.00 – 7.97 (m, 2H), 7.93 – 7.89 (m, 2H), 7.77 – 7.73 (m, 2H), 7.64 – 7.50 (m, 3H), 7.49–7.38 (m, 9H), 7.23 (t,  $J = 7.8$  Hz, 2H), 7.07 (d,  $J = 7.9$  Hz, 2H), 5.99 (d,  $J = 3.2$  Hz, 1H), 5.74 (t,  $J = 9.9$  Hz, 1H), 5.59 (dd,  $J = 9.9, 3.3$  Hz, 1H), 4.98 (d,  $J = 9.9$  Hz, 1H), 4.65 (dd,  $J = 11.3, 6.8$  Hz, 1H), 4.43 (dd,  $J = 11.3, 5.8$  Hz, 1H), 4.37 (t,  $J = 6.4$  Hz, 1H), 2.37 (s, 3H).  $^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  166.18, 165.65, 165.52, 165.30, 138.78, 134.54, 133.67, 133.45, 133.39, 130.16, 129.98, 129.96, 129.90, 129.75, 129.62, 129.50, 129.08, 128.89, 128.66, 128.57, 128.40, 127.54, 86.27, 75.22, 73.17, 68.51, 68.07, 62.63, 21.47.

A similar experimental procedure was applied as described in the literature. [*Carbohydr. Res.* 384 (2014) 1–8]

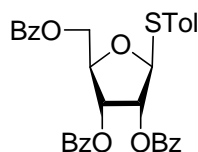
***p*-Tolyl 2,3,4,6-tetra-*O*-benzoyl-1-thio- $\beta$ -D-glucopyranoside (1e)**



$^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  8.04 (d,  $J = 7.3$  Hz, 2H, ArH), 7.98 (d,  $J = 7.4$  Hz, 2H, ArH), 7.89 (d,  $J = 7.4$  Hz, 2H, ArH), 7.79 (d,  $J = 7.4$  Hz, 2H, ArH), 7.60 (t,  $J = 7.4$  Hz, 1H, ArH), 7.57 – 7.31 (m, 11H, ArH), 7.30 – 7.23 (m, 2H, ArH), 6.94 (d,  $J = 7.9$  Hz, 2H, ArH), 5.89 (t,  $J = 9.5$  Hz, 1H, H-3), 5.59 (t,  $J = 9.8$  Hz, 1H, H-4), 5.45 (t,  $J = 9.7$  Hz, 1H, H-2), 4.98 (d,  $J = 10.0$  Hz, 1H, H-1), 4.68 (dd,  $J = 12.1, 2.5$  Hz, 1H, H-6'), 4.47 (dd,  $J = 12.1, 5.8$  Hz, 1H, H-6), 4.21 – 4.13 (m, 1H, H-5), 2.28 (s, 3H,  $\text{PhCH}_3$ ).

The spectroscopic data coincide with the previous report<sup>[4]</sup>

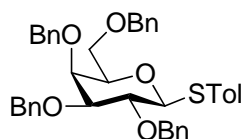
***p*-Tolyl 2, 3, 5-tri-*O*-benzoyl-1-thio- $\beta$ -D-ribofuranoside (1f)**



$^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.04 (d,  $J = 8.0$  Hz, 2H), 7.99 (d,  $J = 8.0$  Hz, 2H), 7.90 (d,  $J = 7.9$  Hz, 2H), 7.59 – 7.49 (m, 3H), 7.48 – 7.30 (m, 8H), 7.05 (d,  $J = 7.8$  Hz, 2H), 5.71 (t,  $J = 4.7$  Hz, 1H), 5.63 (t,  $J = 5.2$  Hz, 1H), 5.54 (d,  $J = 5.2$  Hz, 1H), 4.67 – 4.59 (m, 2H), 4.48 (dd,  $J = 12.4, 4.5$  Hz, 1H), 2.22 (s, 3H).

The spectroscopic data coincide with the previous report<sup>[3]</sup>

***p*-Tolyl 2,3,4,6-tetra-*O*-benzyl-1-thio- $\beta$ -D-galactopyranoside (1g)**

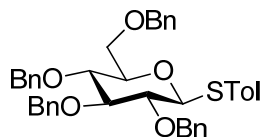


$^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.46 (d,  $J = 8.1$  Hz, 2H, SPh), 7.40 – 7.23 (m, 20H, Bn), 6.98 (d,  $J = 8.0$

Hz, 2H, SPh), 4.95 (d,  $J = 11.5$  Hz, 1H, H-1), 4.79 (d,  $J = 10.2$  Hz, 1H, Bn), 4.76 – 4.67 (m, 3H, Bn), 4.59 (dd,  $J = 10.6, 4.1$  Hz, 2H, Bn), 4.43 (q,  $J = 11.7$  Hz, 2H, Bn), 3.97 (d,  $J = 2.5$  Hz, 1H, H-4), 3.89 (t,  $J = 9.4$  Hz, 1H, H-2), 3.65 (d,  $J = 6.2$  Hz, 2H), 3.59 (dd,  $J = 9.2, 2.7$  Hz, 2H), 2.28 (s, 3H, Me).

The spectroscopic data coincide with the previous report<sup>[5]</sup>

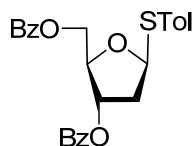
#### ***p*-Tolyl 2,3,4,6-tetra-*O*-benzyl-1-thio- $\beta$ -D-glucopyranoside (1h)**



<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  7.48 (d,  $J = 8.1$  Hz, 2H, ArH), 7.42 – 7.37 (m, 2H, ArH), 7.36 – 7.23 (m, 16H, ArH), 7.21 – 7.17 (m, 2H, ArH), 7.02 (d,  $J = 8.0$  Hz, 2H, ArH), 4.89 (dd,  $J = 10.6, 3.6$  Hz, 2H, Bn), 4.83 (dd,  $J = 10.8, 8.6$  Hz, 2H, Bn), 4.72 (d,  $J = 10.3$  Hz, 1H, Bn), 4.63 – 4.56 (m, 3H, Bn, H-1), 4.53 (d,  $J = 11.9$  Hz, 1H, Bn), 3.78 (dd,  $J = 10.9, 1.9$  Hz, 1H, H-6), 3.75 – 3.66 (m, 2H, H-4, H-6), 3.63 (t,  $J = 9.3$  Hz, 1H, H-3), 3.48 (ddd,  $J = 9.5, 4.4, 2.0$  Hz, 2H, H-2, H-5), 2.30 (s, 3H, PhCH<sub>3</sub>).

The spectroscopic data coincide with the previous report<sup>[4]</sup>

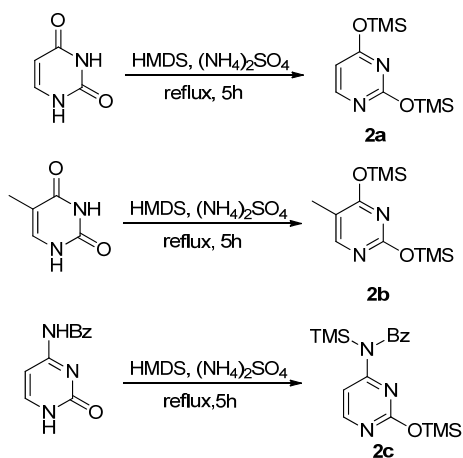
#### ***p*-Tolyl 3,5-di-*O*-benzoyl-1-thio-2-deoxy- $\beta$ -D-ribofuranoside (1i)**



<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  8.14 – 8.10 (m, 2H), 8.04 – 8.00 (m, 2H), 7.62 – 7.53 (m, 2H), 7.50 – 7.40 (m, 6H), 7.11 (d,  $J = 7.9$  Hz, 2H), 5.79 (dd,  $J = 7.7, 2.5$  Hz, 1H), 5.50 (ddd,  $J = 7.5, 3.8, 2.5$  Hz, 1H), 4.78 (q,  $J = 4.0$  Hz, 1H), 4.62 (dd,  $J = 4.0, 2.2$  Hz, 2H), 2.94 (dt,  $J = 15.0, 7.6$  Hz, 1H), 2.36 (dt,  $J = 14.7, 2.5$  Hz, 1H), 2.32 (s, 3H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>)  $\delta$  166.35, 166.28, 137.63, 133.51, 133.28, 132.18, 131.63, 130.02, 129.92, 129.86, 129.78, 129.70, 128.63, 128.56, 88.40, 81.26, 75.00, 64.28, 39.75, 21.24. ESI-MS for C<sub>26</sub>H<sub>24</sub>O<sub>5</sub>SNa [M+Na]<sup>+</sup>: 471.1, found: 471.3.

### 3. Preparation of acceptors

#### (1) General procedure for silylation of pyrimidines



A suspension of nucleobase (5 mmol), (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> (20 mg) in hexamethyldisilazane (HMDS) (7 mmol, 1.5 mL) was refluxed until the reaction mixture became clear. The reaction mixture was then concentrated in vacuo and the crude was kept under high vacuum overnight to remove traces of solvent.

#### (2) Preparation of Boc-protected purines.

The Boc-protected purines **2e** and **2f** were prepared according to the published procedures. [*J. Org. Chem.*, 65 (2000) 7697–7699 ]

#### 4. General procedure for *N*-glycosylations of nucleobases

##### (1) General procedure for *N*-glycosylations of pyrimidines

To a mixture of thioglycoside donor (1.0 equiv, 0.069 mmol), (*p*-Tol)<sub>2</sub>SO (2 – 6 equiv), and activated 3 Å powdered sieves in flame-dried glass vessel was added anhydrous dichloromethane (CH<sub>2</sub>Cl<sub>2</sub>, 2.5 mL), which was freshly distilled over calcium hydride. The resulting mixture was stirred at preactivation temperature (-70°C – -40°C) for 20 min, followed by the addition of trifluoromethanesulfonic anhydride (13.6 μL, 1.2 equiv). After activation for 30 min (for donor **1e**, the activation time was 1 h), a solution of silylated nucleobase (3.0 equiv) in anhydrous CH<sub>3</sub>CN (0.7 ml) was added, and the reaction mixture was maintained at this temperature for 2 h and then raise 20°C higher for another 2 h. The reaction mixture was quenched with saturated NaHCO<sub>3</sub> solution (1 ml), diluted with CH<sub>2</sub>Cl<sub>2</sub>, filtered through Celite, washed with saturated brine, dried and concentrated to leave a residue which was purified by column chromatography on silica gel to afford the pure nucleosides **3-20** and **29-37**.

##### (2) General procedure for *N*-glycosylations of purines

To a mixture of thioglycoside donor (1.0 equiv, 0.069 mmol), (*p*-Tol)<sub>2</sub>SO (2.0 equiv), and activated 3 Å powdered sieves in flame-dried glass vessel was added anhydrous dichloromethane (CH<sub>2</sub>Cl<sub>2</sub>, 2.5 mL), which was freshly distilled over calcium hydride. The resulting mixture was stirred at preactivation temperature (-50°C – -40°C) for 20 min, followed by the addition of trifluoromethanesulfonic anhydride (13.6 μL, 1.2 equiv). After activation for 30 min (for donor **1e**, the activation time was 1 h), a solution of acceptor (3.0 equiv) in anhydrous CH<sub>3</sub>CN (0.7 ml) was added, and the reaction mixture was allowed warming to rt gradually. After which, the reaction mixture was quenched with Et<sub>3</sub>N (0.1 ml), diluted with CH<sub>2</sub>Cl<sub>2</sub>, filtered through Celite, washed with saturated brine, dried and concentrated to leave a residue which was purified by column chromatography on silica gel to afford the pure nucleosides **21 – 28**.

