Carbohydrate Research 346 (2011) 2421–2425

Contents lists available at SciVerse ScienceDirect

Carbohydrate Research

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

Wide sugar substrate specificity of galactokinase from *Streptococcus pneumoniae* TIGR4

Min Chen [a,b], Lei-lei Chen [a,b,c,†], Yang Zou [a,b], Mengyang Xue [a,b], Min Liang [a,b], Lan Jin [a,b], Wan-yi Guan [a,b], Jie Shen [d], Wenjun Wang [d], Lei Wang [a,b], Jun Liu [a,b], Peng George Wang [a,b,e,f,*]

[a] *The State Key laboratory of Microbial Technology, Shandong University, 27# Shanda South Road, Jinan, Shandong 250100, China*
[b] *National Glycoengineering Research Center, Shandong University, 27# Shanda South Road, Jinan, Shandong 250100, China*
[c] *Institute of Agro-Food Science & Technology, Shandong Academy of Agricultural Sciences, Jinan 250100, China*
[d] *College of Pharmacy, Nankai University, Tianjin 300071, China*
[e] *Department of Chemistry, The Ohio State University, Columbus, OH 43202, USA*
[f] *Department of Biochemistry, The Ohio State University, Columbus, OH 43202, USA*

ARTICLE INFO

Article history:
Received 14 June 2011
Received in revised form 27 July 2011
Accepted 12 August 2011
Available online 22 August 2011

Keywords:
Galactokinase
Substrate specificity
Streptococcus pneumoniae

ABSTRACT

Galactokinases (GALK) have attracted significant research attention for their potential application in the enzymatic synthesis of unique sugar phosphates. The galactokinase (GalKSpe4) cloned from *Streptococcus pneumoniae* TIGR4 had a temperature optimum of 45 °C, and a pH optimum of 8.0. The substrate specificity and kinetics studies revealed that GalKSpe4 had moderate activity toward glucose, in contrast with very low or no activity observed in other previously reported GALKs. Most interestingly, GalKSpe4 exhibited activity for GalNAc, which had never been recorded in other GALKs found by now. This is the first time to report that bacterial GALK can recognize GalNAc.

1. Introduction

D-Galactose (Gal), D-glucose (Glc), *N*-acetyl-D-glucosamine (GlcNAc) and *N*-acetyl-D-galactosamine (GalNAc) are prevalent sugar residues in many glycosylated natural products[1,2] and glycoconjugates.[3–7] They also participate in many biological processes. In nature, uridine 5′-diphospho-sugars (UDP-sugars) are used as sugar donor to incorporate the above-mentioned sugars into target molecules. These donors can be obtained from enzymatic pyrophosphorylation of sugar-1-phosphate (sugar-1-P). The enzymatic route that directly phosphorylated sugars to sugar-1-P offers a more promising synthetic method. However, previously reported kinases that could convert a sugar directly into the corresponding sugar-1-P were limited to fucokinases, galactokinases, and *N*-acetylgalactosamine kinases.[8]

The enzyme galactokinase (EC 2.7.1.6, ATP: D-galactose-1-phosphototransferase, GALK) is crucial in the conversion of exogenous galactose (Gal) to UDP-galactose through the Leloir pathway (Fig. 1).[9] It catalyzes Mg ATP-dependent phosphorylation of the C-1 hydroxyl group of α-D-Gal to yield galactose-1-phosphate (Gal-1-P).[10] Defection of GALK in humans resulted in a disease referred to as Type II galactosemia.[11] GALK is also the key enzyme in the in vitro enzymatic glycorandomization, one of the most promising routes to alter glycosylation.[12] Various GALKs from human,[13] pig,[14] insect,[15] plant seeds,[16] yeast[17,18] and bacteria[12,19–22] had been characterized, and they displayed different biochemical characters.

The substrate specificities of galactokinases of diverse sources have attracted significant research attention, due in part to their potential application in enzymatic preparation of unique sugar phosphates. Substrate specificity studies revealed that D-Gal, 2-deoxy-Gal, 3-deoxy-Gal, 2-amino-2-deoxy-Gal, D-fucose, ATP, 2′-dATP, and 3′-dATP were possible substrates of GALK. For

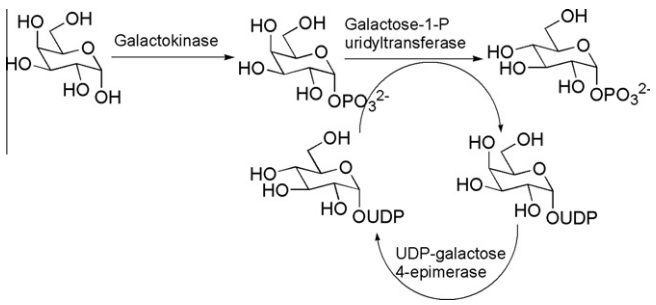

Figure 1. The Leloir pathway.

