Two strictly polyphosphate-dependent gluco(manno)kinases from diazotrophic *Cyanobacteria* with potential to phosphorylate hexoses from polyphosphates

Tomás Albi • Aurelio Serrano*

Instituto de Bioquímica Vegetal y Fotosíntesis, Centro de Investigaciones Científicas Isla Cartuja, CSIC y Universidad de Sevilla, Av. Américo Vespucio 49, 41092 Sevilla, SPAIN

*To whom correspondence should be addressed**:** Aurelio Serrano, Institute for Plant Biochemistry and Photosynthesis, CSIC and University of Seville - 49 Americo Vespucio Avenue, 41092 Seville, SPAIN. Phone: ++ 34 95 4489524. Fax: ++ 34 95 4460165. E-mail: aurelio@ibvf.csic.es Web: http://www.ibvf.csic.es/en/bioenergetics-phosphate

Keywords: Gluco(manno)kinase · Hexokinase · *Cyanobacteria* · PolyP-dependent ·Polyphosphate · Regeneration of ATP

Abstract

The single-copy genes encoding putative polyphosphate-glucose phosphotransferases (PPGK, EC 2.7.1.63) from two nitrogen-fixing *Cyanobacteria*, *Nostoc* sp. PCC7120 and *Nostoc punctiforme* PCC73102, were cloned and functionally characterized. In contrast to their actinobacterial counterparts, the cyanobacterial PPGKs have shown the ability to phosphorylate glucose using strictly inorganic polyphosphates (polyP) as phosphoryl donors. This has proven to be an economically attractive reagent in contrast to the more costly ATP. Cyanobacterial PPGKs had a higher affinity for medium-long sized polyP (>10 phosphoryl residues). Thus, longer polyP resulted in higher catalytic efficiency. Also in contrast to most their homologs in *Actinobacteria*, both cyanobacterial PPGKs exhibited a modest but significant polyP-mannokinase activity as well. Specific activities were in the range of 180-230 and 2-3 μmol min^{-1} mg^{-1} with glucose and mannose as substrates, respectively. No polyP-fructokinase activity was detected. Cyanobacterial PPGKs required a divalent metal cofactor, and exhibited alkaline pH optima (approx. 9.0) and a remarkable thermostability (optimum temperature, 45 ºC). The preference for Mg^{2+} was noted with an affinity constant of 1.3 mM. Both recombinant PPGKs are homodimers with a subunit molecular mass of ca. 27 kDa. Based on databases searches and experimental data from Southern blots and activity assays, closely-related PPGK homologs appear to be widespread among unicellular and filamentous mostly nitrogen-fixing *Cyanobacteria*. Overall, these findings indicate that polyP may be metabolized in these photosynthetic prokaryotes to yield glucose (or mannose) 6-phosphate. They also provide evidence for a novel group-specific subfamily of strictly polyP-dependent gluco(manno)kinases with ancestral features and high biotechnological potential, capable of efficiently using polyP as an alternative and cheap source of energy-rich phosphate instead of costly ATP. Finally, these results could shed new light on the evolutionary origin of sugar kinases.

INTRODUCTION

Inorganic polyphosphate (polyP) is a linear polyanion composed of tens to hundreds of phosphoryl residues, all of them being linked by "high-energy" phosphoanhydride bonds. Found in many diverse organisms in nature, polyP has proven to be essential for the growth, response to stresses and stringencies of cells (Kulaev 1979; Wood and Clark 1988; Kornberg et al. 1999; Rao et al. 2009).

As phosphorylated compounds with a Gibss free energy of hydrolysis similar to the ATP (-30.5 kJ mol^{-1}), inorganic pyrophosphate (diphosphate, PPi) and polyP have been suggested to be used in place of ATP in diverse biological processes (Lipmann 1965). An enzyme known to hydrolyze polyP rather than ATP is the polyphosphate glucokinase (PPGK, polyphosphate-glucose phosphotransferase, EC 2.7.1.63), which catalyzes the phosphorylation of glucose using polyP as a phosphoryl donor to form glucose 6-phosphate as follows:

$$\text{Glucose} + \text{PolyP}_n \rightarrow \text{glucose 6-P} + \text{PolyP}_{n-1}.$$

PPGK was first observed in *Mycobacterium phlei* (Szymona 1957), and later in other Gram-positive bacteria, all of them belonging to the ancient order of *Actinomycetales* (Szymona 1964; Szymona and Widomski 1974; Szymona and Szymona 1978; Szymona and Szymona 1979; Pepin and Wood 1986; Mukai et al. 2003; Tanaka et al. 2003; Lindner et al. 2010a; Hehuan et al. 2012; Koide et al. 2013). However, no PPGK has been described in other sort of bacteria, archaea, fungi, algae, plants or animals to date.

Most actinobacterial PPGKs have been validated as monomers or homodimers with a molecular mass subunit of ca. 30 kDa. A remarkable feature of these enzymes is its dual substrate specificity: PPGK can use both ATP and polyP as donors to phosphorylate glucose to glucose 6-phosphate. Nevertheless, a PPGK from the polyP-accumulating actinobacterium *Microlunatus phosphovorus*, is the only PPGK enzyme solely dependent on polyP as an energized phosphoryl-substrate donor described to date (Tanaka 2003). Concerning this matter, several studies (Hsieh et al. 1993; Phillips et al. 1999) proved that although both enzymatic activities belong to the same protein, the binding site for this protein differs in each phosphate donor substrate. Thus, the enzymes of most phylogenetically ancient species of the *Actinomycetales* order seem to prefer polyP instead of ATP. As a result of this, a higher polyP-glucokinase/ATP-glucokinase ratio is exhibited. Compared to the rest of glucokinases, PPGKs displayed a wider range of NTP as phosphoryl donors (GTP, UTP, TTP, XTP, CTP and dATP), whereas ATP-glucokinases from more evolved organisms are unable to use polyP, and consequently only poorly replace GTP for ATP, as is the case of hexokinases (EC 2.7.1.11) from fungi and mammals, which are exclusively dependent on ATP (Rao 2009).

PPGK belong to the ROK (Repressor ORF Kinase) superfamily (Pfam PF00480) (Finn et al. 2014), a large group of mostly bacterial proteins which also include other sugar kinases and transcriptional repressors, the latter with an extra h-α-h DNA binding domain. Owing to this fact, kinase enzymes within this group (bacterial gluco-, fructo- and manno-kinases, eukaryotic hexokinases and ADP-glucokinases) reveal a significant grade of structural relationship.

PolyP should play important roles in the overcoming of nutrient and heavy-metal stresses by *Cyanobacteria*, a group of Gram-negative oxygenic photoautotrophic prokaryotes which are among the most successful and oldest forms of life (Schopf 2002) and have gained a lot of attention in recent years because of their potential applications in biotechnology (Abed et al. 2009). Accumulation of polyP granules has been described under various culture conditions in both unicellular (Lawry and Jensen 1979) and filamentous (Jensen et al. 1982) *Cyanobacteria*. Moreover, induction of genes involved in polyP metabolism by Pi starvation was reported in the unicellular strain *Synechocystis* sp. (Gómez-García et al. 2003). Interestingly, in the diazotrophic filamentous cyanobacterium *Anabaena flos-aquae* P is stored in different ways depending of the nitrogen source used. Under dinitrogen fixing conditions P is stored as sugar P, whereas with nitrate as the combined N source it is stored as polyP (Thompson et al. 1994). However, the functional relationships between polyP metabolism and dinitrogen fixation in cyanobacteria have not yet been elucidated.

Here, we report the first polyP-gluco(manno)kinases isolated and biochemically characterized from *Cyanobacteria*. The single copy *ppgK* genes of the filamentous nitrogen-fixing strains *Nostoc* sp. PCC7120 and *Nostoc punctiforme* PCC73102 were cloned and overexpressed in *E. coli*, and the corresponding recombinant proteins, hereafter referred as NsPPGK and NpPPGK respectively, were purified and characterized. As shown, these enzymes are smaller proteins and exhibit some novel biochemical features compared to the previously described PPGKs. Additionally, a survey of homologous closely-related PPGKs has been carried out in a wide range of diazotrophic *Cyanobacteria* by several techniques including Southern blots, activity assays, and bioinformatic analyses. Lastly, this study could also offer new evidence towards the matter of hexokinases evolution. Overall, the obtained results provide indications for cyanobacterial PPGKs representing a taxonomic group-specific new subfamily of strictly polyP-dependent gluco(manno)kinases with high biotechnological potential.

MATERIALS AND METHODS

Reagents and PolyP preparation

Restriction endonucleases and T4 DNA ligase were purchased from Takara Bio Inc (Shiga, Japan). ACCUZYME™ Proofreading DNA Polymerase and the gel extraction kit were obtained from Bioline Inc. (MA, USA). Primers were synthesized by Integrated DNA Technologies (Leuven, Belgium). Sodium polyphosphates PPi, tripolyphosphate (P_3), cyclic P_3 (trimetaphosphate, P_{3c}), tetrapolyphosphate (P_4), a polyphosphate mix with an average chain length of 13-18 phosphoryl residues (P_{13-18}) and water-insoluble Maddrell salt (a mixture of crystalline long-chain polyphosphates of very high molecular mass), NTPs (nucleoside 5'-triphosphates), dATP and hexoses (D-glucose, D-mannose, D-fructose) were purchased from Sigma Chemical Co. (St. Louis, MO, USA). Purchased substrates (polyPs and hexoses) were analytical grade reagents, except the Maddrell salt which was of practical grade. P_{60} and P_{150} (polyphosphate mixes purified by polyacrylamide gel electrophoresis; average chain lengths 60 and 150 phosphoryl residues, respectively) were kindly provided by Dr. Toshikazu Shiba (RegeneTiss Co, Japan). Very long chain polyPs with chain lengths of up to approximately 800 phosphoryl residues (P_{LC}) were obtained by fractionation of solubilized Maddrell salt, prepared as described by Van Wazer (1958) on a 2

% (w/v) polyacrylamide/0.8 % (w/v) agarose gel. When necessary, crystalline polyP was washed twice with 70 % (v/v) ethanol, dried overnight in a vacuum dessicator, and resuspended in distilled water. Otherwise stated, the polyP concentration is expressed in terms of polymer, assuming average chain lengths of: 3, 4, 15, 60, 150 and 300 phosphoryl residues for P_3, P_4, P_{13-18}, P_{60}, P_{150} and P_{LC}, respectively. All other chemicals were of analytical grade.

