Bioorganic & Medicinal Chemistry Letters 19 (2009) 6429–6432

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry Letters

journal homepage: www.elsevier.com/locate/bmcl

# Systematic study on the broad nucleotide triphosphate specificity of the pyrophosphorylase domain of the *N*-acetylglucosamine-1-phosphate uridyltransferase from *Escherichia coli* K12

Junqiang Fang [a], Wanyi Guan [a,b], Li Cai [b], Guofeng Gu [a,*], Xianwei Liu [a,*], Peng George Wang [a,b]

[a] National Glycoengineering Research Center and The State Key Laboratory of Microbial Technology, Shandong University, Jinan, Shandong 250100, People's Republic of China
[b] Departments of Chemistry and Biochemistry, The Ohio State University, 484 W. 12th Ave., Columbus, OH 43210, USA

ARTICLE INFO

*Article history:*
Received 26 July 2009
Revised 4 September 2009
Accepted 11 September 2009
Available online 17 September 2009

*Keywords:*
*N*-Acetylglucosamine-1-phosphate uridyltransferase
Pyrophosphorylase domain
Biosynthesis
Substrate specificity

ABSTRACT

*N*-Acetylglucosamine-1-phosphate uridyltransferase (GlmU) from *Escherichia coli* K12 is a bifunctional enzyme that catalyzes both the acetyltransfer and uridyltransfer reactions in the prokaryotic UDP-GlcNAc biosynthetic pathway. In this study, we report the broad substrate specificity of the pyrophosphorylase domain of GlmU during its uridyltransfer reaction and the substrate priority is ranked in the following order: UTP > dUTP > dTTP >> CTP > dATP/dm$^6$ ATP. This pyrophosphorylase domain of GlmU is also a tool to synthesize UDP-GlcNAc analogs, two examples of which were synthesized herein in multiple mg scale in vitro.

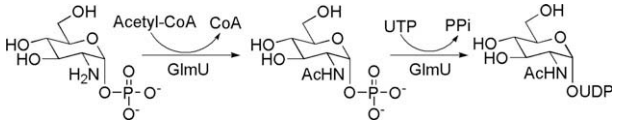


Uridine 5′-diphosphate *N*-acetylglucosamine (UDP-GlcNAc), the key cytoplasmic amino nucleotide sugar, is an essential precursor of the biosynthesis of various cell components including cell wall peptidoglycans,[1] lipopolysaccharides,[2] enterobacterial common antigens,[3] chitin,[4] glycosylphophatidylinositol anchors,[5,6] glycosaminoglycans[7] and glycoproteins.[5] One of the major bottlenecks to the basic research and application of these components is the limited availability of sugar nucleotide analogs and derivatives. Thus, practical synthetic approaches towards UDP-GlcNAc and its analogs are worthwhile for exploitation.

Chemical synthesis of UDP-GlcNAc has been developed,[8] but its application is hindered by low yield and fastidious process. In contrast, enzymatic synthesis by following the biosynthetic pathway was considered more attractive with the advantages of high regio-/stereo-specificity and low cost.[9,10] Several UDP-GlcNAc biosynthetic pathways have been discovered in both eukaryotes[11] and prokaryotes.[12,13] In some prokaryotes, a bifunctional enzyme, *N*-acetylglucosamine-1-phosphate uridyltransferase (GlmU, E.C. 2.3.1.157 and E.C. 2.7.7.23), with both *N*-acetyltransferase and pyrophosphorylase activities, has been found to catalyze the last two sequential reactions in the UDP-GlcNAc biosynthesis.[12,14] It utilizes glucosamine-1-phosphate, UTP and acetyl-CoA to afford the final product UDP-GlcNAc (Scheme 1).[15]

The crystal structure of GlmU from *Escherichia coli* K12 strain (RSB Protein Data Bank 1HV9, Fig. 1a) has already been solved, which indicates that the full-length GlmU has two distinct autonomous catalytic sites:[16] The C-terminal domain, which adopts a left-handed parallel β-helix (LβH) structure[16], catalyzes the CoA-dependent N-acetylation of glucosamine-1-phosphate, while the pyrophosphorylase domain at the N-terminal, which adopts a globular structure, shares similarity with the nucleotide diphosphate sugar pyrophosphorylase family with the strict conservation motif of L-$X_2$-G-X-C-T-X-M-$X_4$-P-K.[17] The *N*-acetyltransferase domain and the pyrophosphorylase domain are connected by a linker containing a 21-amino-acid α-helices.[12] The contacts between these two domains involve only van der Waals interactions between Ala31-Gly32 in the surface loop of the pyrophosphorylase domain and the Arg263 side chain in the *N*-acetyltransferase domain.[17] The kinetic study indicated that acetyl-transfer precedes uridyl-transfer in the formation of UDP-GlcNAc.[18]

