UDP-glucose pyrophosphorylase synthesizes UDP-glucose from UTP and glucose 1-phosphate and exists in almost all species. Most bacteria possess a GalU-type UDP-glucose pyrophosphorylase, whereas many cyanobacteria species do not. In certain cyanobacteria, UDP-glucose is used as a substrate for synthesis of exopolysaccharide cellulose in spite of the absence of GalU-type UDP-glucose pyrophosphorylase. Therefore, there should be an uncharacterized UDP-glucose pyrophosphorylase in cyanobacteria. Here, we show that all cyanobacteria possess a non-GalU-type bacterial UDP-glucose pyrophosphorylase, i.e., CugP, a novel family in the nucleotide triphosphate transferase superfamily. The expressed recombinant Synechocystis sp. strain PCC 6803 CugP had pyrophosphorylase activity that was highly specific for UTP and glucose 1-phosphate. The fact that the CugP gene cannot be deleted completely in Synechocystis sp. PCC 6803 suggests its central role as the substrate supplier for galactolipid synthesis. Galactolipids are major constituents of the photosynthetic thylakoid membrane and important for photosynthetic activity. Based on phylogenetic analysis, this CugP-type UDP-glucose pyrophosphorylase may have recently been horizontally transferred to certain noncyanobacteria.

glycosylation, which is catalyzed by various types of glycosyl transferases, is important for the biosynthesis of many biological molecules, e.g., polysaccharides, glycoproteins, and glycolipids. Nucleotide diphosphate (NDP)-sugars, which are the substrates for glycosylation, are mostly synthesized by nucleotidyl transferases that use a nucleotidyl transferases to form a nucleotide triphosphate (NTP) and a sugar 1-phosphate (sugar-1P) as substrates. Various NDPs and sugars are combined into NDP-sugars in vivo. Of these, UDP-glucose (UDP-Glc) is used for the biosynthesis of cellulose (1) and glyco- gen (2), galactose metabolism (3), and addition of the glucose moiety in heteroglycans, glycoproteins (4–6), and glycolipids (7). Other NDP-sugars are utilized in the biosynthesis of other polysaccharides and in various glycosylation reactions.

UDP-Glc and the by-product inorganic pyrophosphate (PPi) are produced from UTP and Glc-1P by the UDP-Glc pyrophosphorylases (PPases) (EC 2.7.7.9) in the reversible reaction UTP + Glc-1P ↔ UDP-Glc + PPi. This enzyme is ubiquitous although its typical bacterial form, the GalU-type, is widely distributed in amino acid sequence and three-dimensional structure (12–14). The best-known bacterial UDP-Glc PPase is GalU from Escherichia coli (15, 16). The crystal structures and the reaction mechanism of various GalU-type homologs have been elucidated (17–19). These bacterial GalU-type UDP-Glc PPases belong to a large superfamily of nucleotidyl transferases that also includes UDP-N-acetylglucosamine (GlcNAc) PPase, ADP-Glc PPase, CDP-Glc PPase, dTDP-Glc PPase, GDP-mannose (GDP-Man) PPase, and some yet uncharacterized enzymes. These enzymes probably diverged from a common ancestor in evolution (20).

There are many Glc-containing polysaccharides, such as glyco- gen and Glc-enriched exopolysaccharides, in cyanobacteria (21). Certain cyanobacteria also accumulate cellulose (β-1,4-glucan) (22, 23). Consistently, cyanobacterial genomes have been shown to encode many putative nucleotidyl transferases that might supply NDP-sugars (24, 25). Recently the GalU-type protein All3274 from the filamentous cyanobacterium Anabaena sp. strain PCC 7120 was biochemically revealed to be a UDP-Glc PPase (20). However, such a GalU-type enzyme is not present in all cyanobacteria. Specifically, although Thermosynechococcus is reported to possess a cellulose synthase to produce cellulose under conditions of light and low temperature (23), a GalU-type gene that could supply UDP-Glc to this cellulose has not been found in its genome (24). Given these observations, we hypothesized that an uncharacterized non-GalU-type UDP-Glc PPase may exist in such cyanobacteria. To test this hypothesis, we first focused on a common sequence feature that is shared only by GalU-type UDP-Glc PPases, UDP-GlcNac PPases, and a putative enzyme family that is present in every species of cyanobacteria. Here, we report that the members of this previously uncharacterized family are cyano- bacterial UDP-Glc PPases (thus denoted CugP).