Analytical polyacrylamide gel electrophoresis of polyP

Polyacrylamide slab gels (total acrylamide, 30 %, w/v; 70 x 85 mm; 1-mm thick) were prepared which contained a 19.2:0.8 ratio of acrylamide to bisacrylamide. The gel was pre-electrophoresed at 100 V for 3 h to remove contaminating ions. The polyP samples were mixed at a ratio of 1:6 with loading buffer [100 mM Tris-borate buffer (pH 8.3), 30 % (v/v) glycerol and 0.25 % (w/v) bromophenol blue]. Gels were run at 50 mA in TBE as electrophoresis buffer. Once electrophoresed, polyPs were fixed and stained with 0.05 % (w/v) Toluidine blue O, 25 % (v/v) methanol and 1 % (v/v) glycerol in water, followed by destaining in an aqueous mixture containing 25 % (v/v) methanol and 5 % (v/v) glycerol. As a result, the polyP stained dark blue against the colorless or lightly-blue background.

Bacterial strains and culture conditions

The cyanobacterial strains used in this work were obtained as axenic cultures from various microbial culture collections of reference for *Cyanobacteria* (ATCC, American Type Culture Collection, Manassas, VA, USA; PCC, Pasteur Culture Collection, Paris, France; UTEX, Culture Collection of Algae, University of Texas, Austin, USA; the CICCartuja Biological Cultures Service, Instituto de Bioquímica Vegetal y Fotosíntesis, Seville, Spain). The strains were photoautotrophically grown in BG11 liquid medium without combined nitrogen source unless otherwise stated (Rippka et al. 1979), and are described in Online Resource Table S1. Cultures (referred to as bubbled cultures) were supplemented with 10 mM $NaHCO_3$, and bubbled with a mixture of CO_2 and air (2 %, v/v), under continuous fluorescent white light (75 µE m^{-2} s^{-1}). The absence of heterotrophic bacterial contamination was assessed by counts on LB (Luria-Bertani) agar plates incubated in the dark.

Preparation of cyanobacterial cell-free extracts

Cyanobacterial cells were harvested by centrifugation and resuspended in 100 mM Tris-HCl (pH 9.0) buffer supplemented with 5 mM $MgCl_2$, 5 mM DTT, 0.1 mM PMSF and a 1:1000 dilution of a Protein Inhibitor Cocktail for use with bacterial cell extracts (P8465; Sigma-Aldrich, USA), at a ratio of 0.2 g (cells wet wt)/ml. Then cells were ultrasonically disrupted at 0-4 ºC. The cell homogenate was centrifuged at 15,000 x *g* at 4 ºC for 20 min, and the resultant clear supernatant (cell-free extract) was used for enzymatic assays.

DNA methodology

Total DNA was isolated by the following procedure: 50 ml of cyanobacterial cultures in the mid-log phase of growth were harvested and resuspended in a final volume of 400 µl in a microcentrifuge tube with 10 mM Tris-HCl (pH 7.5) buffer with 0.1 mM EDTA. Then, 150 µl of sterile glass beads (0.2 µm

diameter), 20 µl of 10 % (w/v) SDS, and 450 µl of phenol-chloroform-isoamyl alcohol mixture (25:24:1 v/v) were added. The mixture was subjected to six cycles of 1-min vigorous vortexing followed by 1-min cooling on ice. The resulting suspension was centrifuged at 15,000 x *g* for 10 min, then the clear supernatant solution was transferred to a new microcentrifuge tube and DNA was finally ethanol precipitated.

Southern blotting

DNA samples isolated from a number of strains representative of the different taxonomic groups of cyanobacteria were digested with appropriate restriction enzymes and loaded onto agarose gels; then Southern analysis was performed (Ausubel et al. 1992) using GeneScreen Plus membranes (Dupont, USA). DNA probes utilized in the hybridizations (full coding *ppgK* fragments) were obtained by PCR, and were then labeled with [α-^{32}P]-dCTP using the Ready-To-Go© DNA labeling kit (GE Healthcare). Nucleic acid hybridization was carried out at 55 °C with gently shaking. Films were exposed for 4 days and developed using a Cyclone© Storage Phosphor System (Packard, USA).

Construction of recombinant plasmids and gene expression in E. coli

The *ppgK* genes from *Nostoc* sp. PCC7120 and *Nostoc punctiforme* PCC73102 were PCR amplified using specific primers (Online Resource Table S2) and genomic DNA as a template. The unique DNA fragments of ca. 0.72 Kb obtained in both cases were initially cloned into the pGemT-Easy vector (Invitrogen) for sequencing. These plasmids were then digested with *Bam*HI and *Pst*I, and the DNA fragments carrying the native open reading frames of *ppgK* genes were eventually ligated into pQE-80L vector (Quiagen, Germany). In this way, a His_6 tag of 12 amino acid residues in total (MRGSHHHHHHGS; nominal mass 1,420 Da) was added to the N-terminal end of the native proteins. *E. coli* BL21(DE3) cells transformed with the appropriate expression plasmid were cultured at 30 °C in 1 L LB liquid medium supplemented with 100 µg ml^{-1} ampicillin with vigorous shaking. When OD_{600} reached ca. 0.6, protein expression was induced by adding 1 mM IPTG and cultures were then incubated overnight at 20 °C with shaking at 200 rpm.

Purification of recombinant cyanobacterial PPGKs by nickel-nitrilotriacetic acid (Ni-NTA) metal-affinity chromatography

Cells were harvested and resuspended in buffer A (500 mM NaCl, 50 mM Na_2HPO_4, 10 mM imidazole, pH 8.0), and then lysed by sonication at 4 °C. Cell debris were removed by centrifugation at 15,000 x *g* for 15 min. The resultant crude extract was loaded onto a pre-equilibrated HisTrap FF Crude Ni-NTA 1-ml column (GE-Healthcare). Subsequently, non-target proteins were removed by washing the column with buffer B (500 mM NaCl, 50 mM Na_2HPO_4, 50 mM imidazole, pH 8.0) until no more protein elution was observed. Finally, recombinant proteins were eluted by applying a linear gradient with a target concentration of 100 % of buffer C (500 mM NaCl, 50 mM Na_2HPO_4, 500 mM imidazole, pH 8.0). The eluted PPGK proteins were dialyzed three times with 50 mM Tris-HCl (pH 9.0) to remove imidazole and phosphate salts, and eventually concentrated by ultrafiltration using Amicon Ultra-3 kDa filters.

FPLC gel filtration chromatography. Estimation of molecular masses

Partially purified His-tagged PPGK preparations, previously concentrated by ultrafiltration, were further purified by FPLC gel filtration chromatography carried out at 4 °C. The concentrated preparations (0.5-1.0 ml volume) were loaded on to a Superdex© 200 PG (GE Healthcare, Sweden) column equilibrated with 150 mM NaCl, 20 mM KCl, 5 mM $MgCl_2$, 50 mM Tris-HCl (pH 9.0) buffer at a flow rate of 2 ml min^{-1} using an ÄKTA-FPLC system (GE Healthcare, Sweden). The molecular masses (M_m) of oligomeric PPGK proteins were determined using the calibration plot derived from the elution volumes of a series of protein standards including: thyroglobulin (Thy, 669 kDa), ferritin (Fer, 443 kDa), β-amylase (β-Amy, 200 kDa), alcohol dehydrogenase (ADH, 150 kDa), bovine serum albumin (BSA, 66 kDa), carbonic anhydrase (CA, 29 kDa) and cytochrome c (Cyt.c, 12.4 kDa). Subunit molecular masses were determined by denaturing discontinuous SDS-PAGE following the method of Laemmli (Laemmli 1970) using 12 % (w/v) separating and 4 % (w/v) stacking polyacrylamide gels. Protein bands were stained with Coomassie Brilliant Blue R-250. Apparent M_m of monomers under denaturing PAGE was calculated using standard proteins. Absolute M_m values of purified recombinant PPGKs were confirmed by MALDI-TOF mass spectrometry (see below). These purified fractions were used for the *in vitro* kinetics assays and biochemical characterization.

Peptide mass fingerprinting and validation of PPGK proteins by MALDI-TOF mass spectrometry

Protein samples corresponding to high-purity cyanobacterial PPGKs were derived from SDS-PAGE. Proteins were digested with trypsin and the resulting peptides were extracted, then loaded onto a suitable MALDI matrix and eventually processed by a MALDI-TOF mass spectrometer (AutoFlex, Bruker-Daltonics, Proteomics Service of the Instituto de Bioquímica Vegetal y Fotosíntesis, CSIC-University of Seville) which generated peptide mass spectra in the mass range 0.8–2.5 kDa. MASCOT-Matrix Science database was used to analyze the peaks lists for protein identification (Koenig et al., 2008).

Determination of enzymatic activities

Unless otherwise stated sugar-kinase enzymatic activities were determined at 40 °C and pH 9.0, using P_{13-18} as a phosphoryl donor substrate. The polyP glucokinase activity was assayed spectrophotometrically by monitoring the production of NADPH at 340 nm using a glucose 6-phosphate dehydrogenase coupled reaction. The assay mixture (1 ml) contained of 100 mM Tris–HCl buffer (pH 9.0), 5 mM $MgCl_2$, 5 mM glucose, 1.11 mM polyP, 5 mM $NADP^+$, and 0.5 U of yeast glucose 6-phosphate dehydrogenase (Sigma Chem. Co., USA). The reaction was started by the addition 0.5-1.5 μg of purified PPGK or 10-20 μl of cell-free extracts. Concentrations of polyphosphate substrates were calculated as polymers, considering mean chain lengths of 15, 60 and 300 phosphate residues for P_{13-18}, P_{60} and P_{LC}, respectively. NTPs were used at 2 mM concentration when assayed as alternative phosphoryl donor substrates instead of polyP. To determine the dependence on pH, 1.0 μg of purified enzyme was incubated as described above in the following buffers at 100 mM concentration: 2-morpholinoethanesulfonic acid (MES) (pH 5.5-7.0), MOPS (pH 7.0-8.0), Tris (pH 8.0-9.0), N-cyclohexyl-2-aminoethanesulfonic acid (CHES) (pH 9.0-10.0) and 3-[cyclohexylamino]-1-propane sulfonic acid (CAPS) (10.0-10.5). When measuring enzymatic activity in cell-free extracts or when the effects of pH, temperature, divalent metal ions, inhibitors and other factors

on glucokinase activity were examined, the assay was discontinuous and $NADP^+$ and glucose 6-phosphate dehydrogenase were omitted from the assay mixture. The reaction was finished by heating the test tube at 95 ºC for 5 min. Then the assay followed as described above by adding 5 mM $NADP^+$ and 0.5 U of glucose 6-phosphate dehydrogenase. The polyP-mannokinase activity was assayed in a similar way, but glucose was replaced by 50 mM mannose and 0.5 U of mannose 6-phosphate isomerase (from *E. coli*; Sigma Chem. Co., USA). Finally, for fructokinase activity determinations, 50 mM fructose and 0.5 U of yeast glucose 6-phosphate isomerase (Sigma Chem. Co., USA) were added in substitution of glucose. Kinetic parameters (K_m and k_{cat}) were determined from initial velocity data that were fitted by the nonlinear regression software Anemona.xlt (Hernández and Ruiz 1998). One unit (U) of PPGK corresponds to 1 µmol of phosphorylated product per minute at 30 °C. Protein concentration was determined by the Bradford method (Bradford 1976) with ovalbumin as a standard.