## 5. NMR data for coupling products

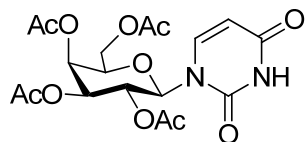
### (1) Reference for known compounds

Compounds	References	<sup>1</sup> H NMR	Compounds	References	<sup>1</sup> H NMR
3	[6]	★	22	[8]	★
4	[7]	★	23	[8]	★
5	[8]	★	25	[8]	★
6	[6]	★	26	[8]	★
7	[9]	★	29	[16]	★
8	[8]	★	30	[16]	★
9	[10]	★	31	[16]	★
10	[11]	★	32	[13]	★
11	[12]	★	33	[13]	★
15	[8]	★	34	[16]	★
16	[13]	★	35	[10]	★
18	[14]	★	36	[17]	★
19	[14]	★	37	[17]	★
20	[14]	★	S1	[8]	★
21	[15]	★			

### (2) New compounds

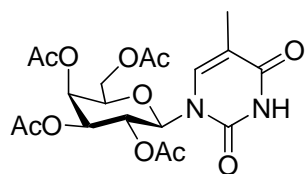
Compounds	<sup>1</sup> H NMR	<sup>13</sup> C NMR	HRMS
12	★	★	★
13	★	★	★
14	★	★	★
17	★	★	★
24	★	★	★
27	★	★	★
28	★	★	★
S2	★	★	★

### 1-(2',3',4',6'-Tetra-*O*-acetyl-β-D-galactopyranosyl)-uracil (3)



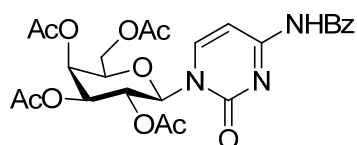
( $R_f = 0.40$ , petroleum: ethyl acetate = 1:3) <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 9.16 (s, broad, 1H, NH), 7.40 (d,  $J = 8.2$  Hz, 1H, H-6), 5.90 – 5.83 (m, 2H, H-1' and H-5), 5.51 (d,  $J = 2.7$  Hz, 1H), 5.31 (t,  $J = 9.6$  Hz, 1H), 5.24 (dd,  $J = 10.2, 3.2$  Hz, 1H), 4.10-4.18 (m, 3H), 2.20 (s, 3H, -OAc), 2.06 (s, 3H, -OAc), 2.02 (s, 3H, -OAc), 2.01 (s, 3H, -OAc).

#### 1-(2',3',4',6'-Tetra-*O*-acetyl- $\beta$ -D-galactopyranosyl)-thymine (4)



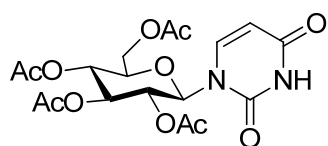
( $R_f$ =0.21, petroleum: ethyl acetate = 2:3)  $^1\text{H NMR}$  (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.86 (s, 1H, NH), 7.16 (d,  $J$  = 1.2 Hz, 1H, H-6), 5.85 (d,  $J$  = 9.2 Hz, 1H, H-1'), 5.51 (d,  $J$  = 3.2 Hz, 1H, H-4'), 5.32 (t,  $J$  = 9.7 Hz, 1H, H-2'), 5.22 (dd,  $J$  = 10.2, 3.3 Hz, 1H, H-3'), 4.22 – 4.05 (m, 3H, H-5' and H-6'), 2.22 (s, 3H, OAc), 2.06 (s, 3H, OAc), 2.01 (s, 3H, OAc), 1.98 (d,  $J$  = 1.0 Hz, 3H,  $\text{CH}_3$ ).

#### *N*<sup>4</sup>-benzoyl-1-(2',3',4',6'-tetra-*O*-acetyl- $\beta$ -D-galactopyranosyl)-cytosine (5)



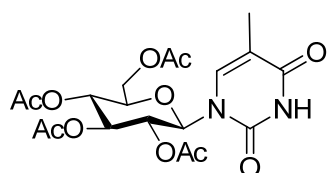
( $R_f$ =0.40, petroleum: ethyl acetate = 1:4)  $^1\text{H NMR}$  (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.91 (d,  $J$  = 7.5 Hz, 2H), 7.84 (d,  $J$  = 7.6 Hz, 1H), 7.58 (dt,  $J$  = 15.3, 7.4 Hz, 4H), 6.11 (d,  $J$  = 8.8 Hz, 1H, H-1'), 5.54 (d,  $J$  = 2.5 Hz, 1H, H-4'), 5.32 (dd,  $J$  = 12.7, 6.3 Hz, 1H, H-2'), 5.26 (dd,  $J$  = 10.2, 3.2 Hz, 1H, H-3'), 4.23 – 4.09 (m, 3H, H-5' and H-6'), 2.22 (s, 3H, OAc), 2.06 (s, 3H, OAc), 2.01 (s, 6H, OAc).

#### 1-(2',3',4',6'-Tetra-*O*-acetyl- $\beta$ -D-glucopyranosyl)-uracil (6)



( $R_f$  = 0.34, petroleum: ethyl acetate = 1:3)  $^1\text{H NMR}$  (400 MHz,  $\text{CDCl}_3$ )  $\delta$  9.35 (s, broad, 1H, NH), 7.35 (d,  $J$  = 8.2 Hz, 1H, H-6), 5.90 (d,  $J$  = 9.5 Hz, 1H, H-1'), 5.84 (dd,  $J$  = 8.2, 2.0 Hz, 1H, H-5), 5.41 (t,  $J$  = 9.5 Hz, 1H, H-2'), 5.17 (dt,  $J$  = 13.5, 9.6 Hz, 2H, H-3' and H-4'), 4.28 (dd,  $J$  = 12.6, 4.9 Hz, 1H, H-6'), 4.13 (dd,  $J$  = 12.6, 2.0 Hz, 1H, H-6'), 3.96 (ddd,  $J$  = 10.2, 4.9, 2.1 Hz, 1H, H-5'), 2.10 (s, 3H, Ac), 2.06 (s, 3H, Ac), 2.03 (s, 3H, Ac), 2.01 (s, 3H, Ac).

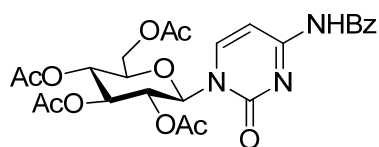
#### 1-(2',3',4',6'-Tetra-*O*-acetyl- $\beta$ -D-glucopyranoside)-thymine (7)



( $R_f$  = 0.49, petroleum: ethyl acetate = 1:3)  $^1\text{H NMR}$  (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.63 (s, 1H, NH), 7.14 (d,  $J$  = 1.2 Hz, 1H, H-6), 5.87 (d,  $J$  = 9.5 Hz, 1H, H-1'), 5.38 (t,  $J$  = 9.5 Hz, 1H, H-3'), 5.24 – 5.11 (m, 2H, H-2', H-4'), 4.28 (dd,  $J$  = 12.6, 5.1 Hz, 1H, H-6'), 4.12 (dd,  $J$  = 12.6, 2.0 Hz, 1H, H-6'), 3.93 (ddd,  $J$  = 10.2, 5.1, 2.1 Hz, 1H, H-5'), 2.10 (s, 3H, OAc), 2.06 (s, 3H, OAc), 2.02 (s, 3H, OAc), 2.00 (s, 3H, OAc), 1.96 (d,  $J$  = 1.1 Hz, 3H,  $\text{CH}_3$ ).

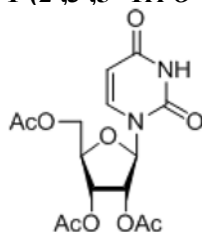


***N*<sup>4</sup>-benzoyl-1-(2',3',4',6'-tetra-*O*-acetyl-β-*D*-glucopyranosyl)-cytosine (8)**



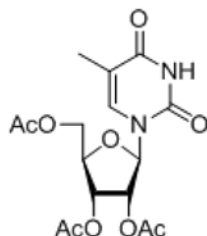
(*R*<sub>f</sub> = 0.30, petroleum: ethyl acetate = 1:3) <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 8.76 (s, 1H), 7.90 (d, *J* = 6.6 Hz, 2H), 7.81 (d, *J* = 7.0 Hz, 1H), 7.66 – 7.50 (m, 4H), 6.14 (d, *J* = 9.3 Hz, 1H), 5.44 (t, *J* = 9.4 Hz, 1H), 5.25 – 5.13 (m, 2H), 4.30 (dd, *J* = 12.6, 4.6 Hz, 1H), 4.14 (d, *J* = 12.5 Hz, 1H), 4.02-3.94 (m, 1H), 2.10 (s, 3H), 2.07 (s, 3H), 2.03 (s, 3H), 1.99 (s, 3H).

**1-(2',3',5'-Tri-*O*-acetyl-β-*D*-ribofuranosyl)-uracil (9)**



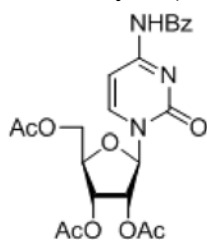
(*R*<sub>f</sub> = 0.28, petroleum: ethyl acetate = 1:3) <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 8.72 (s, 1H, NH), 7.39 (d, *J* = 8.2 Hz, 1H, H-6), 6.04 (d, *J* = 4.9 Hz, 1H, H-1'), 5.79 (dd, *J* = 8.2, 1.8 Hz, 1H, H-5), 5.37 – 5.31 (m, 2H, H-2' and H-3'), 4.40 – 4.32 (m, 3H, H-4' and H-5'), 2.15 (s, 3H, Ac), 2.13 (s, 3H, Ac), 2.11 (s, 3H, Ac).

**1-(2',3',5'-Tri-*O*-acetyl-β-*D*-ribofuranosyl)-thymine (10)**



(*R*<sub>f</sub> = 0.41, petroleum: ethyl acetate = 2:5) <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 8.96 (br s, 1H, N-H), 7.18 (d, *J* = 1.2 Hz, 1H, H-6), 6.12 – 6.06 (m, 1H, H-1'), 5.38 – 5.31 (m, 2H, H-2' and H-3'), 4.41 – 4.31 (m, 3H, H-4' and H-5'), 2.16 (s, 3H, -OAc), 2.13 (s, 3H, -OAc), 2.10 (s, 3H, -OAc), 1.95 (d, *J* = 1.1 Hz, 3H, -CH<sub>3</sub>).