* Corresponding author. Tel.: +86 531 88366078.
E-mail address: pwang@sdu.edu.cn (P.G. Wang).
† Co-first author, Lei-lei Chen and Min Chen had equal contribution to this work.

0008-6215/$ - see front matter © 2011 Published by Elsevier Ltd.
doi:10.1016/j.carres.2011.08.014

example, in addition to Gal, four sugars (2-deoxy-Gal, 3-deoxy-Gal, 2-amino-2-deoxy-Gal and D-fucose) were proved to be substrates for GALK from *Escherichia coli*, while D-Gal and 2-deoxy-Gal were recognized by human GALKs.[10] However, GALK enzymes reported previously could hardly tolerate modifications at C-4 position.[10] Natural anomeric glucokinase capable of D-glucose anomeric phosphorylation was only reported in *L. lactis* GALK with weak activity (k_{cat}/K_m value is 0.7 mM^{-1} min^{-1}).[2] GalNAc could not be utilized by GALK enzymes in previous reports.[2]

Streptococcus pneumoniae (the pneumococcus) is a major human pathogen, which causes more than 1 million deaths each year worldwide.[23] Despite the medical manifestation of the pneumococcus, no reports regarding the GALK in this pathogen were available. Putative *galE*, *galT*, and *galK* genes involved in the Leloir pathway (Fig. 1) are present in the *S. pneumoniae* TIGR4 genome (GenBank accession No. AE005672), and accordingly, *S. pneumoniae* was predicted to be able to utilize exogenous Gal to synthesize Gal-1-P, yet to be experimentally validated. Characterization of the pneumococcal GALK would be of great medical importance. Here, the heterologous expression and biochemical characterization of the enzyme GALK from *S. pneumoniae* TIGR4 were described.

2. Materials and methods

2.1. Bacterial strains and materials

E. coli DH5α (*lacZDM15 hsdR recA*) for plasmid maintenance was purchased from Gibco-BRL (Gaithersburg, MD), and *E. coli* BL21 (DE3) for protein expression was from Novagen (Carlsbad, CA). The genomic DNA of *S. pneumoniae* TIGR4 was kindly provided by Dr. Samantha King (Nationwide Children's Hospital™, Columbus, USA). Plasmid pMCSG7 was obtained from Novagen (Madison, WI). PCR reagents and restriction enzymes were from Invitrogen (Carlsbad, CA). The plasmid extraction kit and gel purification kit were purchased from Qiagen (Valencia, CA). The HisTrap affinity column (5 mL) and HiLoad-16/60-superdex 200 column were from Amersham Pharmacia Biotech (Piscataway, NJ). All kits and enzymes were used following the manufacturer's instruction.

2.2. Cloning of *galK* gene from *S. pneumoniae* TIGR4

The DNA sequence of *galK* gene (GenBank accession No. AAK75925) was extracted from the genome of *S. pneumoniae* TIGR4 (GenBank accession No. AE005672), and a pair of specific primers GalKS (5′-*TACTTCCAATCCA*ATGCGATGGCACAACATCTTACT-3′) and GalKA (5′-*TTATCCACTTCCAATGC*TAGTCAAGGACGCGAG-3′) was accordingly designed. An 1100-bp DNA fragment was amplified by polymerase chain reaction (PCR), and then verified by DNA sequencing. The corrected PCR fragment was cloned into the plasmid pMCSG7. The recombinant plasmid pMCSG7-GALK was subsequently transformed into *E. coli* DH5α cells. Selected clones were characterized by restriction mapping and DNA sequencing. The positive constructs were subsequently transformed into *E. coli* BL21 (DE3) for protein expression.