Computer-aided analysis

Amino acid sequence homology among the PPGK sequences was analyzed online using BLAST searches (Altschul et al. 1990) against the public databases GenBank (Benson et al. 2013), DOE Joint Genome Institute (JGI) (Nordberg et al. 2014) and InterPro (Hunter et al. 2011). The amino acid sequences of putative PPGK orthologs from diverse bacterial strains (Online Resource Table S3) were aligned and phylogenetic trees were constructed with the Evolutionary-distances (Neighbor-joining), Maximum Parsimony, and Maximum Likelihood methods using the SeaView v5.2 software (Gouy et al. 2010).

Nucleotide sequence accession numbers

The nucleotide sequences of the gene constructs reported in this paper have been deposited in the GenBank/EMBL/DDBJ nucleotide sequence databases under accession numbers HG764586 (*ppgK* of *Nostoc* sp. PCC7120) and HG764587 (*ppgK* of *Nostoc punctiforme* PCC73102), respectively.

RESULTS

*all*1371 and Npun_R1878 genes encode functional polyP-dependent glucokinases

BLAST sequence similarity searches in cyanobacterial genomes (Cyanobase, Kazusa DNA Research Institute) (Fujisawa et al. 2014) identified two ORFs, *all*1371 and Npun_R1878 of the diazotrophic filamentous strains *Nostoc* sp. PCC7120 and *Nostoc punctiforme* PCC73102, respectively, with high homology to the *ppgK* gene from *Mycobacterium tuberculosis* $H_{37}Rv$ (Hsieh et al. 1996a). The corresponding predicted proteins, thereafter named NsPPGK and NpPPGK, shared 32 % and 29% identity with their mycobacterial homolog and 91% sequence identity to each other. In addition, each of the genomes of *Nostoc* sp. PCC7120 and *Nostoc punctiforme* PCC73102 possessed a gene encoding a putative glucokinase, *alr*2973 and *Npun*_R5075. They respectively showed 27 % and 14 % sequence identity at the protein level with their corresponding PPGK homolog. Even though both sequences of putative *ppgK* genes were available, Npun_R1878 was wrongly annotated as a transcriptional regulator/sugar kinase (ROK family protein) instead of a PPGK encoding gene. The predicted NsPPGK

and NpPPGK polypeptides have 239 (nominal mass 25,919 Da) and 238 (nominal mass 25,816 Da) amino acid residues, respectively. They are smaller than their actinobacterial homologs (of 260-280 residues) and exhibit in their primary structures the seven regions with structural motifs conserved among the bacterial PolyP/ATP-dependent PPGKs (Mukai et al. 2003), as revealed by protein sequences alignments. Interestingly, when other putative cyanobacterial PPGK sequences were used in the alignment a high level of conservation was found within them, while when cyanobacterial PPGKs are compared to their actinobacterial polyP/ATP-dependent homologs, motifs reported to be involved in phosphoryl-donor and polyphosphate substrate binding (phosphate-1 and -2, connect-1) and the glucose-binding motif are more clearly conserved (Online Resource Fig. S1). Thus, the finding of putative *ppgK* genes led us to investigate whether glucose 6-phosphate synthesis in *Cyanobacteria* could take place enzymatically through a similar way to that previously described in *M. tuberculosis* along with other *Actinobacteria*. To characterize NsPPGK and NpPPGK, their respective putative genes were obtained from genomic DNA by PCR amplification which yielded a single product with the expected size of 0.72 kb in both cases (Fig. 1a). They were lastly cloned into the pQE-80L expression vector and over-expressed in *E. coli* (BL21). Protein expression was induced in early-log phase cultures by addition of IPTG. The heterologous overexpression of cyanobacterial *ppgK* genes conferred high PPGK activity to *E. coli* cells. Thus, crude extracts from induced *E. coli* cells overproducing NsPPGK or NpPPGK showed fairly high glucokinase activity levels with $P_{13\text{-}18}$ as a substrate, in the range of 0.15 to 0.20 μmol min^{-1} mg^{-1} protein, respectively. In contrast, no PPGK activity was detected in extracts from cells containing the pQE-80L plasmid with no insert. Milligram quantities of the respective N-terminal His_6-tagged fusion proteins were subsequently isolated in ca. 95% purity after one-step affinity purification onto a HisTrap FF Crude Ni-NTA column (Online Resources Figs. S2 and S3, and Table S4). Enzyme purity was further enhanced by following FPLC gel-filtration chromatography, which was confirmed by electrophoresis on SDS-PAGE gels (Fig. 1, Online Resource Table S4). Thus, a single protein band of ca. 27 kDa was found in both purified PPGK preparations (Fig 1a), in good agreement with the nominal M_m values of 27,339 and 27,236 Da predicted for the recombinant NsPPGK and NpPPGK polypeptides, respectively. Besides, native M_m values and oligomeric states of oligomeric states of the recombinant proteins were determined by gel-filtration chromatography, and values of 49.4 ± 4 kDa and 55.1 ± 5 kDa (means ± SE of three independent determinations) were obtained for NsPPGK and NpPPGK, respectively (Fig. 1b). Therefore, both proteins adopted a stable dimeric arrangement in solution. In accordance with these results, MALDI-TOF determination of absolute M_m values gave values of 27,287 Da ± 0.1% and 27,236 Da ± 0.1% for the recombinant NsPPGK and NpPPGK subunits, respectively. In addition, the identities of the recombinant NsPPGK and NpPPGK polypeptides were confirmed by peptide mass fingerprinting covering respectively about 55 and 82 % of the natural sequences, and eventual identification by MALDI-TOF MS (Online resource Fig. S4). Together, these active and high purity fractions were used for the subsequent determination of their enzymatic kinetic parameters.

NsPPGK and NpPPGK are strictly polyP-dependent glucokinases with preference for long-chain PolyP

The purified recombinant NsPPGK showed no activity towards ATP, CTP, GTP, TTP, or dATP as compared to sorts of polyP (Fig. 2). The absolute specificity of NsPPGK for inorganic polyphosphates appears to be a common property of PPGK enzymes in other heterocystous filamentous cyanobacteria, since similar results were observed in the characterization of NpPPGK. The substrate specificities concerning polyP as phosphoryl donor to produce glucose 6-phosphate by cyanobacterial PPGKs were probed using synthetic polyP molecules of various chain lengths at saturating glucose levels (Fig. 2a). The rate of sugar phosphorylation for the polyP chain lengths followed a similar trend in both recombinant enzymes, longer polyP result in higher specific glucokinase activity. This indicated that PPGKs from cyanobacteria bound and hydrolyzed long-chain polyP substrates most efficiently (Table 1). This highlights its reasonable consistency with previous findings on polyP/ATP glucokinases of other bacteria (Girbal et al. 1989; Hsieh et al. 1996b; Tanaka et al. 2003; Mukai et al. 2003; Lindner et al. 2010a). Noteworthy, both cyanobacterial PPGKs are also able to use short-chain polyP. With reference to the sole crystal structure of a bacterial polyP/ATP glucomannokinase published to date (Mukai et al. 2004), it has been proposed that there is a minimal length between two phosphoryl groups consisting on a putative pentapolyphosphate-binding site. However, NsPPGK and NpPPGK exhibit modest but significant specific activity levels with P_4 (5-7 μmol min^{-1} mg^{-1} protein) (see Table 1). Analogous experiments revealed that cyanobacterial PPGKs were unable to use shorter polyP than P_4, such as P_3, P_{3c} or PPi. Estimation of the kinetic parameters of NsPPGK with different polyP and hexose substrates revealed that there is a remarkable increase of the catalytic constant k_{cat} (more than 30-fold) with increasing polyP chain-length from P_4 up to P_{LC} while K_m values remain fairly constant, which explained the higher catalytic efficiency of long-chain polyP (Tables 1 and 2, Figs. S5 and S6).

To determine the mechanism of polyP utilization by NsPPGK, P_{150} at saturation concentration was used as a phosphoryl substrate while the progress of the reaction was monitored by collecting sequential aliquots at increasing times. PolyP were isolated and electrophoresed on a preparative polyacrylamide gel, and eventually visualized with Toluidine blue O staining. As shown in Fig. 3, this medium-size polyP was utilized by the cyanobacterial PPGK by an essentially non-processive mechanism, as was evidenced by the non noticeable broadening of the range of polyP sizes with the reaction time. A non-processive mechanism is also consistent with the observed formation of polyP of intermediate sizes from the longest polyP during the reaction progress (see Fig. 3).

A variety of compounds which are analogs to the phosphoryl donors were also tested to assess whether or not they could act as PPGK activity inhibitors (Table 3). P_3 and PPi were fairly strong inhibitors, with K_i values of 0.13 and 0.19 mM respectively, while ATP only modestly inhibited the activity of NsPPGK. In addition, a control experiment with increasing concentrations of NaCl was conducted to determine the effect of the ionic strength on the PPGK activity. Results from Table 3 indicated that NsPPGK was not as severely inhibited by NaCl as by short polyP or ATP, since the observed concentrations required for substantial enzyme inhibition, most probably produced by ionic strength effect, were much higher (50-150 mM range). An inhibitory effect on PPGK activity was also obtained with KCl being even more marked than that of NaCl (Table 3), thus suggesting that electrostatic forces may be involved in the interaction between polyP and the enzyme.