A serendipitous generated N-terminal proteolytic fragment of GlmU (M1-R331, GlmU-Tr) has been isolated with only pyrophos-




**Scheme 1.** Synthesis of UDP-GlcNAc from GlcN-1-P catalyzed by GlmU.

* Corresponding authors. Tel./fax: +86 531 88366078.
*E-mail addresses:* guofenggu@sdu.edu.cn (G. Gu), xianweiliu@mail.sdu.edu.cn (X. Liu).

0960-894X/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2009.09.039

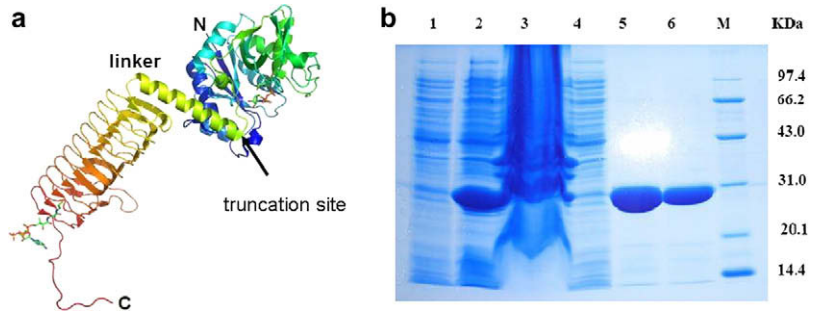



**Figure 1.** (a) Crystal structure of *Escherichia coli* GlmU from RCSB Protein Data Bank (1HV9). N: Pyrophosphorylase domain, M1-R229; C: Acetyltranferase domain, G251-K456; Linkage: L230-A250; Black arrow: truncation site for GlmU-Tr229. (b) SDS–PAGE analysis of the expression and purification process of the pyrophosphorylase domain of GlmU. Lane: 1, whole cell lysate of the cells transformed with vector alone; 2, whole cell lysate of the cells with the expression vector; 3, cell pellet; 4, flow through during purification; 5,6: Eluted target protein; M, protein standards (the molecular weights are indicated in kDa on the right side).

phorylase activity.[18] Detailed studies on this proteolytic fragment have been performed, including determination of kinetic parameters, resolving of GlmU-Tr structure and elucidation of its possible active site. According to its structure, we further truncated GlmU and obtained a fragment (M1-R229, GlmU-Tr229) possessing the pyrophosphorylase domain. The fragment was overexpressed (Fig. 1b) and its substrate specificity towards sugar-1-phosphate (sugar-1-P) and nucleotide triphosphate (NTP) was investigated. And we synthesized two UDP-GlcNAc analogs with the recombinant enzyme in multiple mg scale.

To investigate the substrate specificity towards NTP, GlcNAc-1-P was used as the sugar-1-P substrate. Twelve NTPs were included into the reaction system, respectively (Table 1). Capillary electrophoresis (CE) was used to detect the formation of NTP-GlcNAcs. Typically, enzymatic reactions were carried out under standard conditions and were initiated with the addition of purified GlmU-Tr229 (1.5 mg) to a solution containing 100 mM Tris–HCl (pH 7.5), 100 mM NaCl, 5 mM sugar-1-P, 5 mM NTP, 5 mM $MgCl_2$, 4 mM DTT, 1 U/mL pyrophosphatase and 1% BSA in a final volume of 100 μL. The reaction was performed at 37 °C for 30 min and quenched by boiling at 100 °C for 1 min followed by centrifugation at 13,000 rpm for 15 min. The yields of NTP-GlcNAcs showed in Table 1 indicated a substrate tolerance of GlmU-Tr229 for NTPs in the following order: UTP > dUTP > dTTP >> CTP > dATP/dm$^6$ ATP, suggesting that GlmU-Tr229 prefers pyrimidine nucleotide than purine nucleotide.