2.3. Heterologous expression and enzymatic purification of GalKSpe4

E. coli BL21 (DE3) strain harboring the recombinant plasmids was grown in 1 L of LB medium with shaking at 37 °C, 220 rpm. When OD_{600} reached 0.6–0.8, isopropyl-1-thio-β-D-galactospyranoside (IPTG) was added to a final concentration of 0.2 mM for induction. After expression proceeded overnight at 16 °C, cells were harvested by centrifugation at 4 °C, and the target proteins were detected by SDS–PAGE analysis with Coomassie Blue staining. Then cells were washed, resuspended in the binding buffer (500 mM NaCl, 20 mM Tris–HCl, pH 8.0, 20 mM imidazole), and lysed by sonication for 10 min at 400 W. The supernatant was collected by centrifugation (12,000 rpm, 30 min, 4 °C), and loaded onto a 5-mL HisTrap affinity column which was pre-equilibrated with binding buffer. The target proteins were eluted with elution buffer (500 mM NaCl, 20 mM Tris–HCl, pH 8.0, 500 mM imidazole), and fractions containing the purified enzymes were collected and dialyzed against Tris–HCl buffer (20 mM, pH 8.0). The purity and molecular weight of the purified enzyme was analyzed by 12% SDS–PAGE as described by Laemmli.[24] Protein concentration was determined by the Bradford method.[25] The purified enzyme was named as GalKSpe4 and stored at −20 °C in 20% glycerol for further analysis.

2.4. Enzymatic assay and kinetics characterization

The activity of protein GalKSpe4 toward Gal was assayed as described before[11] with a little modification. Briefly, the reaction mixture containing Gal (8 mM), ATP (10 mM), $MgCl_2$ (5 mM), and GALK (6 μM) in Tris–HCl buffer (20 mM, pH 8.0) was incubated at 45 °C for 180 min. The reducing sugar consumed in this reaction was quantified by the DNS method with D-Gal as standard.[26] The effects of temperature (25–50 °C) and pH (3.0–11.0) on the activity of GalKSpe4 were determined using Gal as the substrate. Kinetic data were obtained by determining the slope of linear phase of the progress curve over 2 min in 30 s intervals and then by calculating the kinetic parameters as described by Yang et al.[12]

2.5. Substrate specificity of the enzyme GalKSpe4

The activity of recombinant GalKSpe4 toward Gal, Glc, *N*-acetylgalactosamine (GalNAc) and *N*-acetylglucosamine (GlcNAc) was assayed in a cocktail containing 50 mM Tris–HCl (pH 8.0), 5 mM $MgCl_2$, 5 mM ATP and proper amount of GalKSpe4 at 45 °C for 3 h. The reaction was terminated by boiling the mixture for 5 min followed by centrifugation. The supernatant was analyzed by normal phase silica gel thin-layer chromatography with the phase being *n*-butanol/acetic acid/water (2:1:1, v/v/v).

2.6. Purification of the phosphorylated products catalyzed by GalKSpe4

The reaction mixture (10 mL), respectively, contained 40 mM Gal, Glc, GalNAc or GlcNAc, 50 mM ATP, 5 mM $MgCl_2$, and 1.5 mg/mL GalKSpe4 in 100 mM Tris–HCl buffer (pH 8.0). After incubation at 45 °C for 3 h, the mixture was briefly boiled for 5 min and then centrifuged to remove protein precipitation. The supernatant was concentrated under reduced pressure and the residue was purified by normal phase silica gel column chromatography using gradient CH_2Cl_2/5 mM NH_4HCO_3 in methanol as elution buffer (CH_2Cl_2/5 mM NH_4HCO_3 in methanol 1:1 to 0:1). The fractions containing products were collected and concentrated under reduced pressure.

2.7. Identification of the phosphorylated products catalyzed by GalKSpe4

Analytical TLC was carried out on Silica Gel 60 F254 aluminum-backed plates (E. Merck). The 200–400 mesh size of the same absorbent was utilized for all chromatographic purifications. All compounds isolated were sufficiently identified by mass spectra (MS) and Proton nuclear magnetic resonance (^{1}H NMR) analysis. MS were recorded on a Bruker MicroTOF spectrometer. ^{1}H NMR spectra were recorded on a Bruker DPX 400 spectrometer at 400 MHz.

3. Results

3.1. Sequence analysis of galactokinase from *S. pneumoniae* TIGR4

The gene of galactokinase cloned from the genome of *S. pneumoniae* TIGR4 has a size of 1179 bp. The complete amino acid sequence of GalKSpe4 showed 53.9% identity with the galactokinase from *L. lactis* (PDB ID: 1PIE) and 37.82% identity with the galactokinase from *E. coli* str. K-12 substr. MG1655 (GenBank ID: 16128725). Sequence analysis showed a high similarity (51.93%) between these three galactokinases. GALK belongs to GHMP superfamily. All three glycine-rich functional motifs conserved in GHMP superfamily members were identified in GalKSpe4 (Fig. 2).