NsPPGK and NpPPGK possess a modest but significant polyP-dependent mannokinase activity

Besides glucose, the cyanobacterial PPGKs phosphorylated mannose as well, but just in a minor extent (Table 2). The specific activity values of NsPPGK for glucose and mannose were 229.1 and 3.1 μmol min^{-1} mg^{-1}, respectively. Values of the same order of magnitude were obtained for NpPPGK. A further study of their catalytic efficiencies evidenced that cyanobacterial PPGKs clearly exhibit a remarkable preference (approx. 100-fold higher) for glucose instead of mannose as a substrate (Table 2). Conversely, fructose was totally inactive as a phosphoryl acceptor.

Cyanobacterial PPGKs are divalent-cation dependent enzymes with distinctive alkaline pH optimum and remarkable thermotolerance

The activities of both cyanobacterial PPGKs were absolutely dependent on the presence of a divalent cation in the reaction mixture. Mg^{2+} was the optimal metal cofactor for glucose phosphorylation by both recombinant PPGKs, while Mn^{2+} and Fe^{2+} ions function in a lesser extent and no activity was detected in the presence of $Co^{2+,}$ Ca^{2+} or Cu^{2+} ions (Fig. 4a). No activity was detected after incubation of the enzyme samples with 10 mM EDTA, and subsequent dialysis to remove all traces of EDTA resulted in complete loss of activity. Addition of 5 mM Mg^{2+} restored the full PPGK activity. The highest specific activity with magnesium ions was found in the concentration range 4-6 mM with an optimum estimated at 5 mM and a calculated K_m value of 1.3 mM (Fig. 4b). Higher concentrations of Mg^{2+} resulted in a decrease of PPGK activity.

Cyanobacterial PPGK activity was optimal at the alkaline pH range, between pH 8.5 and 9.0 (Fig. 5a). Nevertheless, the activity declined quite rapidly at higher pH values with no activity remaining at pH 10.5 or higher. A very similar pH dependence curve was obtained for both enzymes. Alkaline pH optimum is a common distinctive feature of other cyanobacterial enzymes when compared with their orthologs of non-photosynthetic bacteria and eukaryotes (Serrano et al. 1984, Serrano et al. 1992).

Both cyanobacterial PPGKs showed an optimal temperature as high as 45 °C (Fig. 5b). Indeed, PPGK from the actinobacterium *Arthrobacter* sp. (Mukai et al. 2003) exhibits a similar value, but considerably higher as compared to 30 °C for the PPGK from *Microlunatus phosphovorus* and most others actinobacterial polyP/ATP glucokinases (Tanaka et al. 2003). To determine the thermostability of cyanobacterial PPGK, NsPPGK was preheated at 40 ºC, 50 ºC, 60 ºC and 70 ºC for 30 min. No loss of activity was observed after incubation below 50 ºC. At 50 ºC, 45 % activity remained. However, only 7 % of PPGK activity remained at 60 ºC implying that NsPPGK is unable to tolerate these fairly high temperatures. Finally, this PPGK was irreversibly inactivated when exposed to temperatures above 60 ºC for 30 min.

Closely-related PPGK orthologs occur among diazotrophic *Cyanobacteria*

A bioinformatic search was carried out looking for PPGK homologs in a range of filamentous and unicellular strains representative of any of the five cyanobacterial taxonomic sections defined by Rippka et al. (1979). Then, a number of putative *ppgK* genes were identified in strains belonging to all these taxonomic groups. With two exceptions, the nitrogen-fixing strains *Synechococcus* sp. PCC7335 (Bergman et al. 1997) and *Synechococcus* sp. PCC7502, no PPGK-like ORFs were found so far among sequenced genomes of unicellular species from section I, typically non nitrogen-fixing, e.g. *Synechocystis* sp. PCC6803 and *Synechococcus elongatus* PCC7942. Likewise, no hybridization band was observed in Southern blot analysis (Fig. 6) and no PPGK activity was detected in whole-cell extracts of the two latter strains (Table 4). In contrast, clear hybridization bands and significant levels of PPGK activity were detected with several polyPs in a number of strains of sections II to V with sequenced genomes exhibiting predicted *ppgK* genes as expected (such as *Dermocarpa* sp. PCC7437 and *Nostoc* spp.), but also in other diverse *Cyanobacteria* whose genomes have not been sequenced yet, such as section III strain *Pseudanabaena* sp. PCC6903, the section IV strains *Nostoc* sp. PCC6719, *Calothrix* sp. PCC7601, *Calothrix* sp. PCC9327, *Anabaena* sp. ATCC33047 and *Nodularia chucula*, and section V strain *Fischerella muscicola* (Fig. 6 and Table 4). Therefore, closely related putative polyP-dependent PPGKs seem to be widely distributed among diazotrophic, mostly multicellular, cyanobacterial strains.

DISCUSSION

A BLAST sequence similarity search in the genome of *Nostoc sp.* PCC7120 revealed one ORF, *all*1371, with high homology to the well-characterized *Mycobacterium tuberculosis* H37Ra *ppgK* gene (Szymona and Widomski 1974; Hsieh et al. 1996a). A similar approach revealed another putative PPGK encoding gene, Npun_R1878, in the genome of *Nostoc punctiforme* PCC73102, which was annotated as encoding a ROK (transcriptional regulator/sugar kinase) family protein which share a 91 % sequence identity with its homolog of *Nostoc* sp. PCC7120. Subsequent searches in bioinformatic databases identified about other forty putative cyanobacterial PPGK orthologs. They are predicted to be highly-similar proteins of about 230-250 amino acid residues, clearly smaller than their conventional ATP-glucokinase counterparts (290-330 residues), and most of them were unprecisely annotated as ROK family proteins or transcriptional regulators/sugar kinases. This finding together with the deduced ROK family domain architecture characteristic of other previously reported bacterial PPGKs predicted for all cyanobacterial orthologs, led us to investigate whether glucose 6-phosphate synthesis could take place in *Cyanobacteria* through PPGK enzymes.

This presumption was confirmed by the biochemical characterization of two recombinant cyanobacterial PPGK proteins purified by metal-affinity and size-exclusion chromatographies as above described. As a result, both *Nostoc* proteins were functionally validated with the ability to phosphorylate glucose and, to a lesser extent, mannose. These enzymatic reactions occurred using a wide range of polyP with different chain lengths as phosphoryl donors. However, no activity was detected with the shortest chain-length polyPs, namely pyrophosphate (PPi), $P_{3c,}$ or P_3. In fact, P_4 was confirmed as the shortest polyP active as

substrate for PPGK enzymes described so far. Concerning the length of the chain of active polyP substrates, both *Nostoc* proteins seemed to follow a similar pattern, with higher catalytic efficiencies for long-chain polyP ($\geq P_{60}$). Remarkably, reaction rates of polyP utilization increased considerably with the number of phosphate residues per molecule. To our knowledge, Mukai et al. reported the sole crystal structure so far available for a polyP/ATP-glucokinase – from the actinobacterium *Arthrobacter* sp. KM – and it was complexed with one glucose and two phosphate molecules instead of polyP (Mukai et al. 2004). According to this model, P_5 has been claimed as the shortest polyP able to enzymatically phosphorylate glucose. In contrast, we showed that both NsPPGK and NpPPK are able to generate glucose 6-phosphate when using P_4 as a phosphoryl donor, although with a lower efficiency than longer-chain polyPs. Apart from this, it is noteworthy that both cyanobacterial PPGKs were strictly dependent on polyP, as there was no activity detected when ATP or any other NTP were used as phosphoryl substrates. This feature has been only described to date for the PPGK of the primitive, polyP-accumulating actinobacterium *Microlunatus phosphovorus* (Tanaka et al. 2003). Here we describe a novel subfamily of PPGK enzymes characteristic of *Cyanobacteria*, all of them being strictly dependent of polyP as the phosphoryl donor.

Using multiple sequence alignment of the polyP/ATP-glucomannokinase from *Arthrobacter* sp. KM and other actinobacterial polyP/ATP glucokinases an specific extra heptapeptide (PEAPAAG) was identified in the conserved glucose region of the former protein which was proposed as responsible for the mannose-phosphorylating ability of the polyP/ATP-glucomannokinase. In fact, PPGK from *Arthrobacter* sp. KM can phosphorylate fructose as well (Mukai et al. 2003). In addition, Szymona *et al*. have shown that when *Mycobacterium phlei* was grown on fructose, a polyP-fructokinase activity was found. Contrastingly, when grown on mannose, polyP-mannokinase was detected (Szymona and Ostrowski 1964). Nonetheless, despite lacking such heptapeptide, this work shows that both cyanobacterial PPGKs are able to phosphorylate mannose, although with fairly modest levels and a notably reduced catalytic efficiency compared to glucose. However, no significant polyP-fructokinase activity was detected for NsPPGK and NpPPGK.

Some other features of NsPPGK and NpPPGK were in some extent distinct to those previously described for other bacterial PPGKs. Thus, optimum pH was clearly alkaline, 8.5-9.0, while actinobacterial PPGKs have almost neutral optimal pH values (e.g. 7.5 for the *Arthrobacter* enzyme). Also, the notable thermostability of cyanobacterial PPGKs (optimal temperature, ca. 45º C) is an outstanding catalytic feature that, like its alkaline optimal pH, may have biotechnological relevance. Other biochemical features were similar to those of other sugar kinases; thus, both cyanobacterial PPGKs required divalent metal cations, to which Mg^{2+} was preferred. Similarly, they were identified as homodimers although with natural subunit M_m values somewhat lower than those of bacterial polyP/ATP-dependent glucokinases (ca. 30 kDa) and eukaryotic hexoquinases (ca. 35 kDa).