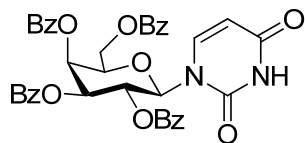
***N*<sup>4</sup>-benzoyl-1-(2',3',5'-tri-*O*-acetyl-β-*D*-ribofuranosyl)- cytosine (11)**



(*R*<sub>f</sub> = 0.22, petroleum: ethyl acetate = 2:7) <sup>1</sup>H NMR (400 MHz, DMSO) δ 11.35 (s, 1H, ex, NH), 8.19 (d, *J* = 7.5 Hz, 1H, H-6), 8.01 (d, *J* = 7.3 Hz, 2H, Ph), 7.63 (t, *J* = 7.4 Hz, 1H, Ph), 7.52 (t, *J* = 7.7 Hz, 2H, Ph), 7.40 (bs, 1H, H-5), 5.94 (d, *J* = 3.8 Hz, 1H, H-1'), 5.54 (dd, *J* = 6.1, 3.9 Hz, 1H, H-2'), 5.40

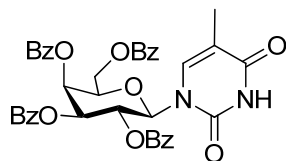
(apt,  $J = 6.3$  Hz, 1H, H-3'), 4.41 – 4.30 (m, 2H), 4.25 (dd,  $J = 11.6, 5.3$  Hz, 1H), 2.08 (s, 3H), 2.08 (s, 3H), 2.06 (s, 3H).

#### 1-(2',3',4',6'-Tetra-*O*-benzoyl- $\beta$ -D-galactopyranosyl)-uracil (12)



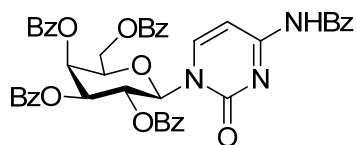
( $R_f = 0.20$ , petroleum: ethyl acetate = 1:1)  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.18 (s, 1H, NH), 8.08 – 8.03 (m, 2H), 8.00 (dd,  $J = 8.3, 1.3$  Hz, 2H), 7.91 – 7.87 (m, 2H), 7.77 (dd,  $J = 8.3, 1.2$  Hz, 2H), 7.70-7.65(m, 1H), 7.61 – 7.40 (m, 8H), 7.36 (t,  $J = 7.8$  Hz, 2H), 7.29 – 7.23 (m, 2H), 6.21 (d,  $J = 9.3$  Hz, 1H), 6.10 (d,  $J = 2.7$  Hz, 1H), 5.95 – 5.87 (m, 2H), 5.82 (dd,  $J = 10.1, 3.3$  Hz, 1H), 4.64 (dd,  $J = 10.9, 6.7$  Hz, 1H), 4.61 – 4.56 (m, 1H), 4.45 (dd,  $J = 10.9, 5.2$  Hz, 1H).  $^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  166.09, 165.62, 165.42, 165.35, 162.62, 150.39, 139.25, 133.98, 133.93, 133.58, 133.48, 130.03, 129.94, 129.90, 129.85, 129.28, 129.02, 128.97, 128.65, 128.57, 128.47, 128.14, 104.06, 81.04, 74.62, 71.79, 68.37, 68.20, 62.16; HRMS (ESI) calcd for  $\text{C}_{38}\text{H}_{30}\text{N}_2\text{O}_{11}\text{Na}$   $[\text{M}+\text{Na}]^+$ : 713.1742, found: 713.1741

#### 1-(2',3',4',6'-Tetra-*O*-benzoyl- $\beta$ -D-galactopyranosyl)-thymine (13)



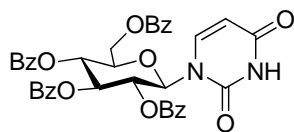
( $R_f = 0.51$ , petroleum: ethyl acetate = 1:1)  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.60 (s, 1H), 8.10 – 8.04 (m, 2H), 8.03 – 7.98 (m, 2H), 7.91 – 7.86 (m, 2H), 7.79 (dd,  $J = 8.3, 1.2$  Hz, 2H), 7.71 – 7.65 (m, 1H), 7.55 (ddd,  $J = 7.6, 4.8, 3.3$  Hz, 3H), 7.51 – 7.31 (m, 7H), 7.26 (t,  $J = 7.8$  Hz, 2H), 6.25 (d,  $J = 9.3$  Hz, 1H), 6.10 (d,  $J = 3.2$  Hz, 1H), 5.93 (t,  $J = 9.7$  Hz, 1H), 5.82 (dd,  $J = 10.1, 3.3$  Hz, 1H), 4.67 – 4.57 (m, 2H), 4.51 – 4.43 (m, 1H), 2.03 (d,  $J = 0.9$  Hz, 3H).  $^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  166.12, 165.61, 165.41, 165.37, 163.18, 150.43, 134.75, 134.02, 133.94, 133.61, 133.53, 130.05, 129.95, 129.91, 129.88, 129.30, 129.10, 128.98, 128.68, 128.60, 128.49, 128.22, 112.41, 81.00, 74.53, 71.91, 68.36, 68.23, 62.24, 12.97; HRMS (ESI) calcd for  $\text{C}_{39}\text{H}_{32}\text{N}_2\text{O}_{11}\text{Na}$   $[\text{M}+\text{Na}]^+$ : 727.1898, found: 727.1897.

#### $N^4$ -benzoyl-1-(2',3',4',5'-tetra-*O*-benzoyl- $\beta$ -D-galactopyranosyl)-cytosine (14)



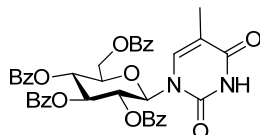
( $R_f = 0.40$ , petroleum: ethyl acetate = 2:3)  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.10 – 7.85 (m, 9H), 7.81 – 7.52 (m, 7H), 7.52 – 7.39 (m, 6H), 7.34 (t,  $J = 7.8$  Hz, 2H), 7.29 – 7.22 (m, 2H), 6.53 (d,  $J = 9.1$  Hz, 1H), 6.13 (d,  $J = 3.2$  Hz, 1H), 5.96 (t,  $J = 9.7$  Hz, 1H), 5.84 (dd,  $J = 10.1, 3.3$  Hz, 1H), 4.68 – 4.60 (m, 2H), 4.48 (q,  $J = 9.2$  Hz, 1H).  $^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  166.06, 165.73, 165.43, 165.32, 162.62, 144.39, 133.98, 133.81, 133.53, 133.45, 133.40, 130.11, 129.93, 129.86, 129.29, 129.12, 129.03, 128.98, 128.67, 128.60, 128.56, 128.44, 127.79, 81.86, 74.86, 71.94, 69.05, 68.44, 62.19; HRMS (ESI) calcd for  $\text{C}_{45}\text{H}_{35}\text{N}_3\text{O}_{11}\text{Na}$   $[\text{M}+\text{Na}]^+$ : 816.2164, found: 816.2165.

### 1-(2',3',4',6'-Tetra-*O*-benzoyl- $\beta$ -D-glucopyranosyl)-uracil (15)



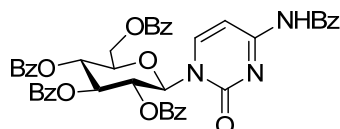
( $R_f$  = 0.30, petroleum: ethyl acetate = 1:1)  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.24 (s, 1H), 8.04 (dd,  $J$  = 8.3, 1.3 Hz, 2H), 7.93 (dd,  $J$  = 8.4, 1.2 Hz, 2H), 7.90 – 7.87 (m, 2H), 7.82 (dd,  $J$  = 8.4, 1.2 Hz, 2H), 7.62 – 7.56 (m, 1H), 7.56 – 7.33 (m, 10H), 7.32 – 7.26 (m, 2H), 6.25 (d,  $J$  = 9.5 Hz, 1H), 6.09 (t,  $J$  = 9.6 Hz, 1H), 5.83 (dd,  $J$  = 8.2, 2.2 Hz, 1H), 5.78 (t,  $J$  = 9.8 Hz, 1H), 5.66 (t,  $J$  = 9.5 Hz, 1H), 4.67 (dd,  $J$  = 12.4, 2.7 Hz, 1H), 4.50 (dd,  $J$  = 12.5, 5.0 Hz, 1H), 4.40 (ddd,  $J$  = 10.0, 5.0, 2.7 Hz, 1H).  $^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  166.15, 165.61, 165.40, 165.29, 161.99, 150.10, 139.13, 134.05, 133.86, 133.61, 133.48, 130.15, 130.05, 129.94, 129.87, 129.52, 128.71, 128.67, 128.63, 128.52, 128.00, 103.96, 80.89, 75.63, 73.00, 70.31, 68.99, 62.72.

### 1-(2',3',4',6'-Tetra-*O*-benzoyl- $\beta$ -D-glucopyranosyl)-thymine (16)



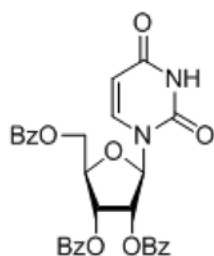
( $R_f$  = 0.41, petroleum: ethyl acetate = 1:1)  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.20 (s, 1H), 8.06 – 8.00 (m, 2H), 7.94 – 7.89 (m, 2H), 7.89 – 7.85 (m, 2H), 7.84 – 7.79 (m, 2H), 7.57 (dd,  $J$  = 10.6, 4.3 Hz, 1H), 7.51 (q,  $J$  = 7.4 Hz, 2H), 7.44 (dd,  $J$  = 10.9, 4.4 Hz, 3H), 7.36 (q,  $J$  = 8.0 Hz, 4H), 7.31 – 7.25 (m, 3H), 6.24 (d,  $J$  = 9.4 Hz, 1H), 6.07 (t,  $J$  = 9.6 Hz, 1H), 5.77 (t,  $J$  = 9.8 Hz, 1H), 5.67 (t,  $J$  = 9.5 Hz, 1H), 4.67 (dd,  $J$  = 12.4, 2.7 Hz, 1H), 4.48 (dd,  $J$  = 12.4, 5.1 Hz, 1H), 4.38 (ddd,  $J$  = 10.0, 5.0, 2.7 Hz, 1H), 1.94 (d,  $J$  = 0.9 Hz, 3H).