3.2. Heterologous expression and purification of the enzyme GalKSpe4

To identify the enzyme GalKSpe4, its encoding gene was cloned into plasmid pMCSG7. The recombinant enzymes were expressed in *E. coli* and purified by nickel affinity chromatography (Fig. 3A). Fractions containing the purified enzymes were collected and dialyzed against Tris–HCl buffer (20 mM, pH 8.0). GalKSpe4 was shown to be purified to homogeneity with an apparent molecular mass of about 37 kDa, which was consistent with the molecular mass calculated from its amino acid sequence (Fig. 3B).

3.3. Enzymatic assay

The effects of temperature and pH value on the activity of GalKSpe4 were investigated. It was found that GalKSpe4 had an optimal activity at about 45 °C (Fig. 4A), and displayed moderate activity at 25 °C. However, the enzyme lost activity sharply when the temperature reached above 45 °C, and became nearly inactive at 55 °C. As shown in Figure 4B, the activity assay of GalKSpe4 under different pH value demonstrated that the enzyme catalyzed the Gal phosphorylating reaction under wide pH conditions (pH 3–11) with its optimal activity at about pH 8.0.

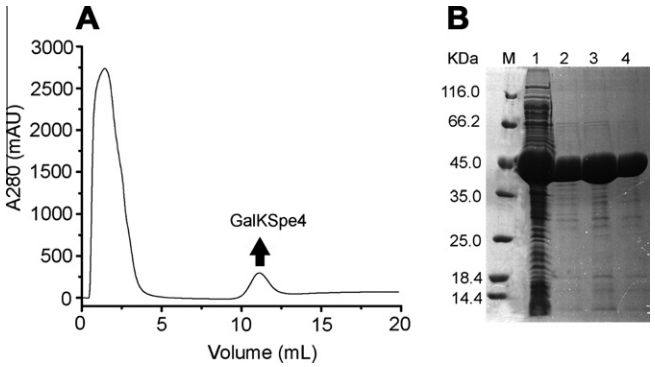

Figure 3. Purification and SDS–PAGE analysis of the purity of GalKSpe4 eluted from affinity column. (A) The elution profile of Ni-affinity column, the peak of the enzyme GalKSpe4 was marked with an arrow. (B) The proteins were stained with Coomassie blue R250. M, protein standards; lane 1, the whole recombinant *E. coli* cell extracts expressing GalKSpe4; lane 2, samples eluted by the washing buffer; lane 3–4, samples eluted by the elution buffer.

3.4. Substrate specificity and kinetics of GalKSpe4

To test the substrate specificity of GalKSpe4, the free sugar/sugar-1-phosphate ratio was assessed by both TLC and DNS assay. As