It has been hypothesized that polyP could be the phosphoryl donors for ancient organisms, and they were later replaced by ATP in the evolution (Lipman 1965). This is based on the assumption that the Gibbs free energy of polyP hydrolysis is similar to the ATP, and their likely occurrence since prebiotic times. An interesting observation is that PPGK activities have been reported to date only in the comparatively

ancient order of *Actinomycetales*. Noteworthy, in bacteria belonging to this order, the ratio polyP-glucokinase *vs.* ATP-glucokinase activities is higher in more phylogenetically ancient representatives (Hsieh et al. 1993; Phillips et al. 1999). According to this hypothesis, the following stage in the evolution of sugar kinases might be played by the dual ATP/polyP glucokinases, like the PPGKs described in *Propionibacterium shermanii*, *Mycobacterium tuberculosis* or other *Actinobacteria* (Pepin and Wood 1986; Kowalczyk et al. 1996; Hsieh et al. 1996b). Lastly in sequence evolution, this role would be played by hexokinases which all are strictly dependent on ATP (Bork et al. 1993). For this reason, it would be expected that PPGK from more primitive bacteria, such as *Microlunatus phosphovorus* or cyanobacterial species were strictly dependent on polyP. It was also expected, therefore, that a similar analysis carried out with PPGKs from this ancient group of photosynthetic prokaryotes may shed light on the origin and evolution sugar kinases.

Likewise PPGK where polyP can be employed instead of ATP, the polyP/ATP-dependent NAD kinase (PPNK, EC 2.7.1.23) forms NADP using either polyP or ATP. Characterized (Lindner et al. 2010b) or putative PPNKs are identified in *Actinobacteria* already described to possess PPGK. Surprisingly, no putative PPNKs were revealed after Blast sequence similarity searches in the *Nostoc* sp. PCC7120 and *Nostoc punctiforme* PCC73102 genomes, as well as in many other cyanobacterial genomes (data not shown). These findings might suggest that the series of genes involved in polyP metabolism of filamentous nitrogen-fixing *Cyanobacteria* are characteristic.

The occurrence of PPGK orthologs in other *Cyanobacteria* was confirmed following a multidisciplinary approach based on Southern blot experiments and PPGK activity level determinations in whole-cell extracts. Thus, using the full *ppgK* gene from *Nostoc* sp. PCC 7120 as a probe putative *ppgK* genes were identified in genomic Southern blot analysis of a number of diverse cyanobacterial species belonging to sections II, III, IV, and V of the classification of Rippka et al. (1979). However, no orthologs were detected in unicellular species from section I, such as *Synechocystis* sp. PCC6803, *Thermosynechococcus enlongatus* BP-1, and others. As a consequence, the occurrence of PPGK might be a characteristic feature of nitrogen-fixing cyanobacterial species, like the heterocystous filamentous species of sections IV and V, as well as the non-heterocystous filamentous and colonial species of section III which fix nitrogen in microaerobiosis.

An amino acid alignment including the aminoacid sequences of NsPPGK and NpPPGK proteins and those of the known PPGKs from other bacteria showed extensive sequence similarity (Online Resource Fig. S1). More importantly, the seven characteristic motifs of this protein family and all amino acid residues shown to be involved in catalysis (Mukai et al. 2003 and 2004) are conserved. These results along with the biochemical characterization presented in this work clearly demonstrate that *all*1371 and Npun_R1878 encode functional polyP-gluco/mannose phosphotransferases, hence its re-annotation. Based on the above sequence similarities, the *ppgK* assignation for both cyanobacterial genes is further supported by molecular phylogenetic analyses (Fig. 7). As molecular phylogenetic data shown, their encoded PPGK proteins form a compact well-supported cluster, clearly divergent from the actinobacterial PPGK assembly, with a number of putative orthologs encoded by the genomes of selected unicellular, colonial and filamentous cyanobacterial strains. I should be noted in this respect that about fifty putative

cyanobacterial PPGK orthologs were identified in databases searches (July 2014) (Online Resource Table S3). Noteworthy, PPGK orthologs of marine cyanobacterial strains such as *Acaryochloris marina* MBIC11017 and *Nodularia spumigena* CCY9414 are also included in this group. This suggests that, cyanobacterial polyP is possibly used as an alternative source of energy in place of ATP in the ocean environment as well. All in all, our sequence comparison and molecular phylogenetic data reveal that cyanobacterial PPGKs are structurally simpler and presumably more ancient than their homologs of *Actinobacteria*. With the exception of the enzyme of the polyP-accumulating actinobacterium *M. phosphovorus*, all the other PPGKs described to date utilize polyP as well as ATP. These findings agree with the ancestral character of *Cyanobacteria*, and suggest that the strictly polyP-dependent PPGKs may represent molecular relicts of a hypothetical ancient world in which polyP could be preferentially used for metabolic functions.

The present work also envisages new perspectives for an innovative costly-effective enzymatic production of glucose 6-phosphate or mannose 6-phosphate by a novel class of strictly-polyP dependent glucokinases from diazotrophic *Cyanobacteria*. Thus, synthesis of sugar-phosphates could be performed by immobilized-engineered cyanobacterial PPGKs from the very-stable inorganic polymer polyP without continuous regeneration of ATP, an expensive cofactor required by conventional hexokinases.

ACKNOWLEDGMENTS

This work was supported by research grants from the Spanish (BFU2004-00843, BFU2007-61887) and Andalusian Regional (PAIDI group BIO-261) Governments, all of them partially funded by the EU FEDER program. PAIDI group BIO-261 belongs to the CeiA3 and AndaluciaTECH University Campuses of International Excellence. Authors thank Dr. Toshikazu Shiba (RegeneTiss Co, Japan) for generously providing highly-purified polyP samples and to Dr M. R. Gómez-García for helpful suggestions and discussions.

REFERENCES

Abed RM, Dobretsov S, Sudesh K (2009) Applications of *cyanobacteria* in biotechnology. J Appl Microbiol 106(1):1-12 doi:10.1111/j.1365-2672.2008.03918.x

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J. Mol. Biol. 215:403-410

Ausubel FM, Brent R, Kingston RE. Moore DD, Seidman JG, Smith JA, Struhl K (1992) Current Protocols in Molecular Biology, Greene Publishing and Wiley-Interscience, New York

Benson DA, Cavanaugh M, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW (2013) GenBank. Nucleic Acids Res 41:D36-D42. doi: 10.1093/nar/gks1195

Bergman B, Gallon JR, Rai AN, Stal LJ (1997) N2 Fixation by non-heterocystous *cyanobacteria*. FEMS Microbiology Reviews 19(3):139-185 doi:10.1111/j.1574-6976.1997.tb00296.x

Bork P, Sander C, Valencia A (1993) Convergent evolution of similar enzymatic function on different protein folds: the hexokinase, ribokinase, and galactokinase families of sugar kinases. Protein Sci 2:31-40. doi: 10.1002/pro.5560020104

Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72:248-54

Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, Heger A, Hetherington K, Holm L, Mistry J, Sonnhammer EL, Tate J, Punta M (2014) Pfam: the protein families database. Nucleic Acids Res 42(Database issue):D222-30 doi:10.1093/nar/gkt1223

Fujisawa T, Okamoto S, Katayama T, Nakao M, Yoshimura H, Kajiya-Kanegae H, Yamamoto S, Yano C, Yanaka Y, Maita H, Kaneko T, Tabata S, Nakamura Y (2014) CyanoBase and RhizoBase: databases of manually curated annotations for cyanobacterial and rhizobial genomes. Nucleic Acids Res 42(Database issue):D666-70 doi:10.1093/nar/gkt1145

Girbal E, Binot RA, Monsan RE (1989) Production, purification, properties and kinetic studies of free and immobilized polyphosphate:glucose-6-phosphotransferase from *Mycobacterium phlei*. Enzyme Microb Technol 11:518-527

Gomez-Garcia MR, Losada M, Serrano A (2003) Concurrent transcriptional activation of *ppa* and *ppx* genes by phosphate deprivation in the cyanobacterium *Synechocystis* sp. strain PCC 6803. Biochem Biophys Res Commun 302(3):601-9

Gouy M, Guindon S, Gascuel O (2010) SeaView 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Mol Biol Evol 27:221-224. doi: 10.1093/molbev/msp259

Hehuan L, Suwan Myung, Zhang YHP (2012) One-step purification and immobilization of thermophilic polyphosphate glucokinase from *Thermobifida fusca* YX: glucose-6-phosphate generation without ATP. Appl Microbiol Biotechnol 93*:*1109–1117. doi: 10.1007/s00253-011-3458-1

Hernandez A, Ruiz MT (1998) An EXCEL template for calculation of enzyme kinetic parameters by non-linear regression. Bioinformatics 14:227–228

Hseih PC, Shenoy BC, Jentoft JE, Phillips NFB (1993) Purification of polyphosphate and ATP glucose phosphotransferase from *Mycobacterium tuberculosis* H_{37}Ra: evidence that poly(p) and ATP glucokinase activities are catalyzed by the same enzyme. Protein Expres Purif 4:76-84

Hsieh PC, Shenoy BC, Samols D, Phillips NFB (1996a) Cloning, expression, and characterization of polyphosphate glucokinase from *Mycobacterium tuberculosis*. J Biol Chem 271:4909-4915

Hsieh PC, Kowalczyk TH, Phillips NFB (1996b) Kinetic mechanisms of polyphosphate glucokinase from *Mycobacterium tuberculosis*. Biochemistry 35:9772-9781

Hunter S, Jones P, Mitchell A, Apweiler R, Attwood TK, Bateman A, Bernard T, Binns D, Bork P, Burge S,de Castro E, Coggill P, Corbett M, Das U, Daugherty L, Duquenne L, Finn RD, Fraser M, Gough J, Haft D, Hulo N, Kahn D, Kelly E, Letunic I, Lonsdale D, Lopez R, Madera M, Maslen J, McAnulla C, McDowall J, McMenamin C, Mi H, Mutowo-Muellenet P, Mulder N, Natale D, Orengo C, Pesseat S, Punta M, Quinn AF, Rivoire C, Sangrador-Vegas A, Selengut JD, Sigrist CJA, Scheremetjew M, Tate J, Thimmajanarthanan M, Thomas PD, Wu CH, Yeats C, Yong SY (2011) InterPro in 2011: new developments in the family and domain prediction database. Nucleic Acids Res. doi: 10.1093/nar/gkr948

Jensen TE, Baxter M, Rachlin JW, Jani V (1982) Uptake of heavy metals by *Plectonema boryanum* (*cyanophyceae*) into cellular components, especially polyphosphate bodies: An X-ray energy dispersive study. Environmental Pollution Series A, Ecological and Biological 27(2):119-127 doi:http://dx.doi.org/10.1016/0143-1471(82)90104-0