### $N^4$ -benzoyl-1-(2',3',4',6'-tetra-*O*-benzoyl- $\beta$ -D-glucopyranosyl)-cytosine (17)



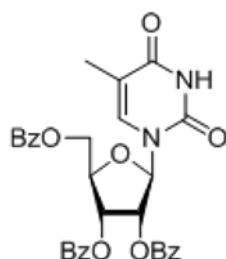
( $R_f$  = 0.42, petroleum: ethyl acetate = 2:3)  $^1\text{H}$  NMR (400 MHz, DMSO)  $\delta$  11.24 (s, 1H), 8.69 (d,  $J$  = 6.3 Hz, 1H), 8.00 (dd,  $J$  = 29.9, 7.5 Hz, 4H), 7.82 (dd,  $J$  = 22.7, 7.5 Hz, 4H), 7.76 – 7.36 (m, 18H), 6.67 (d,  $J$  = 9.1 Hz, 1H), 6.27 (t,  $J$  = 9.4 Hz, 1H), 6.09 (t,  $J$  = 9.3 Hz, 1H), 5.96 (t,  $J$  = 9.7 Hz, 1H), 4.87 (d,  $J$  = 9.7 Hz, 1H), 4.53 (s, 2H).  $^{13}\text{C}$  NMR (126 MHz, DMSO)  $\delta$  167.39, 165.42, 165.10, 164.80, 164.72, 163.54, 154.25, 146.35, 133.99, 133.84, 133.79, 133.47, 132.94, 132.82, 129.52, 129.27, 128.95, 128.80, 128.55, 128.43, 128.13, 97.24, 80.47, 73.34, 71.60, 68.43, 62.48; HRMS (ESI) calcd for  $\text{C}_{45}\text{H}_{35}\text{N}_3\text{O}_{11}\text{Na}$   $[\text{M}+\text{Na}]^+$ : 816.2164, found: 816.2164

**1-(2',3',5'-Tri-*O*-benzoyl- $\beta$ -D-ribofuranosyl)-uracil (18)**



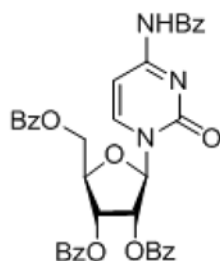
( $R_f$  = 0.52, petroleum: ethyl acetate = 2:3 )  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.47 (bs, 1H, NH), 8.10 (dd,  $J$  = 5.2, 3.3 Hz, 2H), 7.98 (dd,  $J$  = 8.3, 1.2 Hz, 2H), 7.94 (dd,  $J$  = 8.3, 1.2 Hz, 2H), 7.65 – 7.53 (m, 3H), 7.49 (t,  $J$  = 7.7 Hz, 2H), 7.44 – 7.34 (m, 5H), 6.32 (d,  $J$  = 5.6 Hz, 1H, H-1'), 5.89 (dd,  $J$  = 5.9, 4.5 Hz, 1H, H-3'), 5.75 (t,  $J$  = 5.8 Hz, 1H, H-2'), 5.61 (dd,  $J$  = 8.1, 2.1 Hz, 1H, H-5), 4.84 (dd,  $J$  = 12.0, 2.7 Hz, 1H, H-5'), 4.74 – 4.70 (m, 1H, H-4'), 4.67 (dd,  $J$  = 12.1, 3.8 Hz, 1H, H-5').

**1-(2',3',5'-Tri-*O*-benzoyl- $\beta$ -D-ribofuranosyl)-thymine (19)**



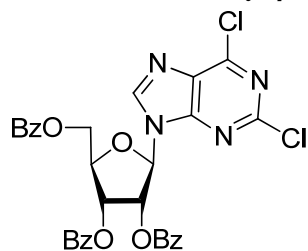
( $R_f$  = 0.58, petroleum: ethyl acetate = 4:5 )  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.63 (bs, 1H, NH), 8.17 – 8.11 (m, 2H), 7.99 (dd,  $J$  = 8.3, 1.2 Hz, 2H), 7.95 (dd,  $J$  = 8.3, 1.2 Hz, 2H), 7.66 – 7.47 (m, 5H), 7.42-7.35 (m, 4H), 7.16 (d,  $J$  = 1.2 Hz, 1H, H-6), 6.43 (d,  $J$  = 6.4 Hz, 1H, H-1'), 5.92 (dd,  $J$  = 6.0, 3.7 Hz, 1H, H-3'), 5.76 (t,  $J$  = 6.2 Hz, 1H, H-2'), 4.89 (dd,  $J$  = 12.1, 2.6 Hz, 1H, H-5'), 4.70 (dd,  $J$  = 6.3, 3.4 Hz, 1H), 4.65 (dd,  $J$  = 12.1, 3.5 Hz, 1H), 1.60 (d,  $J$  = 1.0 Hz, 3H).

**$N^4$ -benzoyl-1-(2',3',5'-tri-*O*-benzoyl- $\beta$ -D-ribofuranosyl)-cytosine (20)**



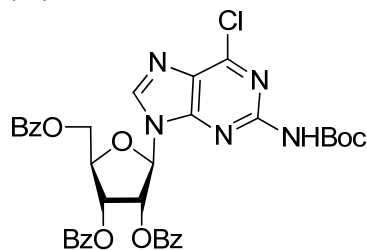
( $R_f$  = 0.28, petroleum: ethyl acetate = 2:3 )  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.14 – 8.09 (m, 2H), 8.06 – 7.89 (m, 7H), 7.65-7.45 (m, 9H), 7.41 - 7.35 (m, 4H), 6.47 (d,  $J$  = 4.6 Hz, 1H, H-1'), 5.92 (t,  $J$  = 5.5 Hz, 1H), 5.88 – 5.82 (m, 1H), 4.87 (dd,  $J$  = 12.2, 2.7 Hz, 1H), 4.80 (dd,  $J$  = 8.3, 3.3 Hz, 1H), 4.74 (dd,  $J$  = 12.2, 3.9 Hz, 1H).

**2',3',5'-Tri-*O*-benzoyl- $\beta$ -D-ribofuranosyl-2,6-dichloropurine (21)**



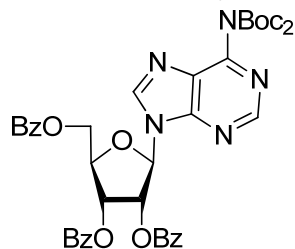
( $R_f$  = 0.39, petroleum: ethyl acetate = 2:1 )  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.30 (s, 1H, H-8), 8.08 – 8.00 (m, 4H, -OBz), 7.93 (d,  $J$  = 7.3 Hz, 2H, -OBz), 7.63 – 7.53 (m, 3H, -OBz), 7.45 (dd,  $J$  = 15.1, 7.5 Hz, 4H, -OBz), 7.37 (t,  $J$  = 7.8 Hz, 2H, -OBz), 6.49 (d,  $J$  = 5.4 Hz, 1H, H-1'), 6.19 (t,  $J$  = 5.6 Hz, 1H, H-3'), 6.16 – 6.12 (m, 1H, H-2'), 4.93 (dd,  $J$  = 12.1, 3.2 Hz, 1H, H-5'), 4.88 (dd,  $J$  = 7.5, 3.9 Hz, 1H, H-4'), 4.73 (dd,  $J$  = 12.1, 4.0 Hz, 1H, H-5').  $^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  166.16, 165.41, 165.25 (3 Bz), 153.57, 152.73, 152.42, (C-6, C-2, C-4), 143.98 (C-8), 134.12, 134.01, 133.76, 131.46, 130.02, 129.98, 129.73, 129.15, 128.85, 128.75, 128.72, 128.65, 128.19, (3Bz, C-5) 87.12 (C-1'), 81.61 (C-4'), 74.41 (C-3'), 71.66 (C-2'), 63.60 (C-5').

**2-*tert*-Butoxycarbonylamino-6-chloro-9-(2',3',5'-tri-*O*-benzoyl- $\beta$ -D-ribofuranosyl)purine (22)**



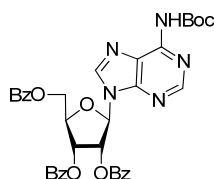
( $R_f$  = 0.44, petroleum: ethyl acetate = 2:1 )  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.08 (s, 1H), 8.04 – 7.98 (m, 4H), 7.96 – 7.91 (m, 2H), 7.62 – 7.52 (m, 4H), 7.44 – 7.34 (m, 6H), 6.61 (t,  $J$  = 5.9 Hz, 1H), 6.32 – 6.26 (m, 2H), 4.97 (dd,  $J$  = 11.9, 3.8 Hz, 1H), 4.92 – 4.87 (m, 1H), 4.79 (dd,  $J$  = 11.9, 5.4 Hz, 1H), 1.48 (s, 9H).

***N,N*-Di-*tert*-butoxycarbonyl-9-(2',3',5'-tri-*O*-benzoyl- $\beta$ -D-ribofuranosyl)adenine (23)**



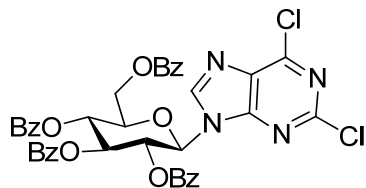
( $R_f$  = 0.52, petroleum: ethyl acetate = 1:1 )  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.76 (s, 1H), 8.25 (s, 1H), 8.11 (d,  $J$  = 7.3 Hz, 2H), 8.02 (d,  $J$  = 7.3 Hz, 2H), 7.92 (d,  $J$  = 7.3 Hz, 2H), 7.63 – 7.52 (m, 3H), 7.49 – 7.33 (m, 6H), 6.51 (d,  $J$  = 5.3 Hz, 1H), 6.41 (t,  $J$  = 5.5 Hz, 1H), 6.27 (t,  $J$  = 5.2 Hz, 1H), 4.92 (dd,  $J$  = 12.1, 3.2 Hz, 1H), 4.86 (dd,  $J$  = 7.9, 4.1 Hz, 1H), 4.73 (dd,  $J$  = 12.1, 4.2 Hz, 1H), 1.44 (s, 18H).

***N*-tert-Butoxycarbonyl-9-(2',3',5'-tri-*O*-benzoyl- $\beta$ -D-ribofuranosyl)adenine (S1)**



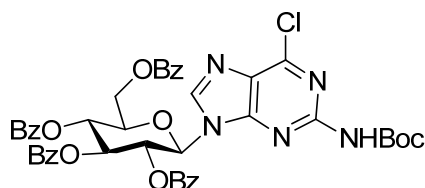
$^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.67 (s, 1H), 8.10 (d,  $J = 6.8$  Hz, 4H), 8.01 (d,  $J = 7.2$  Hz, 2H), 7.92 (d,  $J = 7.2$  Hz, 2H), 7.62 – 7.52 (m, 3H), 7.48 – 7.33 (m, 6H), 6.47 (d,  $J = 5.2$  Hz, 1H), 6.38 (t,  $J = 5.5$  Hz, 1H), 6.25 (t,  $J = 5.3$  Hz, 1H), 4.93 (dd,  $J = 12.2, 3.2$  Hz, 1H), 4.84 (dd,  $J = 7.9, 4.1$  Hz, 1H), 4.70 (dd,  $J = 12.2, 4.2$  Hz, 1H), 1.55 (s, 9H).

**2',3',4',6'-Tetra-*O*-benzoyl- $\beta$ -D- glucopyranosyl -2,6-dichloropurine (24)**



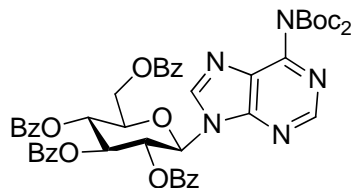
( $R_f = 0.32$ , petroleum: ethyl acetate = 2:1)  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.48 (s, 1H, H-8), 8.02 (d,  $J = 7.4$  Hz, 2H), 7.94 (d,  $J = 7.4$  Hz, 2H), 7.82 (d,  $J = 7.4$  Hz, 2H), 7.74 (d,  $J = 7.4$  Hz, 2H), 7.60 - 7.50 (m, 2H), 7.48 - 7.34 (m, 6H), 7.32 - 7.26 m, 4H), 6.28 (d,  $J = 9.3$  Hz, 1H, H-1'), 6.19 (t,  $J = 9.6$  Hz, 1H), 6.05 (t,  $J = 9.5$  Hz, 1H), 5.92 (t,  $J = 9.4$  Hz, 1H), 4.70 (dd,  $J = 14.5, 4.8$  Hz, 1H), 4.54 (d,  $J = 9.4$  Hz, 2H).  $^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  166.09, 165.62, 165.21, 164.96, (4 Bz), 153.67, 153.10, 152.37(C-2, C-4, C-6), 143.45(C-8), 134.13, 133.90, 133.69, 133.50, 130.72(C-5), 130.02, 129.97, 129.89, 129.86, 129.34, 128.67, 128.61, 128.53, 128.43, 128.39, 127.44, 81.39(C-1'), 76.01, 72.83, 71.21, 68.80, 62.53; HRMS (ESI) calcd for  $\text{C}_{39}\text{H}_{29}\text{N}_4\text{O}_9\text{Cl}_2$   $[\text{M}+\text{H}]^+$ :767.1312, found: 767.1307.