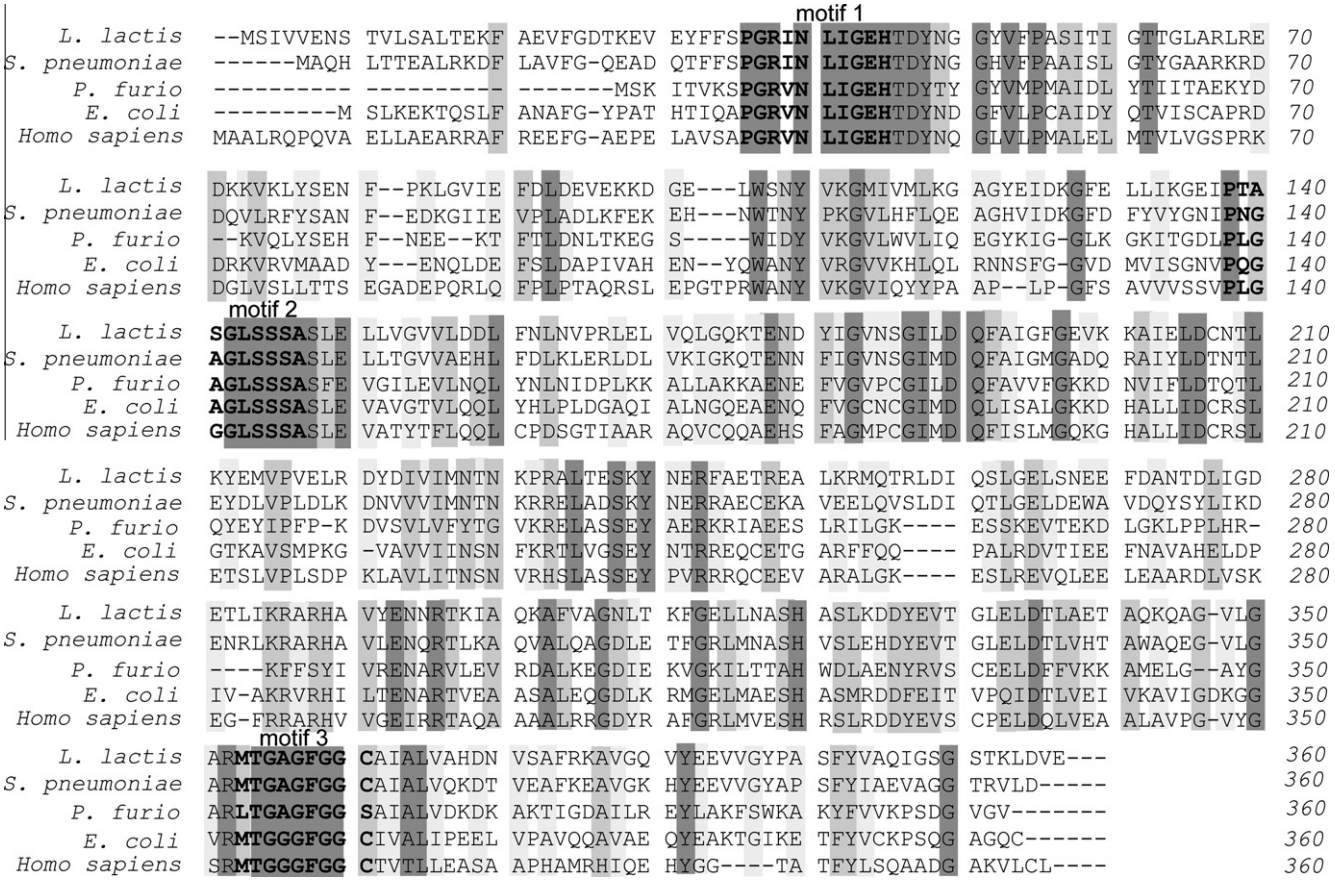

Figure 2. Sequence alignment of galactokinases of different sources.

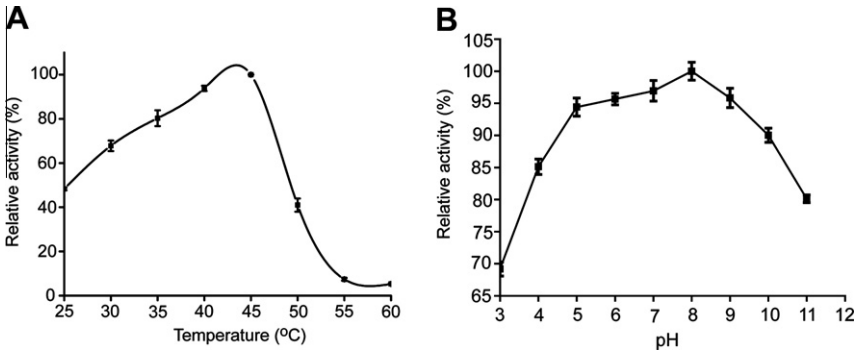

Figure 4. The effect of temperature (A) and pH (B) on the activity of GalKSpe4. To determine the optimal temperature, the galactokinase activity of GalKSpe4 was assayed at pH 8.0 in a range of 25–60 °C. To determine the optimal pH, the galactokinase activity of GalKSpe4 was assayed at 45 °C in a range of pH 3.0–11.0.

Table 1
Kinetic parameters of purified galactokinase GalKSpe4 from *S. pneumoniae* TIGR4 toward different acceptor sugars

Substrate	V_{max} (mmol min^{-1})	K_m (mM)	k_{cat} (min^{-1})	k_{cat}/K_m (mM^{-1} min^{-1})
Gal	0.94	37.23	157.08	4.22
Glc	0.6	49.37	70.28	1.42
GalNAc	0.13	15.76	21.7	1.37

confirmation, the products were isolated from the supernatant by normal phase silica gel column chromatography and further confirmed via LC–MS and NMR (Supplementary data). The substrate specificity studies revealed that GalKSpe4 was highly active with Gal, and interestingly, was also active with Glc and GalNAc, whereas no reaction toward GlcNAc was observed.