Koenig T, Menze BH, Kirchner M (2008) Robust prediction of the MASCOT score for an improved quality assessment in mass spectrometric proteomics. J Proteome Res 7:3708–3717. doi: 10.1021/pr700859x

Koide M, Miyanaga A, Kudo F, Eguchi T (2013) Characterization of polyphosphate glucokinase SCO5059 from *Streptomyces coelicolor* A3(2). Biosci Biotechnol Biochem 77:130498-1-3

Kornberg A, Rao NN, Ault-Riché D (1999) Inorganic polyphosphate: a molecule with many functions. Ann Rev Biochem 68:89-125

Kowalczyk TH, Horn PJ, Pan WH, Phillips NFB (1996) Initial rate and equilibrium isotope exchange studies on the ATP-dependent activity of polyphosphate Glucokinase from *Propionibacterium shermanii*. Biochemistry 35:6777-6785

Kulaev IS (1979) The biochemistry of inorganic polyphosphates. John Wiley & Sons, Inc., New York

Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227(5259):680-5

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) ClustalW and ClustalX version 2. Bioinformatics 23: 2947-2948

Lawry N, Jensen T (1979) Deposition of condensed phosphate as an effect of varying sulfur deficiency in the cyanobacterium *Synechococcus* sp. (*Anacystis nidulans*). Arch Microbiol 120(1):1-7 doi:10.1007/bf00413264

Lindner SN, Knebel S, Pallerla SR, Schoberth SM, Wendisch VF (2010a) Cg2091 encodes a polyphosphate/ATP-dependent glucokinase of *Corynebacterium glutamicum*. Appl Microbiol Biotechnol 87:703–713. doi: 10.1007/s00253-010-2568-5

Lindner SN, Niederholtmeyer H, Schmitz K, Schoberth SM, Wendisch VF (2010b) Polyphosphate/ATP-dependent NAD kinase of *Corynebacterium glutamicum*: biochemical properties and impact of *ppnK* overexpression on lysine production. Appl Microbiol Biotechnol 87:583-93. doi: 10.1007/s00253-010-2481-y.

Lipmann F (1965) The origins of prebiological system and their molecular matrices. In: Fox SW (ed) Academic Press, New York, pp 259-280

Mukai T, Kawai, S, Matsukawa H, Matuo Y, Murata K (2003) Characterization and molecular cloning of a novel enzyme, inorganic polyphosphate/ATPglucomannokinase, of *Arthrobacter* sp. strain KM. Appl Environ Microbiol 69:3849-3857

Mukai T, Kawai S, Mori S, Mikami B, Murata K (2004) Crystal Structure of Bacterial Inorganic Polyphosphate/ATP-glucomannokinase. Insights into kinase evolution. J Biol Chem 279:50591-50600. doi: 10.1074/jbc.M408126200

Nordberg H, Cantor M, Dusheyko S, Hua S, Poliakov A, Shabalov I, Smirnova T, Grigoriev IV, Dubchak I (2014) The genome portal of the Department of Energy Joint Genome Institute: 2014 updates. Nucleic Acids Res 42(Database issue):D26-31 doi:10.1093/nar/gkt1069

Pepin CA, Wood HG (1986) Polyphosphate glucokinase from *Propionibacterium shermanii*. Kinetics and demonstration that the mechanism involves both processive and nonprocessive type reactions. J Biol Chem 261:4476-4480

Phillips NF B, Hsieh PC, Kowalczyk TH (1999) Polyphosphate glucokinase. Prog Mol Subcell Biol 23:101-125.

Rao NN, Gómez-García MR, Kornberg A (2009) Inorganic polyphosphate: Essential for growth and survival. Annu Rev Biochem 78:605–647. doi: 10.1146/annurev.biochem.77.083007.093039

Rippka R, Deruelles J, Waterbury JB, Hermann M, Stainer RY (1979) Generic assignment, strains histories and properties of pure cultures of *cyanobacteria*. J Gen Microbiol 111:1-16

Schopf JW (2002) The Fossil Record: Tracing the Roots of the Cyanobacterial Lineage. In: Whitton B, Potts M (eds) The Ecology of *Cyanobacteria*. Springer Netherlands, pp 13-35

Serrano A, Rivas J, Losada M (1984) Purification and properties of glutathione reductase from the cyanobacterium *Anabaena* sp. strain 7119. J Bacteriol 158(1):317-24

Serrano A (1992) Purification, characterization and function of dihydrolipoamide dehydrogenase from the cyanobacterium *Anabaena* sp. strain P.C.C. 7119. Biochem J 288 (Pt 3):823-30

Szymona M (1957) Utilization of inorganic polyphosphates for phosphorylation of glucose in *Mycobacterium phlei*. Bull. Acad. Pol. Sci. Ser. Sci. Biol. 5:379-381

Szymona M, Ostrowski W (1964) Inorganic polyphosphate glucokinase of *Mycobacterium phlei*. Biochim Biophys Acta 85:283-295

Szymona M, Widomski J (1974) A kinetic study on inorganic polyphosphate glucokinase from *Mycobacterium tuberculosis* H37RA. Physiol Chem Phys 6:393-404

Szymona O, Szymona M (1978) Multiple forms of polyphosphate-glucose phosphotransferase in various *Mycobacterium* strains. Acta Microbiol Pol 27:73-76

Szymona O, Szymona M (1979) Polyphosphate- and ATP-glucose phosphotransferase activities of *Nocardia minima*. Acta Microbiol Pol 28:153-160

Tanaka S, Lee SO, Hamaoka K, Kato J, Takiguchi N, Nakamura K, Ohtake H, Kuroda A (2003) Strictly polyphosphate-dependent glucokinase in a polyphosphate-accumulating bacterium, *Microlunatus phosphovorus*. J Bacteriol 185: 5654-5656

Thompson PA, Oh H-M, Rhee G-Y (1994) Storage of phosphorus in nitrogen-fixing *Anabaena flos-aquae* (*Cyanophyceae*). J Phycol 30:267-273

Van Wazer JR (1958) Phosphorus and its compounds, Vol 1. Interscience Publishers, New York

Wood HG, Clark JE (1988) Biological aspects of inorganic polyphosphates. Annu Rev Biochem 57:235-260

FIGURE CAPTIONS

Fig. 1 (a) *Upper panel*. PCR amplification of cyanobacterial *ppgK* genes. An electrophoretic analysis of PCR-amplified DNA fragments corresponding to the *ppgK* genes of *Nostoc* sp. PCC7120 (lane 1) and *Nostoc punctiforme* PCC73102 (lane 2), and DNA size markers (M), is shown. Amplification reactions were performed with specific primers pairs and cyanobacterial genomic DNA as a template, as described in Materials and Methods, and subsequently loaded onto 1.2 % agarose-TBE gel. As shown, a single DNA band of approximately 0.72 kb was obtained in each case (arrow). ***Lower panel.*** SDS-PAGE (12 %,

w/v, polyacrylamide, 0.5 %, w/v, SDS) analysis of recombinant NsPPGK and NpPPGK purified after FPLC gel filtration. Approx. 10 µg of NsPPGK (27.34 kDa, nominal subunit M_m) and NpPPGK (27.24 kDa, nominal subunit M_m) were applied per lane. *M*, protein standards. Numerals on the left indicate the molecular masses (kDa). Arrow indicates the PPGK protein bands. (b) FPLC gel filtration chromatography analyses of native M_m and oligomeric states of the PPGKs from *Nostoc* spp. Aliquots (0.5 ml) of metal-chelated chromatography purified preparations of recombinant NsPPGK and NpPPGK were applied to a Superdex© 200 PG column. Calibration curves with protein standards (Thy, thyroglobulin; Fer, ferritine; Amy, β-amylase; ADH, alcohol dehydrogenase; BSA, bovine seroalbumin; CA, carbonic anhydrase; Cyt.c, cytochrome c) are displayed on the left upper corner of the chromatography elution profile figures. A SDS-PAGE analysis of selected fractions around the central peak fraction (50 µl aliquots applied per lane) is also shown. Note that single elution peaks, corresponding to absorbance at 280 nm (broken line) and polyP-glucokinase activity (filled circles, solid line), overlapped in both cases. The asterisks indicate the fraction peaks of recombinant PPGKs as determined by their enzymatic activity and absorbance at 280 nm. Native M_m values of 49.4 and 55.1 kDa were estimated for NsPPK and NpPPGK, respectively. K_{av}, phase distribution coefficient of the analyzed proteins

Fig. 2 Substrate specificity of cyanobacterial recombinant PPGKs. PolyP-glucokinase activity levels of purified NsPPGK (black bars) and NpPPGK (white bars) were determined using polyPs of different chain lengths (panel a) or diverse NTPs (panel b) as phosphoryl donor substrates. Activity levels were obtained from three independent experiments and are shown as means ± S.E. Note that both cyanobacterial PPGKs are strictly polyP dependent glucokinases, and long-chain polyPs are their optimal substrates. No significant activity was detected with either NTPs, PPi, P_{3c} or P_3

Fig. 3 Non-processive utilization of P_{150} by NsPPGK. 2.5 mM of P_{150} was used as a substrate for purified NsPPGK (approx. 3 µg/ml) following the standard assay conditions, as described in the Material and Methods section. At different time intervals, sequential aliquots were collected and polyP was isolated, electrophoresed on a preparative PAGE gel, and finally stained with Toluidine blue O. *Lane 1* is zero time, *lanes 2 to 10* correspond to 3, 6, 8, 10, 12, 14, 16, 18 and 20 min, respectively

Fig. 4 Biochemical characterization of recombinant NsPPGK (black bars) and NpPPGK (white bars) regarding to metal cations dependence of polyP (P_{60}) glucokinase activity. (a) Metal cofactor specificity. Several divalent metal cations were added at 5 mM concentration to the assay mixtures. No detectable activity was measured with Mg^{2+} in the presence of 10 mM EDTA. Bars represent activity levels from three independent experiments and are shown as means ± S.E. Activity is expressed in relative units (100 % percentage assigned to the optimum condition in each case). 100 % activity levels correspond to 81.7±7.4 and 95.8±12.5 µmol min^{-1} mg^{-1} for NsPPGK and NpPPGK, respectively. (b) NsPPGK activity dependence on Mg^{2+} concentration. Each point represents the mean activity value ± S.E. of three independent experiments. As shown, no activity was detected either in the absence of a divalent cation or with an excess of the chelating agent EDTA