The crystal structure of GlmU showed that the active site pocket of the pyrophosphorylase domain was bound by two lobes. The first lobe (residues Asn3-Val111 and His216-227) included strands β1–β4, which interacted primarily with the nucleotide moiety of UDP-GlcNAc. The second lobe (residues Glu112-Val215) included strands β5–β7, which interacted primarily with GlcNAc portion of the sugar nucleotide. Uracil base was recognized by a hydrogen bond between its ring N3 and Gln76 residue, and by two hydrogen bonds between its exocyclic oxygen O4 and Gly81 and Gln76, respectively. The ribose group was recognized by a hydrogen bond between its 2′-hydroxyl group and Gly14.[14] Our results indicated that pyrophosphorylase domain of GlmU had a notable substrate tolerance, small changes on ribose 2′-hydroxyl group (dUTP) or uracil base C5 (dTTP) do not affect the pyrophosphorylase activity. The hydrogen bond between ribose group 2′-hydroxyl group and Gly14 was not critical for substrate recognition. The first lobe which contained the pyrophosphorylase active site was suitable to bear the existence of a methyl group in uracil base C5.

We further investigated the substrate tolerance of the recombinant GlmU-Tr229 for sugar-1-P. Other than its natural substrate GlcNAc-1-P, glucose-1-phoshate (Glc-1-P) and *N*-azidoacetyl-D-glucosamine-1-phosphate (GlcNAcZ-1-P) were tested. These sugar-1-Ps were synthesized as previously described.[19] For each sugar-1-P substrate, UTP, dUTP and dTTP were used as the pyrophosphate donor, respectively. All reactions were carried out and detected the same as for the NTP specificity study.

GlmU-Tr229 exhibited activity for all three sugar-1-Ps with different NTPs to form nine UDP-GlcNAc analogs, however, with diverse yields (Table 2). The yields for GlcNAcZ-1-P decreased to 86%, 92% and 87% of those for GlcNAc-1-P with UTP, dUTP and dTTP as pyrophosphate donor. While the conversion efficiencies for Glc-1-P which lacks the C2 *N*-acetyl group dropped to around 30% of those for GlcNAc-1-P. Previous study suggested that GlcNAc-1-P interacted with the enzyme by hydrogen bonds between the *N*-acetyl arm and Thr82 and Glu154, as well as by van der Waals interaction between the methyl group and Tyr197.[17] The significant decline of GlmU-Tr229 activity for Glc-1-P compared with that for GlcNAc-1-P was consistent with the structure study, which demonstrated the important role of Thr82 and Glu154 residues in sugar-1-P recognition. In contrast, GlmU-Tr229 had better tolerance for the modifications on the *N*-acyl group, as indicated by the result for GlcNAcZ-1-P.

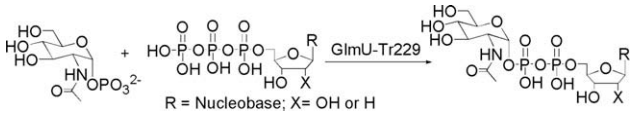
To compare the enzymatic activities on different substrates, the rates of the forward reaction with different enzyme concentration were measured in small amount reactions (100 μL). Figure 2 demonstrated the relationship of the rate of forward reaction to the lineage range of enzyme dependency. The results showed the GlmU-Tr229 activities rank of forward reactions as UDP-GlcNAc > UDP-GlcNAcZ > dUDP-GlcNAc > dUDP-GlcNAcZ > dTDP-GlcNAc > dTDP-GlcNAcZ > NDP-Glc, which were in line with the data from the preparation reactions in Table 2.

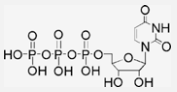
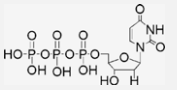
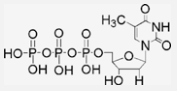
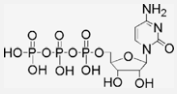
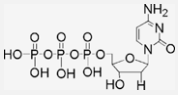
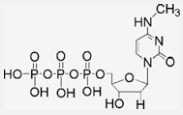
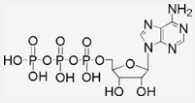
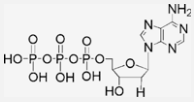
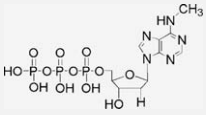
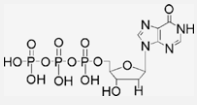
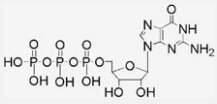
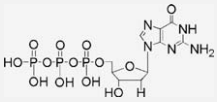
The activity of the full-length GlmU with GlcNAc-1-P, GlcNAcZ-1-P or Glc-1-P as sugar-1-P substrate were investigated and compared with that of GlmU-Tr229. The expression level of the full-length protein (35 mg from 1 L culture) was lower than that of GlmU-Tr229 (55 mg from 1 L culture). Furthermore, GlmU-Tr229 and the full-length GlmU showed comparable activity towards all three sugar-1-Ps (Table 3). Thus the truncation of the C-terminal domain and the linker increases its expression, but does not affect its pyrophosphorylase activity.