The kinetic parameters of GalKSpe4 toward different substrates (Gal, Glc, and GalNAc) were determined (Table 1), the result of which was consistent with that of substrate specificity study. The k_{cat}/K_m value for Gal (4.22 mM^{-1} min^{-1}) was about threefold of that for Glc (1.42 mM^{-1} min^{-1}) and GalNAc (1.37 mM^{-1} min^{-1}).

4. Discussion

D-*Gluco*-configured scaffolds are prevalent in many glycosylated natural products[1,2] and glycoconjugates.[3–7] However, glucokinase which is capable of D-glucose anomeric phosphorylation is still elusive. The first anomeric glucokinase was reported in wild type GALK from *L. lactis* with weak activity[2] (k_{cat}/K_m value is 0.7 mM^{-1} min^{-1}), and its efficiency was improved by direct evolution to k_{cat}/K_m value of 1.43 mM^{-1} min^{-1}. Partly because of their promising application for providing unique sugar phosphates facilely, a lot of research has focused on the substrate specificities of galactokinases from diverse sources. However, similar example was rare. For instance, GALK from *E. coli* can tolerate some C-2, C-3 or C-6 changes, yet neither the wild type nor the *E. coli* GALK mutant, which had relatively relaxed substrate specificity, recognized D-Glc.[2]

In this report, GalKSpe4 recognized Glc with moderate activity. The k_{cat}/K_m value for Gal (4.22 mM^{-1} min^{-1}) is about threefold of that for Glc (1.42 mM^{-1} min^{-1}), whereas the wild type *L. lactis* GALK has a k_{cat}/K_m value for Gal (18.1 mM^{-1} min^{-1}) about 26-fold greater than that for Glc (0.7 mM^{-1} min^{-1}).[2] As such, this is the first reported naturally existing GlcK with moderate activity.

N-Acetylgalactosamine (GalNAc), an important sugar residue, widely exists in living systems and involves in many biological processes. GalNAc kinases that can convert GalNAc to GalNAc-1-P have been discovered in mammals.[27] However, GalNAc kinases in bacteria are limited. The first bacterial GalNAc kinase has been reported as *N*-acetylhexosamine kinase (NahK) from *Bifidobacterium longum*.[28] No GALK found by now could use GalNAc as its substrate. Most interestingly, in this report, GalKSpe4 exhibited activity for GalNAc (k_{cat}/K_m is 1.37 mM^{-1} min^{-1}). This is the first time to report that GALK in bacteria can recognize GalNAc. These results also suggested that enzyme from different sources might display different advantages in enzymatic synthesis.

To better understand the nature of GalKSpe4, the sequence alignment was determined. In amino acid sequence assay, GalKSpe4 exhibited 53.9% identity with the galactokinase from *L. lactis* (PDB accession No. 1PIE), 38% with that from human (gi accession No. 4503895), 37.82% with that from *E. coli* (gi ID: 16128725), 27% with *N*-acetylgalactosamine kinase from human (gi accession No. 4503897) and 27% with Gal1p from *Saccharomyces cerevisiae* S288c (gi ID: 6319494). The three-dimensional structure of galactokinase from *L. lactis* (PDBID: 1PIE) was reported. Amino acids responsible for sugar ligand binding included Arg36, Glu42, Asp45, Asp183 and Tyr233. Both Arg36 and Asp183 were thought to be strictly conserved in the amino acid sequences for galactokinases.[10] Asp45 and Tyr233 are thought to be bound with the critical galactose C-4 axial hydroxyl group.[21] In the sequence alignment assay, the equivalent residues within GalKSpe4 were found as Arg31, Glu37, Asp40, Asp178 and Tyr 228. To further uncover the mechanism for the recognition of both Gal and Glc by GalKSpe4, its crystal structure will be of significant assistance.

5. Conclusion

A novel galactokinase from *Streptococcus pneumoniae* TIGR4 (GalKSpe4) was identified, characterized, and successfully overexpressed in *E. coli* with high activity toward galactose. There have been many galactokinase identified but few of them used Glc as substrate to form Glc-1-P. GalKSpe4 is the second GlcK identified in bacteria, however, the first GlcK with moderate activity. Most importantly, it is the first galactokinase identified in bacteria which could recognize GalNAc as substrate.