MATERIALS AND METHODS

Sequence alignment and phylogram. The sequences of the cyanobacte- rial proteins categorized as belonging to the bacterial nucleotidyl trans- ferase superfamily (Pfam PF00483, NTP_transferase) (http://pfam.sanger.ac.uk) were obtained from NCBI (http://www.ncbi.nlm.nih.gov/) and Cyanobase (http://genome.microbedb.jp/cyanobase) databases. Sequence alignment was performed by the neighbor-joining method using Clustal X (26).

Cloning of sll1558 and single-amino acid substitution. sll1558 from Synechocystis sp. strain PCC 6803, which encodes the hypothetical UDP- Glc PPase Sll1558 was PCR amplified using PrimeSTAR Max DNA Poly-
merase (TaKaRa Bio, Otsu, Japan) and then directly cloned into the expression vector pET-28a (+) (Merck, Darmstadt, Germany) using In-Fusion HD Cloning kit reagents (TaKaRa) to produce sll1558-pET28a. The primers were pET28a-1F (5′-GGAG TCC GAA TTC GAG CTC CGT C-3′), pET28a-2R (5′-CAT ATG GCT GCC GCG GAC CAC-3′), sll1558-1F28a (5′-GGCC GGC AGC CAT ATG AAA GCC ATG ATT TTG GCCC-3′), and sll1558-2R28a (5′-CTC GAA TTC GGA TCT TTA TTC CGT CTG GAC AAG-3′). To generate a single amino acid substitution within the Sll1558 protein product, the plasmid was amplified by two sets of primers (sll1558-3F/A8G, 5′-GGCC GGC AGC CAT ATG AAG GCC ATG ATT TTG GCCC-3′; sll1558-4R/A8G, 5′-GGCC CTT GCC ACC GGC CAA AAT CAT GCG TTT CAT-3′) and self-combined using the aforementioned cloning reagents. These proteins were His tagged from the pET28a vector. The resulting DNA constructs were verified by DNA sequencing. These proteins were N-terminally His tagged, derived from the pET28a vector.

Expression, purification, and SDS-PAGE analysis of recombinant proteins. E. coli strain C41 (DE3) carrying sll1558-pET28a or sll1558(A8G)-pET28a, which carries the sll1558 gene encoding the amino acid change A to G at residue 8, was cultured in 1 liter of LB medium at 37°C. When each culture reached an optical density at 600 nm (OD600) of 0.4 to 0.8, 1 mM isopropyl β-D-1-thiogalactopyranoside was added (final concentration, 100 μM), and the cells were grown at 18°C for 16 h to achieve induction. The cells from each culture were then centrifuged at 4,220 g for 15 min, suspended in 20 mM HEPES (pH 7.5) containing 100 mM NaCl and 10% (wt/vol) glycerol, and homogenized three times with a French press at 1,500 kg/cm². The soluble proteins were collected by centrifugation at 50,000 g for 30 min. His-tagged Sll1558 and Sll1558 with the A8G substitution (Sll1558(A8G)) were purified by Ni-affinity column chromatography (HisTrap HP; GE Healthcare, Little Chalfont, United Kingdom), with the eluent consisting of a gradient of 30 to 430 mM imidazole in the aforementioned buffer system. Proteins in each fraction were subjected to SDS-PAGE, followed by Coomassie staining. The fraction in each fraction were subjected to SDS-PAGE, followed by Coomassie staining. The fraction

RESULTS AND DISCUSSION

Sequence features of nucleotidyl transferases. The phylogeny of the nucleotidyl transferase superfamily (protein family [Pfam], NTP_transferase) in cyanobacterial species was analyzed based on alignment of full-length amino acid sequences of the transferases (Fig. 1). It is of note that GalU-type UDP-Glc PPase is found in only a few species (Anabaena sp. PCC 7120, Arthrospira platensis NIES-39, and Nostoc punctiforme ATCC 29133). On the other hand, many other enzymes such as dTDP-Glc PPase, CDP-Glc PPase, ADP-Glc PPase, UDP-GlcNac PPase, and as yet uncharacterized nucleotidyl transferases are found in most species of cyanobacteria. We searched for common residues in the GalU-type UDP-Glc PPase and UDP-GlcNac PPase sequences. As reported previously, the G-X-G-T-R motif is highly conserved in all sequences as a part of the (d)NTP-binding site. We found that one residue just prior to the motif is characteristically conserved: Ala is present in GalU-type UDP-Glc PPase, UDP-GlcNac PPase, and one uncharacterized family (provisionally annotated GDP-Man PPase), whereas Gly is present in all of the others (Fig. 2). Since this residue was the only candidate that might be responsible for substrate discrimination, we expressed and characterized the activities of Sll1558, the putative Synechocystis sp. PCC 6803 GDP-Man PPase, and its Ala-8-Gly mutant Sll1558(AG8), which we chose because that cyanobacterium is a well-studied model for cyanobacteria in general.