In addition, the product inhibition effect of the reaction byproduct inorganic pyrophosphate (IPP) was studied. Pyrophosphatase which catalyzes the hydrolysis of IPP to two orthophosphates was omitted to the reaction mixture. UDP-GlcNAc conversion ratio decreased to 65% in contrast with 95% at maximum when pyrophosphatase existed in the reaction. The result showed inhibition of IPP to the reaction and hydrolysis of IPP by pyrophosphatase could facilitate the uridyltransfer reaction.

To demonstrate the application of the recombinant GlmU-Tr229 for the synthesis of UDP-GlcNAc analogs, we performed a synthetic

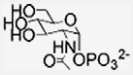
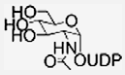
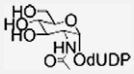
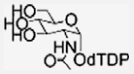
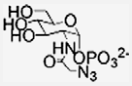
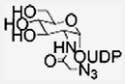
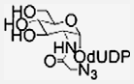
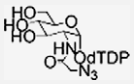
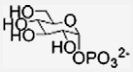
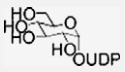
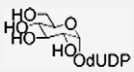
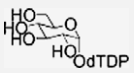
**Table 1**
Yields for enzymatic NDP-GlcNAc reaction




| Entry | Substrate structure | Yield[a] (%) |
|---|---|---|
| 1 |  | 90.9 |
| 2 |  | 72.4 |
| 3 |  | 65.8 |
| 4 |  | 13.9 |
| 5 |  | 8.9 |
| 6 |  | NP[b] |
| 7 |  | 8.5[c] |
| 8 |  | 5.7[c] |
| 9 |  | NP |
| 10 |  | 1.7[c] |
| 11 |  | NP |
| 12 |  | NP |

[a] Yield from profiles of capillary electrophoresis.
[b] No product detected by CE or MS.
[c] Yield of a prolonged reaction.

**Table 2**
Yields for the reactions of GlmU pyrophosphorylase domain with different sugar-1-phosphates.

| Entry | Sugar 1-Phosphate | NTP | Product | Yield[a] (%) |
|---|---|---|---|---|
| 1 |  | UTP |  | 90.9 |
| 2 | | dUTP |  | 72.4 |
| 3 | | dTTP |  | 65.8 |
| 4 |  | UTP |  | 77.7 |
| 5 | | dUTP |  | 66.8 |
| 6 | | dTTP |  | 56.5 |
| 7 |  | UTP |  | 35.3 |
| 8 | | dUTP |  | 24.2 |
| 9 | | dTTP |  | 15.3 |

Entries 1–3, using GlcNAc-1-P as sugar-1-P substrate; entries 4–6, using GlcNAcZ-1-P as sugar-1-P substrate; entries 7–9, using Glc-1-P as sugar-1-P substrate.
[a] Yield from profiles of capillary electrophoresis.

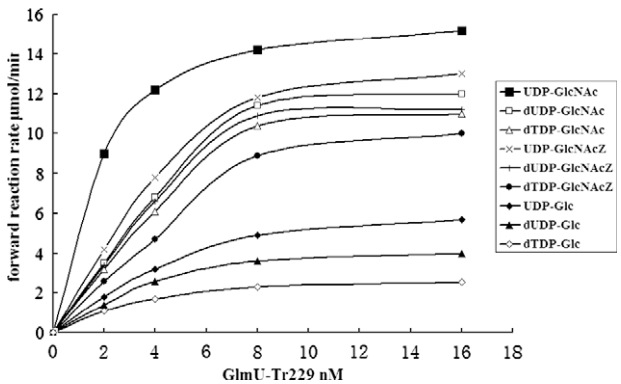
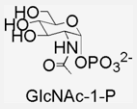
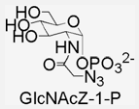
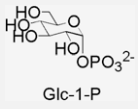