**2-*tert*-Butoxycarbonylamino-6-chloro-9-(2',3',4',6'-tetra-*O*-benzoyl- $\beta$ -D-glucopyranosyl)purine (25)**



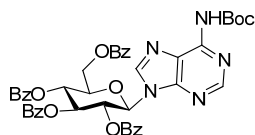
$^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.31 (s, 1H), 8.05 – 7.99 (m, 2H), 7.97 – 7.93 (m, 2H), 7.85 – 7.79 (m, 2H), 7.78 – 7.71 (m, 2H), 7.58 – 7.36 (m, 9H), 7.31 – 7.22 (m, 4H), 6.27 (d,  $J = 9.3$  Hz, 1H, H-1'), 6.18 (t,  $J = 9.6$  Hz, 1H), 6.02 (t,  $J = 9.5$  Hz, 1H), 5.89 (t,  $J = 9.5$  Hz, 1H), 4.68 (q,  $J = 4.4$  Hz, 1H), 4.57 – 4.48 (m, 2H), 1.58 (s, 9H, Boc).

***N,N*-Di-*tert*-butoxycarbonyl-9-(2',3',4',6'-tetra-*O*-benzoyl- $\beta$ -D-glucopyranosyl)adenine (26)**



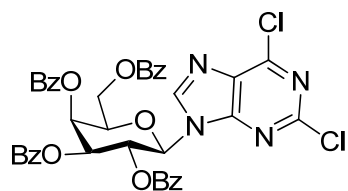
<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 8.79 (s, 1H), 8.42 (s, 1H), 8.04 – 7.99 (m, 2H), 7.97 – 7.93 (m, 2H), 7.84 – 7.80 (m, 2H), 7.71 – 7.67 (m, 2H), 7.58 – 7.51 (m, 2H), 7.47 – 7.36 (m, 6H), 7.32 – 7.22 (m, 4H), 6.34 (d, *J* = 8.9 Hz, 1H), 6.19 – 6.10 (m, 2H), 5.93 (t, *J* = 9.5 Hz, 1H), 4.70 (dd, *J* = 12.9, 3.3 Hz, 1H), 4.56 – 4.48 (m, 2H), 1.33 (s, 18H).

***N*-tert-Butoxycarbonyl-9-(2',3',4',6'-tetra-*O*-benzoyl-β-D-glucopyranosyl)adenine (S2)**



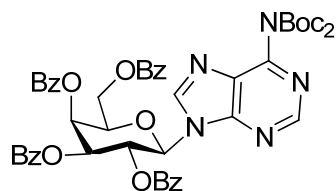
<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 8.67 (s, 1H), 8.31 (s, 1H), 8.02 (d, *J* = 7.2 Hz, 2H), 7.96 – 7.91 (m, 2H), 7.84 – 7.79 (m, 2H), 7.70 (d, *J* = 7.3 Hz, 2H), 7.59 – 7.50 (m, 2H), 7.47 – 7.35 (m, 6H), 7.31 – 7.23 (m, 4H), 6.31 (d, *J* = 9.2 Hz, 1H), 6.16 (t, *J* = 9.6 Hz, 1H), 6.08 (t, *J* = 9.4 Hz, 1H), 5.92 (t, *J* = 9.6 Hz, 1H), 4.69 (d, *J* = 9.8 Hz, 1H), 4.56 – 4.47 (m, 2H), 1.53 (s, 9H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 166.14, 165.67, 165.27, 165.02, 153.10, 151.22, 149.92, 149.50, 140.21, 133.97, 133.85, 133.62, 133.46, 130.05, 129.92, 129.89, 129.47, 128.67, 128.59, 128.55, 128.52, 127.70, 121.14, 82.70, 81.08, 75.86, 73.11, 71.33, 69.02, 62.66, 28.21. HRMS (ESI) calcd for C<sub>44</sub>H<sub>40</sub>N<sub>5</sub>O<sub>11</sub> [M+H]<sup>+</sup>:814.2724, found: 814.2717.

**2',3',4',6'-Tetra-*O*-benzoyl-β-D-galactopyranosyl-2,6-dichloropurine (27)**



(*R*<sub>f</sub> = 0.51, petroleum: acetone = 2:1) <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 8.50 (s, 1H, H-8), 8.17 (d, *J* = 7.4 Hz, 2H), 7.98 (d, *J* = 7.4 Hz, 2H), 7.77 (dd, *J* = 18.5, 7.5 Hz, 4H), 7.68 (d, *J* = 7.4 Hz, 1H), 7.55 (dd, *J* = 14.4, 7.2 Hz, 3H), 7.49 – 7.38 (m, 4H), 7.34 – 7.23 (m, 4H), 6.40 (t, *J* = 9.7 Hz, 1H, H-2'), 6.26 – 6.15 (m, 2H, H-1' and H-4'), 5.92 (dd, *J* = 10.1, 3.1 Hz, 1H, H-3'), 4.76 – 4.62 (m, 2H, H-5' and H-6'), 4.52 (dd, *J* = 11.4, 5.4 Hz, 1H, H-6'). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 166.09, 165.45, 165.43, 165.05(4 Bz), 153.67, 153.11, 152.40(C-2, C-4 and C-6), 143.82(C-8), 134.12, 134.09, 133.72, 133.57, 130.96(C-5), 130.12, 129.91, 129.88, 129.18, 129.04, 128.81, 128.64, 128.60, 128.53, 128.45, 127.62, 82.25(C-1'), 75.03(C-5'), 71.72(C-3'), 68.87(C-2'), 68.06(C-4'), 62.12(C-6'). HRMS (ESI) calcd for C<sub>39</sub>H<sub>29</sub>N<sub>4</sub>O<sub>9</sub>Cl<sub>2</sub> [M+H]<sup>+</sup>:767.1312, found: 767.1305.

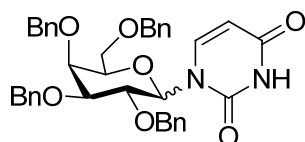
***N,N*-Di-*tert*-butoxycarbonyl-9-(2',3',4',6'-tetra-*O*-benzoyl-β-D-galactopyranosyl)adenine (28)**



<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 8.85 (s, 1H, H-2), 8.49 (s, 1H, H-8), 8.22 – 8.16 (m, 2H), 8.02 – 7.97 (m, 2H), 7.83 – 7.77 (m, 2H), 7.69 (dd, *J* = 7.2, 4.7 Hz, 3H), 7.56 (dd, *J* = 16.1, 7.9 Hz, 3H), 7.48 – 7.38 (m, 4H), 7.29 – 7.22 (m, 4H), 6.49 (t, *J* = 9.7 Hz, 1H, H-2'), 6.31 (d, *J* = 9.4 Hz, 1H, H-1'), 6.19 (d, *J* = 3.0

Hz, 1H, H-4'), 5.90 (dd,  $J = 10.1, 3.2$  Hz, 1H, H-3'), 4.72-4.62 (m, 2H, H-5' and H-6'), 4.51 (dd,  $J = 11.1, 5.7$  Hz, 1H, H-6'), 1.35 (s, 18H, Boc).  $^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  166.10, 165.45, 165.42, 164.98(4Bz), 153.26(C-4), 152.56(C-2), 150.82(C-6), 150.17(Boc), 142.51(C-8), 134.06, 133.86, 133.65, 133.52, 130.12, 129.93, 129.91, 129.79, 129.25, 128.99, 128.96, 128.76, 128.58, 128.49, 128.46, 127.85(C-5), 83.85(Boc), 81.68(C-1'), 74.70(C-5'), 72.05(C-3'), 68.87(C-2'), 68.12(C-4'), 62.04(C-6'), 27.76(Boc). HRMS (ESI) calcd for  $\text{C}_{49}\text{H}_{48}\text{N}_5\text{O}_{13}$   $[\text{M}+\text{H}]^+$ :914.3249, found: 914.3247.

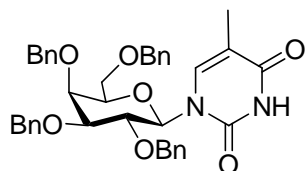
### 1-(2',3',4',6'-Tetra-*O*-benzyl- $\beta$ -D-galactopyranosyl)-uracil (29)



**For the  $\alpha$ -isomer:**  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.06 (s, 1H, NH), 7.59 (d,  $J = 8.2$  Hz, 1H, H-6), 7.39 – 7.24 (m, 18H, benzyl), 7.02 (dd,  $J = 6.6, 2.9$  Hz, 2H, benzyl), 6.04 (d,  $J = 1.8$  Hz, 1H, H-1'), 5.60 (dd,  $J = 8.2, 2.1$  Hz, 1H, H-5), 4.65 – 4.47 (m, 7H), 4.32 (d,  $J = 11.7$  Hz, 1H), 4.16 (dt,  $J = 11.9, 4.6$  Hz, 2H), 4.05 (dd,  $J = 6.4, 2.8$  Hz, 1H), 3.83 – 3.77 (m, 2H), 3.68 (dd,  $J = 3.9, 1.9$  Hz, 1H).

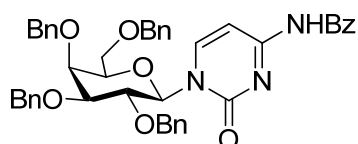
**For the  $\beta$ -isomer:**  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.13 (s, 1H, NH), 7.41 – 7.22 (m, 18H, benzyl), 7.17 (dd,  $J = 6.5, 3.1$  Hz, 2H, benzyl), 6.94 (d,  $J = 8.2$  Hz, 1H, H-6), 5.58 (d,  $J = 8.9$  Hz, 1H, H-1'), 5.34 (dd,  $J = 8.1, 2.1$  Hz, 1H, H-5), 4.95 (d,  $J = 11.3$  Hz, 1H,  $\text{CH}_2\text{-Ph}$ ), 4.84 – 4.72 (m, 3H,  $\text{CH}_2\text{-Ph}$ ), 4.57 (dd,  $J = 16.6, 11.6$  Hz, 2H,  $\text{CH}_2\text{-Ph}$ ), 4.44 (q,  $J = 11.9$  Hz, 2H,  $\text{CH}_2\text{-Ph}$ ), 3.99 (d,  $J = 2.0$  Hz, 1H, H-4'), 3.87 (t,  $J = 9.2$  Hz, 1H, H-3'), 3.78 – 3.70 (m, 2H, H-2', H-5'), 3.58 – 3.47 (m, 2H, H-6'a, H-6'b).