Purified Sll1558 has a specific UDP-Glc PPase activity. We expressed Sll1558 as a His-tagged protein, which we purified by nickel-affinity chromatography (see Fig. S1 in the supplemental material). According to our SDS-PAGE study, the molecular mass of Sll1558 is that predicted by the translated sequence of sll1558 (45 kDa). The enzyme activity of this product was measured through the liberation of a by-product PPI using a combination of various NTPs (UTP, dTTP, CTP, ATP, and GTP) and sugar-1Ps (Glc-1P, GlcNac-1P, Gal-1P, and Man-1P). We found that Sll1558 showed activity in the presence of UDP and Glc-1P. Notably, it showed practically no activity toward any other substrates or combinations, including Man-1P, even though Sll1558 and related homologs had provisionally been considered to be GDP-Man PPasses, as annotated in Cyanbase. dTTP does not substitute for UTP in reactions with Sll1558, a finding that contrasts with the ability of dTTP to act as a substrate of the GalU-type UDP-Glc PPase from Xanthomonas campestris (8) and Sphingomonas elodea ATCC 31461 (31). Activity and substrate affinity comparison between Sll1558 and Sll1558(AG8). To further examine our prediction about nucleoside discrimination at the Ala residue, we evaluated the possible role of the Ala-8 residue of Sll1558 by site-directed replacement of this residue with Gly. After purification, the substrate...
specificity and enzyme kinetics of Sll1558(A8G) were measured as described above for Sll1558. Sll1558(A8G) also exhibited the PPase activity only for a combination of UTP and Glc-1P, as with wild-type Sll1558. According to Lineweaver-Burk plots (Fig. 3), the \( K_m \) for UTP and \( K_m \) for Glc-1P for Sll1558(A8G) were 0.35 mM and 0.088 mM, and for Sll1558 they were 0.18 mM and 0.085 mM, respectively (Table 1). Thus, the \( K_m \) for UTP was found to be approximately 2-fold higher than that of the wild type, and the \( K_m \) for Glc-1P was comparable to the wild-type level. This finding suggests that the Ala-8 residue plays a slight role in the UTP binding. Notably, the \( V_{max} \) and \( k_{cat} \) values for Sll1558(A8G) were comparable to those of the wild-type enzyme (Table 1). The \( K_m \) values for Sll1558 were similar to those of GalU-type enzymes from bacteria other than cyanobacteria, whereas the \( V_{max} \) and \( k_{cat} \) were about 60- to 200-fold lower than those of these counterpart enzymes (8–10). On the other hand, the kinetics parameters of Sll1558 were found to be similar to those of All3274, the GalU-type UDP-Glc PPase from Anabaena sp. PCC 7120 (20).

Distribution of the GalU-type and CugP-type enzymes, two types of UDP-Glc PPase. Herein, we designated Sll1558 and its homologs as CugP-type UDP-Glc PPase after cyanobacterial UDP-Glc PPase. Because all cyanobacterial genomes examined to date encode a CugP-type UDP-Glc PPase, this enzyme appears to be present in all cyanobacteria (Fig. 1), whereas the GalU-type UDP-Glc PPase is present in only certain species (20). As shown in Fig. 1, a phylogram of all related enzymes in the cyanobacterial nucleotidyl transferase superfamily confirmed such a distribution in cyanobacteria. It must be mentioned also that the cyanobacterial species that possess the GalU-type also possess the CugP-type UDP-Glc PPase.