Acknowledgments

We were grateful to Feng-cheng Xia for the help in NMR assay. Funding was provided by the National High Technology Research and Development Program ('863'Program) of China (2007AA10Z 404), and National Natural Science Foundation of China (310708 24).

Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.carres.2011.08.014.

References

1. Thorson, J. S.; Vogt, T. Glycosylated Natural Products, 1st ed. In *Carbohydrated Based Drug Discovery*; Wong, C. H., Ed.; Academic press: New Jersey, 2003; Vol. 2, pp 685–711.
2. Yang, J.; Liu, L.; Thorson, J. S. *Chembiochem* **2004**, *5*, 992–996.
3. Schmidt, M. A.; Riley, L. W.; Benz, I. *Trends Microbiol.* **2003**, *11*, 554–561.
4. Upreti, R. K.; Kumar, M.; Shankar, V. *Proteomics* **2003**, *3*, 363–379.
5. Messner, P. *Glycoconjugate J.* **1997**, *14*, 3–11.
6. Raetz, C. R.; Whitfield, C. *Annu. Rev. Biochem.* **2002**, *71*, 635–700.
7. Chatterjee, D. *Curr. Opin. Chem. Biol.* **1997**, *1*, 579–588.
8. Cai, L.; Guan, W. Y.; Kitaoka, M.; Shen, J.; Xia, C. F.; Chen, W. L.; Wang, P. G. *Chem. Commun.* **2009**, *28*, 2944–2946.
9. Fery, P. A. *FASEB J.* **1996**, *10*, 461–470.
10. Holden, H. M.; Thoden, J. B.; Timson, D. J.; Reece, R. J. *Cell. Mol. Life Sci.* **2004**, *61*, 2471–2484.
11. Slepak, T. I.; Tang, M. S.; Slepak, V. Z.; Lai, K. *Mol. Genet. Metab.* **2007**, *92*, 78–87.
12. Yang, J.; Fu, X.; Jia, Q.; Shen, J.; Biggins, J. B.; Jiang, J. J.; Zhao, J. J.; Schmidt, J. J.; Wang, P. G.; Thorson, J. S. *Org. Lett.* **2003**, *5*, 2223–2226.
13. Thoden, J. B.; Timson, D. J.; Reeceet, R. J.; Holden, H. M. *J. Biol. Chem.* **2005**, *280*, 9662–9670.
14. Ballard, F. J. *Biochem. J.* **1966**, *98*, 347–352.
15. Lavine, J. E.; Cantlay, E.; Roberts, C. T., Jr.; Morse, D. E. *Biochim. Biophys. Acta* **1982**, *717*, 76–85.
16. Dey, P. M. *Eur. J. Biochem.* **1983**, *136*, 155–159.
17. Heinrich, M. R. *J. Biol. Chem.* **1964**, *239*, 50–53.
18. Timson, D. J.; Reece, R. J. *Biochimie* **2002**, *84*, 265–272.
19. Gulbinsky, J. S.; Cleland, W. W. *Biochemistry* **1968**, *7*, 566–575.
20. Hoffmeister, D.; Thorson, J. S. *Chembiochem* **2004**, *5*, 989–992.
21. Thoden, J. B.; Holden, H. M. *J. Biol. Chem.* **2003**, *278*, 33305–33311.
22. Hartley, A.; Glynn, S. E.; Barynin, V.; Baker, P. J.; Sedelnikova, S. E.; Verhees, C.; de Geus, D.; van der Oost, J.; Timson, D. J.; Reece, R. J.; Rice, D. W. *J. Mol. Biol.* **2004**, *337*, 387–398.
23. Bernatoniene, J.; Finn, A. *Drugs* **2005**, *65*, 229–255.
24. Laemmli, U. K. *Nature* **1970**, *227*, 680–685.
25. Bradford, M. M. *Anal. Biochem.* **1976**, *72*, 248–254.
26. Miller, G. L. *Anal. Chem.* **1959**, *31*, 426–428.
27. Pastuszak, I.; Drake, R.; Elbein, A. D. *J. Biol. Chem.* **1996**, *271*, 20776–20782.
28. Nishimoto, M.; Kitaoka, M. *Appl. Environ. Microbiol.* **2007**, *73*, 6444–6449.