### 1-(2',3',4',6'-Tetra-*O*-benzyl- $\beta$ -D-galactopyranosyl)-thymine (30)



( $R_f = 0.31$ , petroleum: ethyl acetate = 1:1)  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.37 (s, 1H, NH), 7.42 – 7.13 (m, 20H, benzyl), 6.65 (d,  $J = 1.2$  Hz, 1H, H-6), 5.58 (d,  $J = 8.9$  Hz, 1H, H-1'), 4.97 (d,  $J = 11.4$  Hz, 1H,  $\text{CH}_2\text{Ph}$ ), 4.83 – 4.72 (m, 3H,  $\text{CH}_2\text{Ph}$ ), 4.58 (dd,  $J = 20.0, 11.7$  Hz, 2H,  $\text{CH}_2\text{Ph}$ ), 4.45 (q,  $J = 11.9$  Hz, 2H,  $\text{CH}_2\text{Ph}$ ), 4.00 (d,  $J = 1.9$  Hz, 1H), 3.85 (t,  $J = 9.1$  Hz, 1H), 3.79 – 3.70 (m, 2H), 3.53 (ddd,  $J = 15.0, 9.2, 6.6$  Hz, 2H, H-6'a, H-6'b), 1.63 (d,  $J = 1.1$  Hz, 3H,  $\text{CH}_3$ .thymine).

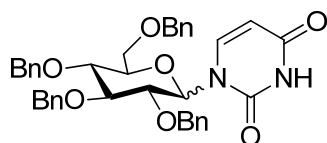
### $N^4$ -benzoyl-1-(2',3',4',6'-tetra-*O*-benzyl- $\beta$ -D-galactopyranosyl)-cytosine (31)



( $R_f = 0.49$ , petroleum: ethyl acetate = 1:2)  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.94 (s, 2H), 7.62 (t,  $J = 7.4$  Hz, 1H), 7.53 (t,  $J = 7.6$  Hz, 2H), 7.42 – 7.10 (m, 22H), 5.88 (d,  $J = 8.2$  Hz, 1H, H-1'), 4.98 (d,  $J = 11.3$  Hz, 1H,  $\text{CH}_2\text{Ph}$ ), 4.84 – 4.72 (m, 3H,  $\text{CH}_2\text{Ph}$ ), 4.61 (d,  $J = 11.3$  Hz, 1H,  $\text{CH}_2\text{Ph}$ ), 4.55 – 4.38 (m, 3H,  $\text{CH}_2\text{Ph}$ ), 4.01 (d,  $J = 2.0$  Hz, 1H), 3.91 (t,  $J = 9.2$  Hz, 1H), 3.84 – 3.74 (m, 2H), 3.61 – 3.48 (m, 2H).



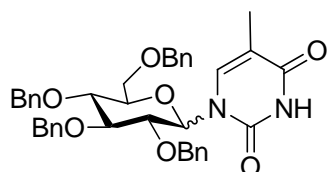
### 1-(2',3',4',6'-Tetra-*O*-benzyl-*D*-glucopyranosyl)-uracil (32)



**For the  $\alpha$ -isomer:**  $^1\text{H NMR}$  (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.00 (s, 1H), 7.62 (d,  $J = 8.2$  Hz, 1H), 7.39 – 7.27 (m, 16H), 7.19 (dd,  $J = 7.0, 2.4$  Hz, 2H), 7.14 (dd,  $J = 6.6, 2.7$  Hz, 2H), 6.03 (d,  $J = 3.2$  Hz, 1H), 5.62 (dd,  $J = 8.2, 2.2$  Hz, 1H), 4.59 – 4.40 (m, 7H), 4.35 – 4.28 (m, 2H), 4.05 (t,  $J = 3.0$  Hz, 1H), 3.83 (t,  $J = 3.1$  Hz, 1H), 3.73 (dd,  $J = 10.6, 2.7$  Hz, 1H), 3.70 – 3.62 (m, 2H).

**For the  $\beta$ -isomer:**  $^1\text{H NMR}$  (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.44 (d,  $J = 1.5$  Hz, 1H), 7.38 – 7.26 (m, 16H), 7.17 (td,  $J = 8.1, 3.1$  Hz, 4H), 6.89 (d,  $J = 8.2$  Hz, 1H), 5.61 (d,  $J = 9.1$  Hz, 1H), 5.37 (dd,  $J = 8.1, 2.2$  Hz, 1H), 4.93 (s, 2H), 4.82 (dd,  $J = 22.9, 11.3$  Hz, 2H), 4.60 (d,  $J = 10.8$  Hz, 1H), 4.56 – 4.45 (m, 3H), 3.86 (t,  $J = 9.0$  Hz, 1H), 3.75 – 3.58 (m, 4H), 3.47 (t,  $J = 9.0$  Hz, 1H)

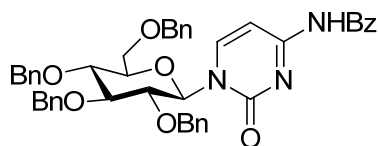
### 1-(2',3',4',6'-Tetra-*O*-benzyl-*D*-glucopyranosyl)-thymine (33)



**For the  $\alpha$ -isomer:**  $^1\text{H NMR}$  (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.90 (s, 1H), 7.45 (d,  $J = 1.1$  Hz, 1H), 7.39 – 7.27 (m, 16H), 7.20 (dd,  $J = 7.0, 2.4$  Hz, 2H), 7.12 (dd,  $J = 7.3, 2.0$  Hz, 2H), 6.04 (d,  $J = 3.2$  Hz, 1H), 4.59 – 4.49 (m, 5H), 4.44 (t,  $J = 11.1$  Hz, 2H), 4.36 – 4.27 (m, 2H), 4.03 (t,  $J = 3.0$  Hz, 1H), 3.83 (t,  $J = 3.1$  Hz, 1H), 3.74 (dd,  $J = 10.5, 3.0$  Hz, 1H), 3.69 – 3.63 (m, 2H), 1.81 (d,  $J = 0.9$  Hz, 3H).

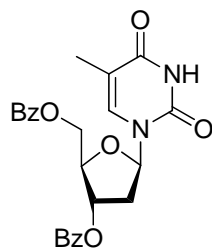
**For the  $\beta$ -isomer:**  $^1\text{H NMR}$  (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.07 (s, 1H), 7.38 – 7.20 (m, 16H), 7.17 (dd,  $J = 5.7, 3.7$  Hz, 4H), 6.65 (d,  $J = 1.1$  Hz, 1H), 5.59 (d,  $J = 9.2$  Hz, 1H), 4.94 (s, 2H), 4.81 (dd,  $J = 21.7, 11.4$  Hz, 2H), 4.59 (d,  $J = 10.7$  Hz, 1H), 4.55-4.45 (m, 3H), 3.86 (t,  $J = 9.0$  Hz, 1H), 3.77 – 3.68 (m, 2H), 3.65 (dd,  $J = 11.0, 1.7$  Hz, 1H), 3.62 – 3.56 (m, 1H), 3.46 (t,  $J = 9.0$  Hz, 1H), 1.67 (d,  $J = 0.9$  Hz, 3H).

### *N*<sup>4</sup>-benzoyl-1-(2',3',4',6'-tetra-*O*-benzyl- $\beta$ -*D*-glucopyranosyl)-cytosine (34)



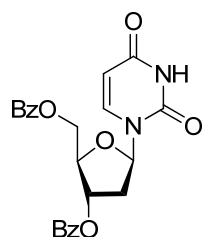
$^1\text{H NMR}$  (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.94 (s, 2H), 7.64 (t,  $J = 7.4$  Hz, 1H), 7.54 (t,  $J = 7.6$  Hz, 2H), 7.38 – 7.27 (m, 14H), 7.23 – 7.10 (m, 8H), 5.91 (d,  $J = 7.7$  Hz, 1H, H-1'), 4.97 – 4.89 (m, 2H,  $\text{CH}_2\text{Ph}$ ), 4.95-4.80 (m, 1H,  $\text{CH}_2\text{Ph}$ ), 4.74 (d,  $J = 11.8$  Hz, 1H,  $\text{CH}_2\text{Ph}$ ), 4.61 (d,  $J = 10.7$  Hz, 1H,  $\text{CH}_2\text{Ph}$ ), 4.54 (d,  $J = 12.2$  Hz, 1H,  $\text{CH}_2\text{Ph}$ ), 4.47 (dd,  $J = 12.0, 3.2$  Hz, 2H,  $\text{CH}_2\text{Ph}$ ), 3.91 (t,  $J = 8.9$  Hz, 1H, H-4'), 3.82 – 3.72 (m, 2H, H-3' and H-2'), 3.70 – 3.62 (m, 2H), 3.52 (t,  $J = 8.7$  Hz, 1H).

**1-(3',5'-Di-*O*-benzoyl-2'-deoxy- $\beta$ -D-ribofuranosyl)-thymine (35)**



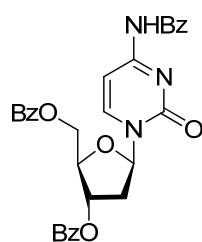
( $R_f$  = 0.50, petroleum: ethyl acetate = 1:2 )  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.77 (s, 1H), 8.06 (t,  $J$  = 8.2 Hz, 4H), 7.62 (t,  $J$  = 7.4 Hz, 2H), 7.48 (t,  $J$  = 7.0 Hz, 4H), 7.26 (s, 1H, H-6), 6.47 (dd,  $J$  = 8.6, 5.6 Hz, 1H, H-1'), 5.66 (d,  $J$  = 6.4 Hz, 1H, H-3'), 4.80 (dd,  $J$  = 12.2, 2.8 Hz, 1H, H-5'), 4.69 (dd,  $J$  = 12.2, 3.3 Hz, 1H, H-5'), 4.54 (d,  $J$  = 2.2 Hz, 1H, H-4'), 2.72 (dd,  $J$  = 14.2, 5.2 Hz, 1H, H-2'), 2.41 – 2.29 (m, 1H, H-2'), 1.62 (s, 3H, - $\text{CH}_3$ ).

**1-(3',5'-Di-*O*-benzoyl-2'-deoxy- $\beta$ -D-ribofuranosyl)-uracil (36)**



( $R_f$  = 0.43, petroleum: ethyl acetate = 1:2 )  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.10 – 8.05 (m, 3H), 8.02 (dd,  $J$  = 8.4, 1.3 Hz, 2H), 7.65 – 7.59 (m, 2H), 7.54 – 7.45 (m, 5H), 6.40 (dd,  $J$  = 8.4, 5.6 Hz, 1H), 5.63 (dt,  $J$  = 6.5, 2.1 Hz, 1H), 5.59 (dd,  $J$  = 8.2, 2.3 Hz, 1H), 4.77 – 4.68 (m, 2H), 4.56 (dd,  $J$  = 6.0, 3.4 Hz, 1H), 2.76 (ddd,  $J$  = 14.4, 5.7, 2.0 Hz, 1H), 2.33 (ddd,  $J$  = 14.6, 8.3, 6.7 Hz, 1H).