According to the structure of the phylogram, the GalU-type
and CugP-type enzyme sequences diverged early on from each other or from other related enzymes during the evolution of the nucleotidyl transferase superfamily. It is thus possible that CugP-type UDP-Glc PPases evolved independently within this superfamily. When a BLAST search was performed for the whole NCBI database, the GalU-type PPases were found to be widely distributed throughout the bacterial kingdom, whereas the CugP-type could be found in only very limited species other than cyanobacteria. These limited species included certain, but not all, of the Beta- and Gammaproteobacteria. The sequence similarity among the cyanobacterial and proteobacterial CugP-type enzymes is much greater (E value, 1e^{-11002} to 1e^{-147}) than that of their GalU-type counterparts (E value, 1e^{-28} to 1e^{-32}). Notably, the proteobacterial sequences have clearly diverged from the inside of the cyanobacterial CugP cluster (e.g., Yersinia intermedia, Methylotherma mobilis, and Vibrio ichthyoenteri) (Fig. 1), suggesting that the proteobacterial CugP-type enzymes have been acquired by horizontal gene transfer from cyanobacteria rather recently. Conversely, although the GalU-type enzymes were, in general, widely distributed in Bacteria, they were found in only some filamentous cyanobacteria and may, therefore, have also been acquired through another horizontal transfer from bacteria (e.g., Poribacterium).
teria and Anaerolinea thermophila) to the limited species of cyanobacteria or may have been lost in many cyanobacterial lineages. We speculate that the GalU-type enzyme may supply UDP-Glc to certain glycosyltransferases specific for cell differentiation in those filamentous cyanobacteria.

Structural implications of the motif for substrate specificity. According to the crystal structure and sequence similarities, the overall structure of the whole catalytic domain seems to be conserved between GalU-type UDP-Glc PPase, CDP-Glc PPase, dTDP-Glc PPase, and other subfamilies of the nucleotidyl transferases. Because CugP-type UDP-Glc PPase is also included in this superfamily, we discuss structural implications of the motif for substrate specificities based on the known structures. These enzymes bind (d)NTP and sugar-1P sequentially in the substrate-binding pocket, which can be divided into a (d)NTP-binding site and a sugar-binding site (18, 31). In the former site, three highly conserved residues directly interact with the pyrimidine/purine base of (d)NTP by hydrogen bonds (Fig. 4; see also Fig. S2 in the supplemental material). However, the specificity determination for (d)NTP has not clearly been elucidated. For example, the amide NH of Gly-117 appears to discriminate uracil of UTP in UDP-Glc PPase from cytosine of CTP (see Fig. S2) but is also conserved in other nucleotidyl transferases. The Ala residue that is highlighted in this study as a possible marker for the UTP-utilizing enzymes directly interacts by hydrogen bonding with the C-2 carbonyl of uracil, but very similar hydrogen bonding is found between the Gly-8 and C-2 carbonyl of cytosine in CDP-Glc PPase (see Fig. S2). Thus, the precise mechanism to determine the specificity for (d)NTP has not clearly been elucidated. For example, the amide NH of Gly-117 appears to discriminate uracil of UTP in UDP-Glc PPase from cytosine of CTP (see Fig. S2) but is also conserved in other nucleotidyl transferases. The Ala residue that is highlighted in this study as a possible marker for the UTP-utilizing enzymes directly interacts by hydrogen bonding with the C-2 carbonyl of uracil, but very similar hydrogen bonding is found between the Gly-8 and C-2 carbonyl of cytosine in CDP-Glc PPase (see Fig. S2). Thus, the precise mechanism to determine the specificity for UTP still remains elusive. However, it must be mentioned that the spatial arrangement of the Gly residue (Gly-21/Gly-9/Gly-9 in GalU/CugP/CDP-Glc PPase) next to the mutated Ala position is markedly different, resulting in distinct hydrogen bonding to cytosine in CDP-Glc PPase or to ribose in GalU-type UDP-Glc PPase.

![Fig 3](http://jb.asm.org/)

**FIG 3** Lineweaver-Burk plots of the enzyme activity of Sll1558 (A and B) and A8G-Sll1558 (C and D). As a fixed substrate, 1 mM Glc-1P was included in the experiments for panels A and C, or 2 mM UTP was included in the experiments for panels B and D. Protein concentrations: 4.5 μg ml⁻¹ (A and B), 18.0 μg ml⁻¹ (C), and 9.0 μg ml⁻¹ (D). The error bars indicate standard deviations (for A, B, and C, n = 3; for D, n = 6).