Fig. 5 Effect of the pH (panel a) and temperature (panel b) on the polyP (P_{60}) glucokinase activity of NsPPGK (filled circles) and NpPPGK (open circles). Data are shown as relative units (100 % percentage) and were assigned to the optimum condition in each case. Activity levels were obtained from three independent experiments and are shown as means ± S.E. 100 % activity values correspond to 80.6±7.3 and 93.5±9.2 µmol min^{-1} mg^{-1} (panel a) and 80.4±6.7 and 92.7±7.3 µmol min^{-1} mg^{-1} (panel b) for NsPPGK and NpPPGK, respectively

Fig. 6 Experimental evidence for the widespread occurrence of homologs of *Nostoc* spp. *ppgK* genes among diazotrophic *Cyanobacteria*. A search of putative *ppgK* genes was carried out by S*outhern* blot analysis with diverse cyanobacterial strains representatives of the taxonomic sections (roman numerals) in the classification of Rippka et al. (1979). The strains are identified by their collection numbers. The positions of *Eco*RI-*Hin*dIII-restricted λ DNA fragments used as standards (in the range of 21 to 2 kb) are indicated on the left side. Genomic DNAs (approx. 5 µg) were digested with *Hin*dIII (left panel) or *Eco*RI (right panel) restriction enzymes. The full coding sequence of the *ppgK* gene from *Nostoc* sp. PCC7120

was used as a probe under heterologous hybridization conditions at 55°C. As shown, no hybridization band was observed only in the lanes corresponding to unicellular non-diazotrophic *Cyanobacteria Synechocystis* sp. PCC6803 and *Synechococcus* sp. PCC7942

Fig. 7 Molecular phylogenetic analysis of cyanobacterial PPGKs. Unrooted Neighbor-Joining (a) and Maximum Parsimony (b) phylogenetic trees, obtained from amino acid sequence alignments of selected bacterial orthologs, are shown. A similar topology was obtained for a Maximum Likelihood tree (not shown). Numbers in selected nodes are bootstrap percentages based on 1,000 replicates. Scale bar indicates number of changes per amino acid site. Most cyanobacterial strains are identified by their PCC numbers. Biochemically characterized PPGKs are shown boxed, and the three strictly polyP-dependent enzymes characterized so far are moreover shown in boldface. Other predicted PPGK sequences were obtained from public databases (UniProtKB and IMG-JGI databases) and their details are summarized in Online Resource Table S3. Note the well-defined and robust cyanobacterial cluster (shaded) which is clearly divergent from the actinobacterial assembly of dual ATP/polyP-dependent homologs, as well as the two deeply-branched clusters of uncharacterized putative PPGKs from α- and β-*proteobacteria* closely related to the cyanobacterial assembly

TABLES

Table 1. Kinetic parameters of purified recombinant polyP-gluco(manno)kinase from *Nostoc* sp. PCC7120 with different polyPs and hexoses as substrates

Substrate ($polyP_n$)	V_{max} (μmol min^{-1} mg^{-1})	K_m [a] (μM)	k_{cat} (s^{-1})	Catalytic efficiency k_{cat}/K_m ($mM^{-1}s^{-1}$)
P_4	5.7	29.9	4.9	164
P_{13-18}	31.4	37.7	27.0	717
P_{60}	81.7	39.8	70.2	1764
P_{LC}	176.3	49.5	151.4	3059
Glucose (P_{LC})	239.3	67.7	196.0	2895
Mannose (P_{LC})	3.1	2,360	1.6	0.7

[a] K_m values are calculated as polyP.

Table 2. PolyP-hexokinase activities of purified recombinant NsPPGK and NpPPGK

Hexose	Specific activity (μmol min^{-1} mg^{-1}) NsPPGK	NpPPGK
Glucose (5 mM)	229.1±14.0	174.3±14.7
Mannose (50 mM)	3.1±0.2	2.0±0.1
Fructose (50 mM)	ND [a]	ND

[a] ND, not detected; the minimum level of detection was ca. 0.05 nmol min^{-1} mg^{-1}.

Table 3. Effect of different compounds on the activity of NsPPGK towards glucose and P_{13-18} as substrates

Chemicals	Activity (%)
NaCl (1 mM)	100
NaCl (5 mM)	99.7
NaCl (10 mM)	99.5
NaCl (50 mM)	84.5

NaCl (150 mM)	66.0
KCl (1 mM)	95.8
KCl (5 mM)	82.3
KCl (10 mM)	70.4
KCl (50 mM)	65.2
KCl (150 mM)	49.9
ATP (1 mM)	96.3
ATP (5 mM)	71.0
ATP (50 mM)	12.2
PPi (1 mM)	99.8
PPi (5 mM)	55.6
PPi (10 mM)	28.4
P_3 (1 mM)	87.5
P_3 (5 mM)	39.9
P_3 (10 mM)	12.0
None	100

Specific activity in the absence of inhibitor (set as 100 %) was 31.8 μmol min^{-1} mg^{-1}.

Table 4. PPGK specific activities levels with different polyP substrates in whole-cell extracts from diverse *Cyanobacteria*

Cyanobacterial strain [a]	P_4	P_{13-18}	P_{60}	P_{LC}
	($nmol^{-1}min^{-1}$ mg^{-1})			
Synechococcus elongatus PCC7942(I)	ND [b]	ND	ND	ND
Synechocystis sp. PCC6803 (I)	ND	ND	ND	ND
Dermocarpa sp. PCC7437 (II)	9.0±0.7	8.4±0.7	8.5±0.7	10.9±1.3
Pseudanabaena sp. PCC6903 (III)	3.6±0.4	15.1±1.3	11.1±0.9	16.3±1.4
Anabaena sp. ATCC29413 (IV)	6.8±0.7	27.4±1.7	18.2±1.5	16.0±1.8
Calothrix sp. PCC9327 (IV)	0.1±0.3	2.1±0.2	3.1±0.2	4.6±0.3
Nostoc punctiforme PCC73102 (IV)	10.5±0.9	15.4±1.2	13.7±1,1	19.0±1,4
Nostoc sp. PCC7120 (IV)	0.8±0.1	17.9±1.3	18.4±1.5	10.9±0.9
Scytonema sp. PCC7110 (IV)	0.4±0.1	5.7±0.8	3.1±0.5	12.8±1.6
Chlorogloeopsis sp. PCC6912 (V) [c]	ND	ND	ND	ND
Fischerella muscicola UTEX1829 (V) [c]	ND	ND	ND	ND

[a] ATCC (American Type Culture Collection); PCC (Pasteur Culture Collection); UTEX (University of Texas at Austin Culture Collection). Roman numerals in parentheses indicate the sections of the taxonomic classification of Rippka et al. (1979). The two unicellular strains of section I were grown in the presence of 2 mM $NaNO_3$ as a nitrogen source.

[b] ND, not detected activity; the minimum level of detection is ca. 0.05 nmol min^{-1} mg^{-1}.

[c] The presence of large amounts of extracellular mucous material made very difficult achieving reliable measurements with these colonial strains.

Figure

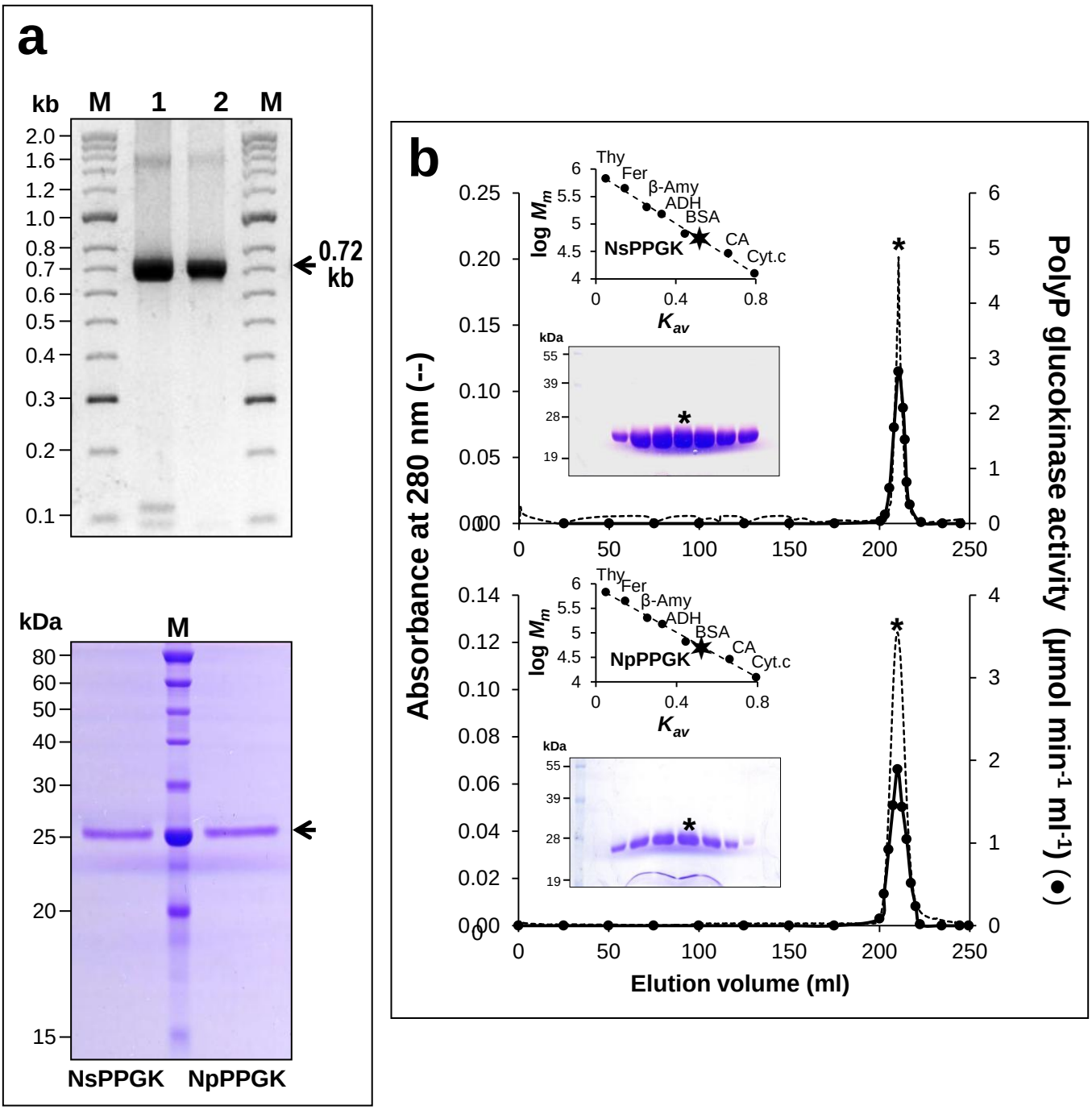
a
kb
M
1
2
M
2.0
1.6
1.2
1.0
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.72 kb
kDa
M
80
60
50
40
30
25
20
15
NsPPGK
NpPPGK
b
Absorbance at 280 nm (--)
PolyP glucokinase activity (µmol min-1 ml-1) (●)
Elution volume (ml)
log Mm
Kav
Thy
Fer
β-Amy
ADH
BSA
CA
Cyt.c
NsPPGK
NpPPGK
kDa
55
39
28
19