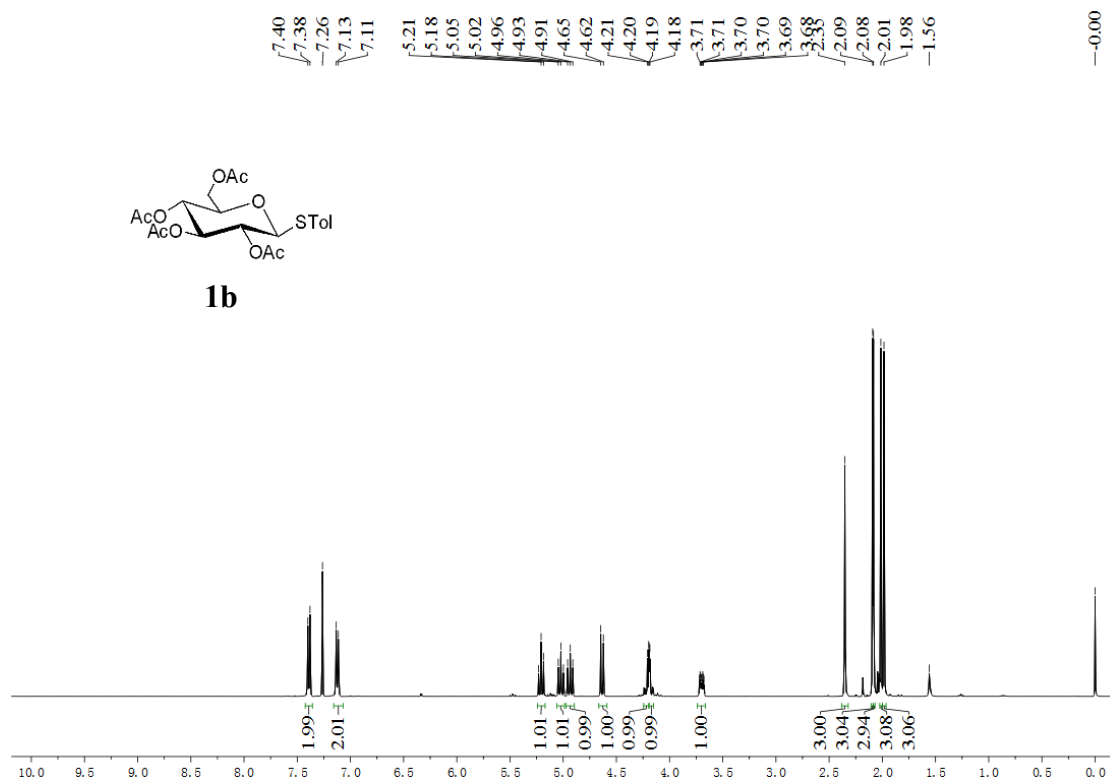
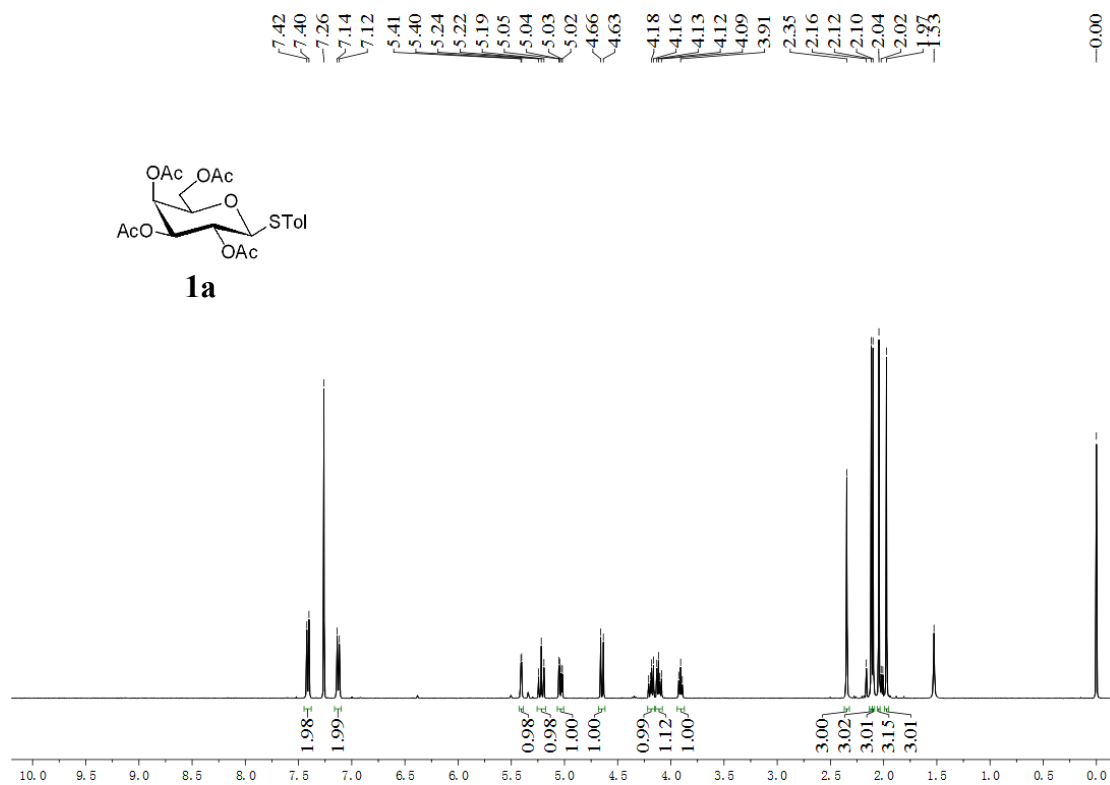
**$N^4$ -benzoyl-1-(3',5'-Di-*O*-benzoyl-2'-deoxy- $\beta$ -D-ribofuranosyl)-cytosine (37)**

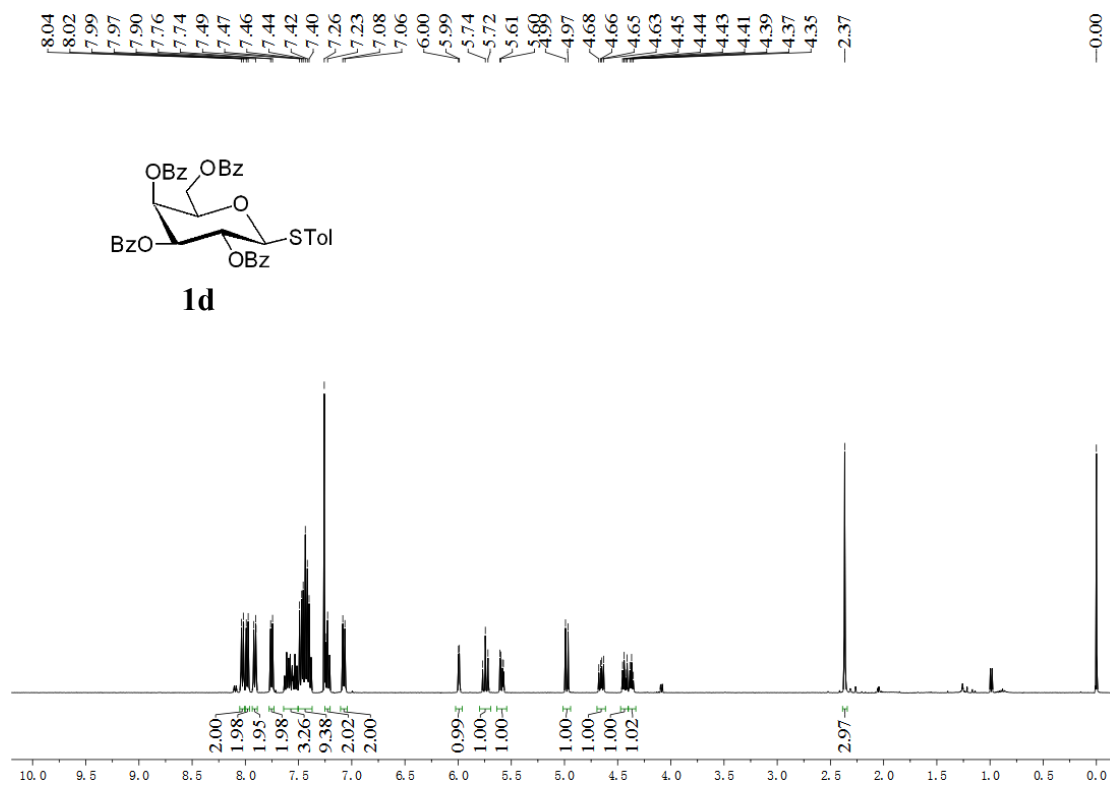
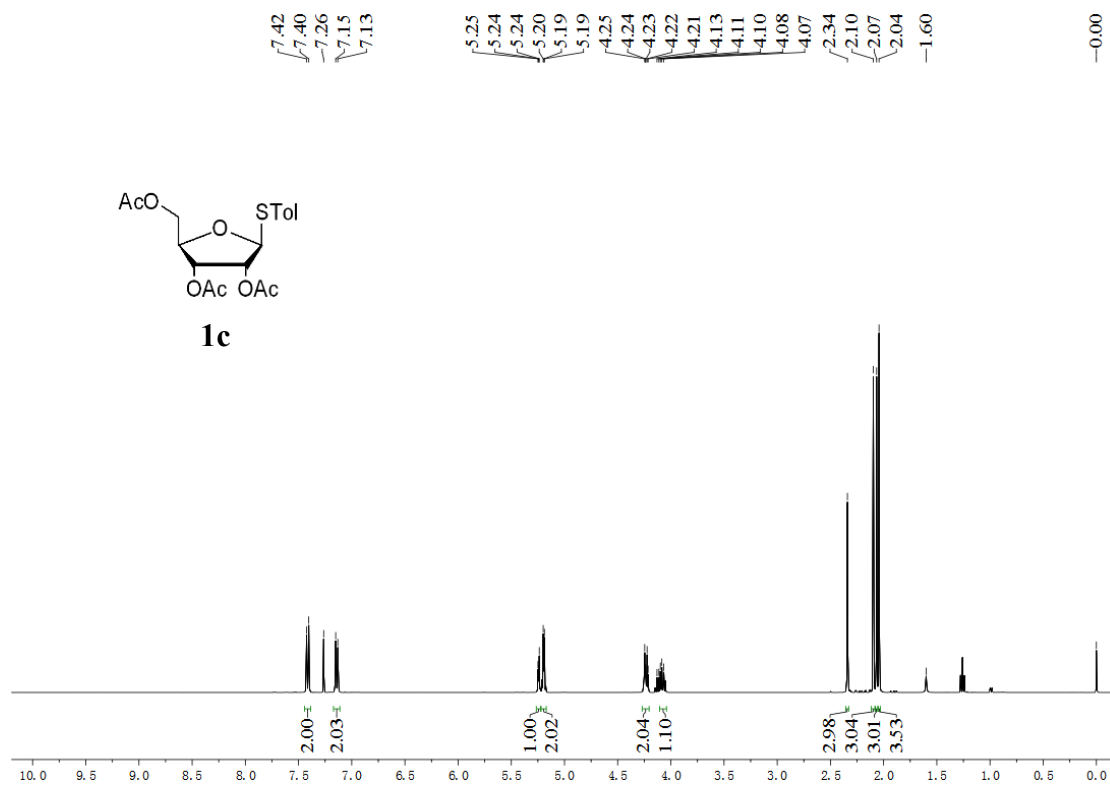