<table>
<thead>
<tr>
<th>Protein and substrate</th>
<th>( V_{\text{max}} ) (U mg⁻¹)</th>
<th>( K_m ) (mM)</th>
<th>( k_{\text{cat}} ) (min⁻¹)</th>
<th>( k_{\text{cat}}/K_m ) (min⁻¹ mM⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sll1558</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UTP</td>
<td>0.63</td>
<td>0.18</td>
<td>28</td>
<td>160</td>
</tr>
<tr>
<td>Glc-1P</td>
<td>0.61</td>
<td>0.085</td>
<td>27</td>
<td>320</td>
</tr>
<tr>
<td><strong>Sll1558(A8G)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UTP</td>
<td>0.64</td>
<td>0.35</td>
<td>29</td>
<td>83</td>
</tr>
<tr>
<td>Glc-1P</td>
<td>0.64</td>
<td>0.088</td>
<td>29</td>
<td>330</td>
</tr>
</tbody>
</table>
(Fig. 4; see also Fig. S2 in the supplemental material). This may be the reason why the $K_m$ of SII1558 (A8G) for UTP increased slightly. Thus, it is still possible that the highlighted Ala residue, which is exclusively conserved in the UTP-utilizing transferases, is somehow involved in discrimination of the uracil of UTP as a substrate. To test this hypothesis, it will be necessary to study the substrate discrimination of the UTP-utilizing transferases by a thorough mutagenesis survey.

**Role of CugP-type UDP-Glc PPase in cyanobacteria.** To better understand the role of cyanobacterial CugP-type UDP-Glc PPases, we attempted to disrupt sII1558 in the *Synechocystis* sp. PCC6803 genome. However, we could not produce such a mutant because cyanobacteria typically contain multiple copies of their chromosomes, and, consequently, the mutated *Synechocystis* sp. PCC6803 did not cleanly segregate, suggesting that the CugP-type UDP-Glc is essential for survival of *Synechocystis* (see Fig. S3 in the supplemental material). In this context, it must be mentioned that a galactolipid, monogalactosyldiacylglycerol (MGDG), which is essential in *Synechocystis*, is supplied from UDP-Glc via mono-galactosyldiacylglycerol (32). The galactolipids MGDG and digalactosyldiacylglycerol are major constituents of the photosynthetic thylakoid membrane in cyanobacteria and in plant chloroplasts. They bind to the membrane-spanning photosynthetic complexes and thereby maintain them in a functional state (33). Hence, the role of UDP-Glc PPases in providing UDP-Glc must be essential for growth and survival in cyanobacteria.

As a preliminary survey, we compared this nonsegregated mutant with the wild type. The chlorophyll (Chl) content was lower in the mutant than in the wild type (mutant, $3.47 \pm 0.51 \mu g$ Chl/10^8 cells; wild type, $4.22 \pm 0.04 \mu g$ Chl/10^8 cells; n = 3). The whole photosynthetic activity from water to CO$_2$ was higher in the mutant than in the wild type (mutant, 263 $\pm$ 40 $\mu$mol O$_2$/mg Chl/h; wild type, 138 $\pm$ 7 $\mu$mol O$_2$/mg Chl/h; n = 4). These facts may reflect modified membrane biogenesis in the partially disrupted mutant (see Fig. S3 in the supplemental material). To further characterize the phenotype, it is essential to achieve the complete segregation of the mutant.

Because galactolipids are specific to photosynthetic membranes, the unique CugP-type enzyme, found mainly in cyanobacteria, may be the result of an evolutionary event in oxygenic photosynthesis. Certain exopolysaccharides secreted by cyanobacteria, including *Synechocystis*, often contain glucose as the major constituent (21). However, it is unknown if they are essential for survival. Cellulose is selectively deposited in a thermophilic cyanobacterium, *Thermosynechococcus*, under conditions of low temperature and light (23). So, a CugP-type Tlr0611 appears to be the sole enzyme to provide UDP-Glc for biosynthesis of cellulose and might be induced under such conditions in this cyanobacterium. Regional distribution of GalU-type PPases in certain cyanobacteria may account for variation in the requirement of exopolysaccharides. Further functional studies of the CugP-type and GalU-type enzymes are needed to understand the role of UDP-Glc in cyanobacteria.

**ACKNOWLEDGMENTS**

This work was supported by Grants-in-Aid for Scientific Research from MEXT and PRESTO from JST (to R.N.) and by Grants-in-Aid for Scientific Research and the GCOE program “From the Earth to Earths” from MEXT and CREST from JST (to M.I.).

**REFERENCES**