**Figure 2.** Kinetic characterization of GlmU-Tr229 at different enzyme concentration in the uridyltransfer reactions.

**Table 3**
Conversion ratios of different sugar-1-phosphates with UTP in the reactions of full-length GlmU and GlmU pyrophosphorylase domain

| |  GlcNAc-1-P |  GlcNAcZ-1-P |  Glc-1-P |
|---|---|---|---|
| GlmU[a] | 88.3% | 74.2% | 34% |
| GlmU-Tr229[a] | 90.9% | 77.7% | 35.3% |

[a] Enzyme concentration ≈ 1.5 mg/100 μL reaction system.

reaction in multiple mg scale for dUDP-GlcNAc and UDP-GlcNAcZ. The reaction system contained 5 mM sugar-1-P and 5 mM NTP in a final volume of 3 mL. Mono Q ion-exchange column (GE Healthcare) was used to isolate the products from the reaction mixture. The column was eluted with a linear gradient of NaCl from 0 to 500 mM and the fractions containing sugar nucleotides were pooled and concentrated. The products were further desalted by P2 gel filtration column (Bio-rad). The isolated dUDP-GlcNAc (5.1 mg, 57.6%) and UDP-GlcNAcZ (4.3 mg, 44.4%) were identified by ESI-MS and NMR spectroscopy.[20]

*N*-Acetylglucosamine-1-phosphate uridyltransferase is regarded as an essential enzyme involved in prokaryote UDP-GlcNAc biosynthetic pathway. Based on the previously reported crystal structure, we cloned a truncated enzyme only with the pyrophosphorylase domain. The truncated GlmU showed substrate tolerance for both sugar-1-P and NTP, providing a feasible approach for the synthesis of UDP-GlcNAc analogs. With this enzyme, two UDP-GlcNAc derivatives, dUDP-GlcNAc and UDP-GlcNAcZ, were in vitro synthesized. The former could be used to investigate the NDP-sugar substrate specificity of glycosyltransferases, while the latter could introduce an azide group into glycan structures, either to facilitate glycosylation process detection or to be an intermediate for the synthesis of more complex glycan analogs with an active group. Further broadening of the activity of this enzyme by enzyme engineering based on its structure is ongoing in our lab.

## Supplementary data

Supplementary data contain the protein expression and purification, CE profiles and ESI-MS data for substrates and products, and NMR spectra for dUDP-GlcNAc and UDP-GlcNAcZ. Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bmcl.2009.09.039.