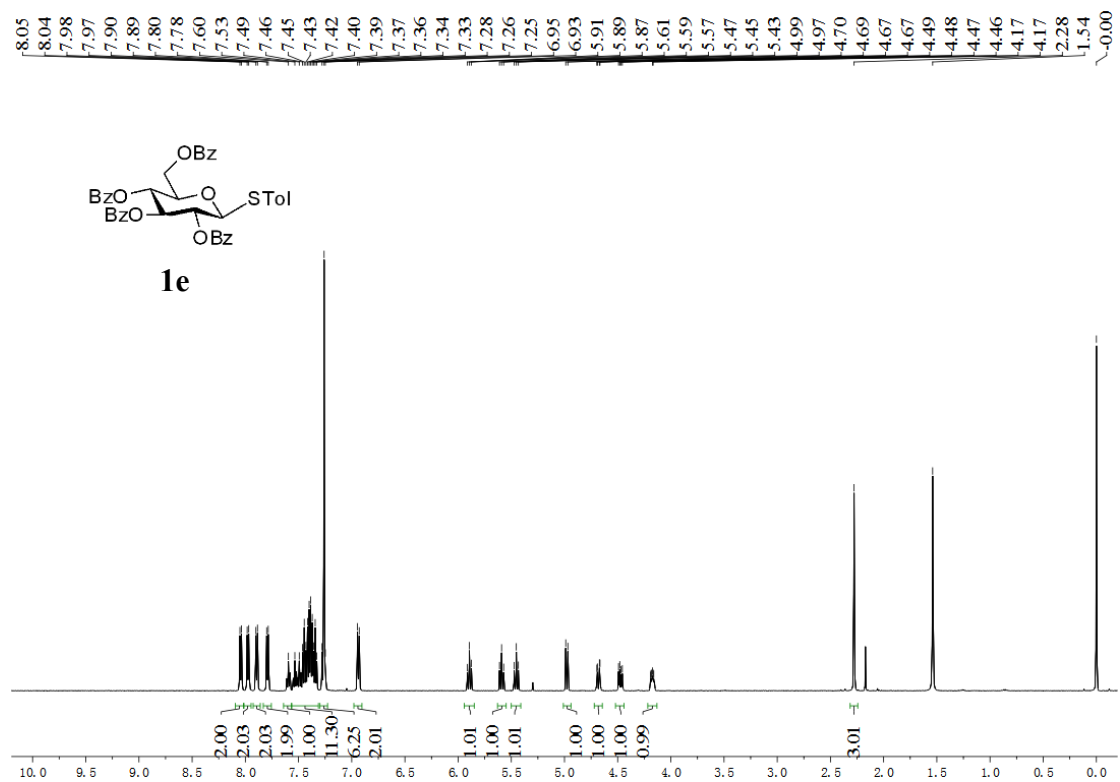
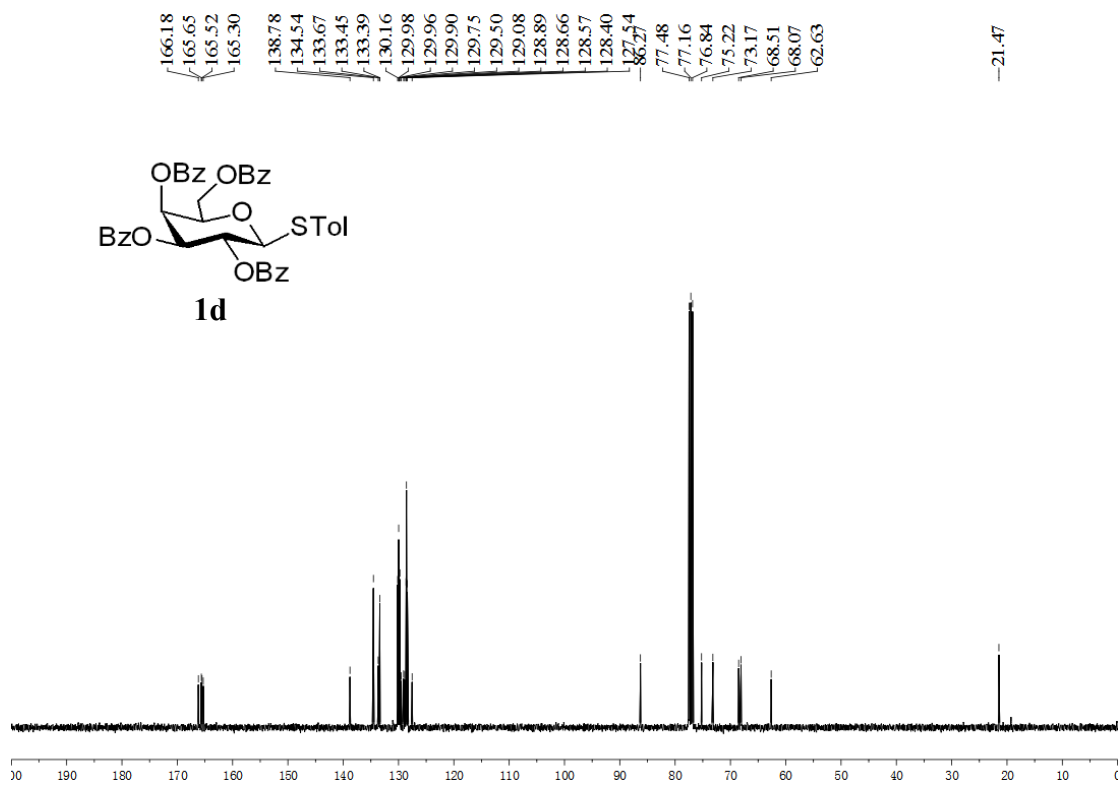
( $R_f$  = 0.32, petroleum: ethyl acetate = 1:2 )  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.14 - 8.05 (m, 3H), 8.03 – 7.97 (m, 2H), 7.89 (d,  $J$  = 7.5 Hz, 2H), 7.65 - 7.42 (m, 10H), 6.41 (dd,  $J$  = 7.9, 5.6 Hz, 1H, H-1'), 5.64 (dt,  $J$  = 6.4, 1.9 Hz, 1H), 4.76 (qd,  $J$  = 12.2, 3.5 Hz, 2H), 4.67 (dd,  $J$  = 5.8, 3.3 Hz, 1H), 3.16 – 3.06 (m, 1H, H-2'), 2.31 (ddd,  $J$  = 14.5, 7.8, 6.6 Hz, 1H, H-2').

## References

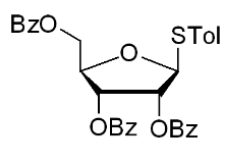
1. C. Chao, M. Chen, S. Lin and K. T. Mong, *Carbohydr. Res.*, 2008, **343**, 957-964.
2. F. Zhang, W. Zhang, Y. Zhang, D. P. Curran and G. Liu, *J. Org. Chem.*, 2009, **74**, 2594-2597.
3. M. Kurosu and K. Li, *J. Org. Chem.*, 2008, **73**, 9767-9770.
4. R. R. France, N. V. Rees, J. D. Wadhawan, A. J. Fairbanks and R. G. Compton, *Org. Biomol. Chem.*, 2004, **2**, 2188-2194.
5. S. Hsieh, M. Jan, L. N. Patkar, C. Chen and C. Lin, *Carbohydr. Res.*, 2005, **340**, 49-57.
6. K. W. C. G. Roland Haeckel and M. W. A. F. Joseph Eisenbarth, *Journal of Labelled Compounds and Radiopharmaceuticals*, 1996, 1061-1070.
7. B. Gerland, A. Goudot, C. Ligeour, G. Pourceau, A. Meyer, S. Vidal, T. Gehin, O. Vidal, E. Souteyrand, J. Vasseur, Y. Chevolut and F. Morvan, *Bioconjugate Chem.*, 2014, **25**, 379-392.
8. Q. Zhang, J. Sun, Y. Zhu, F. Zhang and B. Yu, *Angew. Chem. Int. Ed.*, 2011, **50**, 4933-4936.
9. J. Huchting and C. Meier, *Eur. J. Org. Chem.*, 2014, **2014**, 3423-3429.
10. I. Nowak and M. J. Robins, *Org. Lett.*, 2005, **7**, 4903-4905.
11. K. C. A. S. Danishefsky, *J. Org. Chem.*, 1990, **55**, 4211-4214.
12. J. W. Rigoli, M. E. Østergaard, K. M. Canady, D. C. Guenther and P. J. Hrdlicka, *Tetrahedron Lett.*, 2009, **50**, 1751-1753.
13. J. Liao, J. Sun and B. Yu, *Tetrahedron Lett.*, 2008, **49**, 5036-5038.
14. T. M. T. M. Naoyuki SHIMOMURA, *Bull. Chem. Soc. Jpn*, 1994, **67**, 3100-3106.
15. B. Fraser-Reid, P. Ganney, C. V. S. Ramamurty, A. M. Gómez and J. C. López, *Chem. Commun*, 2013, **49**, 3251-3253.
16. J. J. C. H. Stephen Hanessian, *Tetrahedron*, 1996, **52**, 10827-11083.
17. A. K. Prasad, V. Kumar, J. Maity, W. Zhiwei, V. T. Ravikumar, Y. S. Sanghvi and V. S. Parmar, *Synth. Commun.*, 2005, **35**, 935-945.



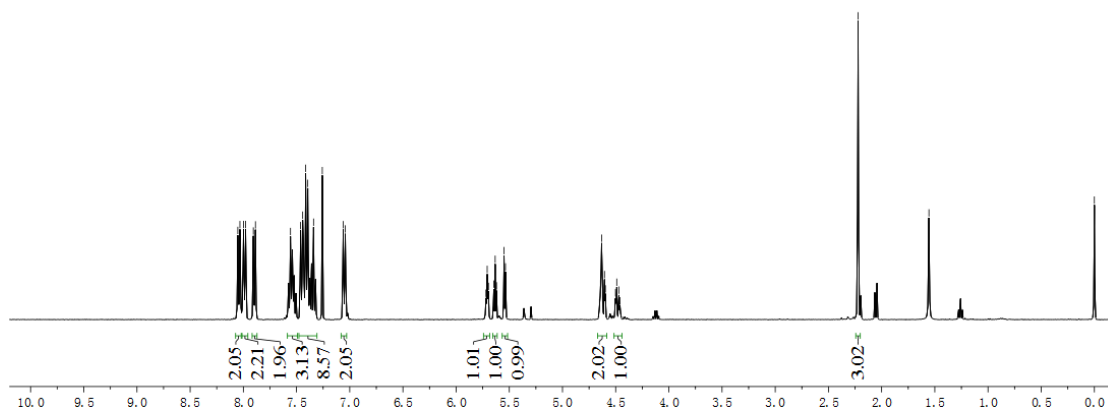




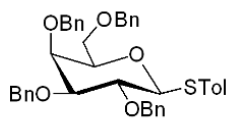
8.05  
8.03  
8.00  
7.98  
7.91  
7.89  
7.56  
7.46  
7.44  
7.41  
7.39  
7.34  
7.26  
7.06  
7.04  
5.71  
5.65  
5.63  
5.62  
5.55  
5.54  
4.64  
4.63  
4.61  
4.60  
4.50  
4.49  
4.47  
4.46  
-2.22  
-1.56  
-0.00



**1f**



7.47  
7.45  
7.38  
7.34  
7.33  
7.33  
7.32  
7.32  
7.30  
7.29  
7.29  
7.27  
7.25  
7.25  
6.99  
6.97  
4.78  
4.74  
4.72  
4.71  
4.59  
4.57  
4.45  
4.42  
3.97  
3.96  
3.89  
3.65  
3.64  
3.60  
3.59  
3.58  
3.28  
-0.00



**1g**