## References and notes


1. Nakagawa, J.; Tamaki, S.; Tomioka, S.; Matsuhashi, M. *J. Biol. Chem.* **1984**, *259*, 13937.
2. Williamson, J. M.; Anderson, M. S.; Raetz, C. R. *J. Bacteriol.* **1991**, *173*, 3591.
3. Kuhn, H. M.; Meier-Dieter, U.; Mayer, H. *FEMS Microbiol. Rev.* **1988**, *4*, 195.
4. Chang, R.; Moquist, P.; Finney, N. S. *Carbohydr. Res.* **2004**, *339*, 1531.
5. Herscovics, A.; Orlean, P. *FASEB J.* **1993**, *7*, 540.
6. Namboori, S. C.; Graham, D. E. *J. Bacteriol.* **2008**, *190*, 2987.
7. Levy, P.; Picard, J. *Eur. J. Biochem.* **1976**, *61*, 613.
8. Feng, F.; Okuyama, K.; Niikura, K.; Ohta, T.; Sadamoto, R.; Monde, K.; Noguchi, T.; Nishimura, S. *Org. Biomol. Chem.* **2004**, *2*, 1617.
9. Shao, J.; Zhang, J.; Nahálka, J.; Wang, P. G. *Chem. Commun.* **2002**, 2586.
10. Milewski, S.; Gabriel, I.; Olchowy, J. *Yeast* **2006**, *23*, 1.
11. Mio, T.; Yabe, T.; Arisawa, M.; Yamada-Okabe, H. *J. Biol. Chem.* **1998**, *273*, 14392.
12. Mengin-Lecreulx, D.; van Heijenoort, J. *J. Bacteriol.* **1994**, *176*, 5788.
13. Bulter, T.; Elling, L. *Glycoconjugate J.* **1999**, *16*, 147.
14. Olsen, L. R.; Roderick, S. L. *Biochemistry* **2001**, *40*, 1913.
15. Mengin-Lecreulx, D.; van Heijenoort, J. *J. Bacteriol.* **1993**, *175*, 6150.
16. Parisi, G.; Echave, J. *BMC Evol. Biol.* **2004**, *4*, 41.
17. Brown, K.; Pompeo, F.; Dixon, S.; Mengin-Lecreulx, D.; Cambillau, C.; Bourne, Y. *EMBO J.* **1999**, *18*, 4096.
18. Gehring, A. M.; Lees, W. J.; Mindiola, D. J.; Walsh, C. T.; Brown, E. D. *Biochemistry* **1996**, *35*, 579.
19. Cai, L.; Guan, W.; Kitaoka, M.; Shen, J.; Xia, C.; Chen, W.; Wang, P. G. *Chem. Commun.* **2009**, 2944.
20. *Spectroscopic data of dUDP-GlcNAc*: $^1$H NMR (600 MHz, $D_2O$): $\delta$ 7.79 (d, 1H, $J$ = 7.8 Hz, H-6″), 6.16 (t, 1H, $J$ = 6.6 Hz, H-1′), 5.78 (d, 1H, $J$ = 7.8 Hz, H-5″), 5.34 (dd, 1H, $J$ = 3.6, 7.2 Hz, H-1), 4.44 (m, 1H, H-3′), 3.98–4.06 (m, 3H, H-4′, H-5′a,b), 3.83 (dt, 1H, $J$ = 3.6, 10.2 Hz, H-2), 3.76 (m, 1H, H-5), 3.71 (dd, 1H, $J$ = 2.4, 12.6 Hz, H-6a), 3.65 (t, 1H, $J$ = 10.2 Hz, H-3), 3.64 (dd, 1H, $J$ = 4.2, 12.6 Hz, H-6b), 3.39 (t, 1H, $J$ = 10.2 Hz, H-4), 2.20–2.25 (m, 2H, H-2′a,b), 1.96 (s, 3H, Ac); $^{13}$C NMR (150 MHz, $D_2O$): $\delta$ 174.6, 166.2, 151.5, 141.8 (C-6″), 102.3 (C-5″), 94.4 (C-1), 85.4 (C-1′), 85.3 (C-4′), 72.9 (C-5), 70.8 (C-3), 70.7 (C-3′), 69.3 (C-4), 65.3 (C-5′), 60.1 (C-6), 53.5 (d, C-2, $J$ = 8.2 Hz), 38.7 (C-2′), 21.9 ($CH_3CO$). ESIMS (negative ion): Calcd for $C_{17}H_{27}N_3O_{16}P_2$: 591.09 [M]; Found: 590.2 $[M-H]^-$. *Spectroscopic data of UDP-GlcNAcZ*: $^1$H NMR (600 MHz, $D_2O$): $\delta$ 7.78 (d, 1H, $J$ = 8.4 Hz, H-6″), 5.81 (d, 1H, $J$ = 5.4 Hz, H-1′), 5.79 (d, 1H, $J$ = 8.4 Hz, H-5″), 5.36 (dd, 1H, $J$ = 3.0, 7.2 Hz, H-1), 4.17–4.22 (m, 2H, H-2′, H-3′), 4.11 (m, 1H, H-4′), 4.07 (m, 1H, H-5a′), 4.01 (m, 1H, H-5b′), 3.97, 3.90 (2d, 2H, $J$ = 16.2 Hz, $COCH_2N_3$), 3.89 (dt, 1H, $J$ = 3.0, 10.2 Hz, H-2), 3.76 (m, 1H, H-5), 3.70 (dd, 1H, $J$ = 2.4, 12.6 Hz, H-6a), 3.65 (t, 1H, $J$ = 10.2 Hz, H-3), 3.64 (dd, 1H, $J$ = 4.2, 12.6 Hz, H-6b), 3.39 (t, 1H, $J$ = 10.2 Hz, H-4); $^{13}$C NMR (150 MHz, $D_2O$): $\delta$ 170.8, 166.2, 151.7, 141.5 (C-6″), 102.5 (C-5″), 94.2 (d, C-1, $J$ = 6.0 Hz), 88.4 (C-1′), 83.0 (d, C-4′, $J$ = 9.0 Hz), 73.6 (C-3′), 72.9 (C-5), 70.7 (C-3), 69.5 (C-2′), 69.3 (C-4), 64.9 (d, C-5′, $J$ = 6.0 Hz), 60.1 (C-6), 53.6 (d, C-2, $J$ = 9.0 Hz), 51.4 ($COCH_2N_3$). ESIMS (negative ion): Calcd for $C_{17}H_{26}N_6O_{17}P_2$: 648.08 [M]; Found: 647.6 $[M-H]^-$, 669.5 $[M-2H+Na]^-$, 323.5 $[M-2H]^{2-}$.