Fig. 1

Figure

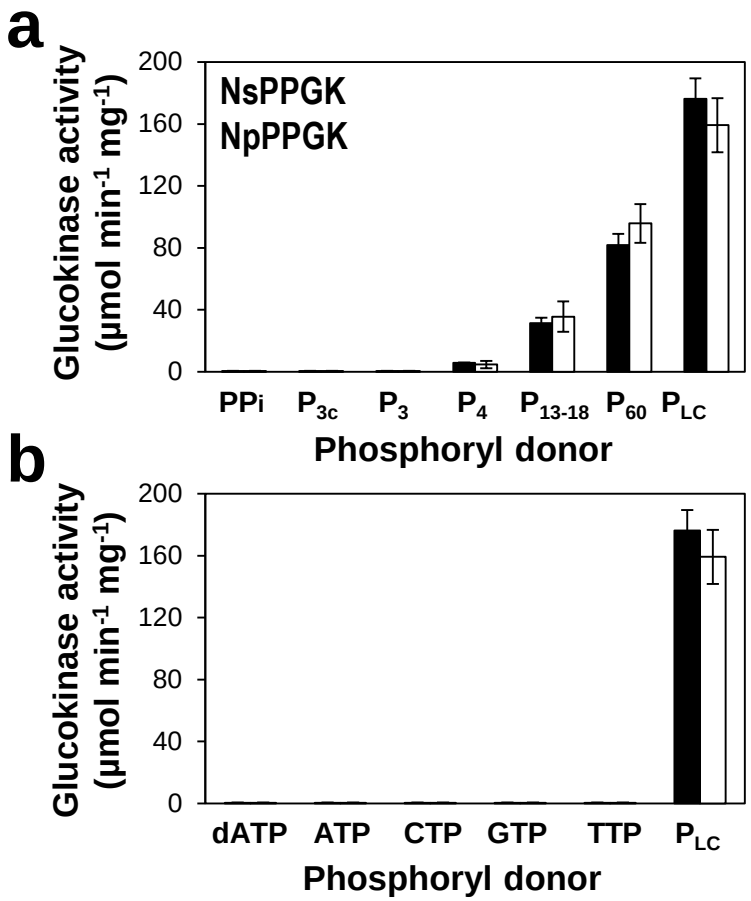
a
Glucokinase activity
(µmol min-1 mg-1)
200
160
120
80
40
0
NsPPGK
NpPPGK
PPi
P3c
P3
P4
P13-18
P60
PLC
Phosphoryl donor
b
Glucokinase activity
(µmol min-1 mg-1)
200
160
120
80
40
0
dATP
ATP
CTP
GTP
TTP
PLC
Phosphoryl donor

Fig. 2

Figure

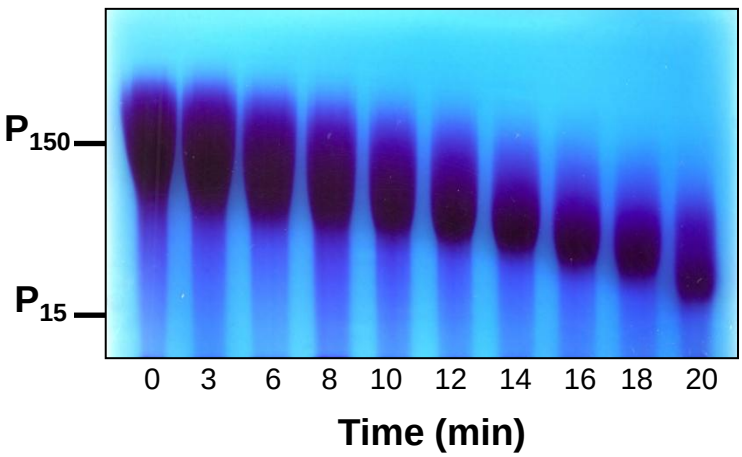

Fig. 3

Figure

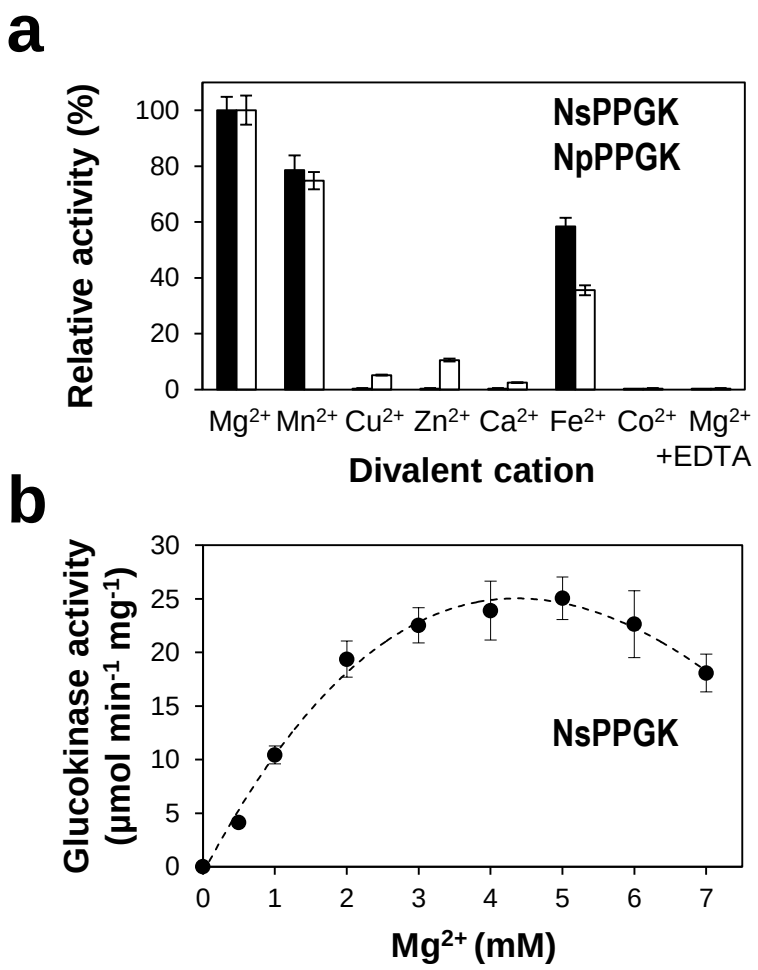
a
Relative activity (%)
100
80
60
40
20
0
NsPPGK
NpPPGK
Mg2+ Mn2+ Cu2+ Zn2+ Ca2+ Fe2+ Co2+ Mg2+ +EDTA
Divalent cation
b
Glucokinase activity (μmol min-1 mg-1)
30
25
20
15
10
5
0
0 1 2 3 4 5 6 7
NsPPGK
Mg2+ (mM)

Fig. 4

Figure

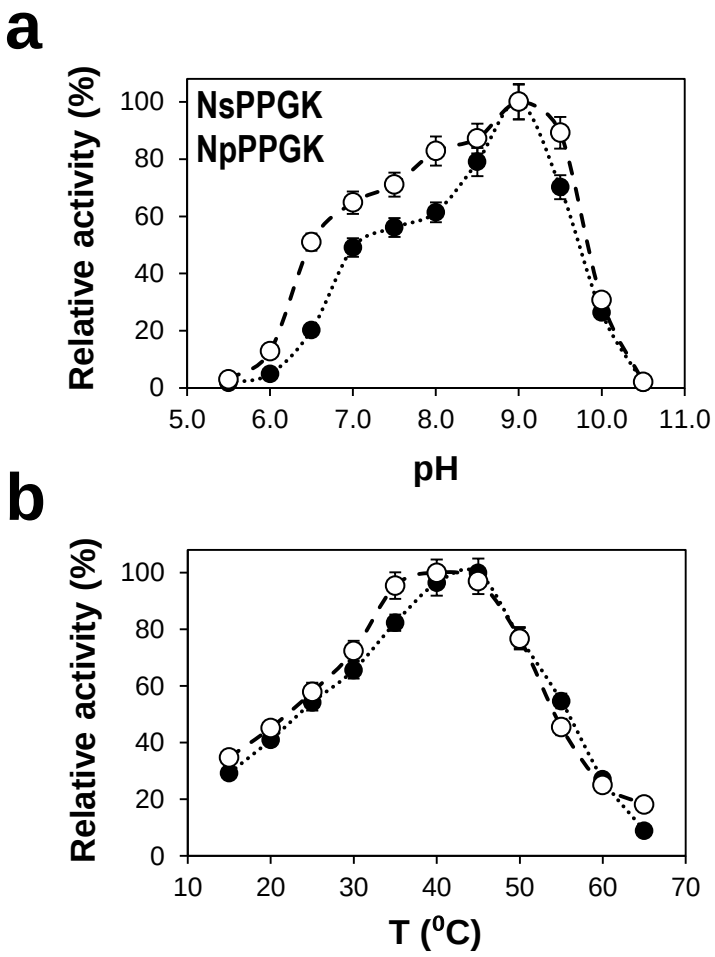
a
Relative activity (%)
NsPPGK
NpPPGK
100
80
60
40
20
0
5.0
6.0
7.0
8.0
9.0
10.0
11.0
pH
b
Relative activity (%)
100
80
60
40
20
0
10
20
30
40
50
60
70
T (⁰C)

Fig. 5

Figure

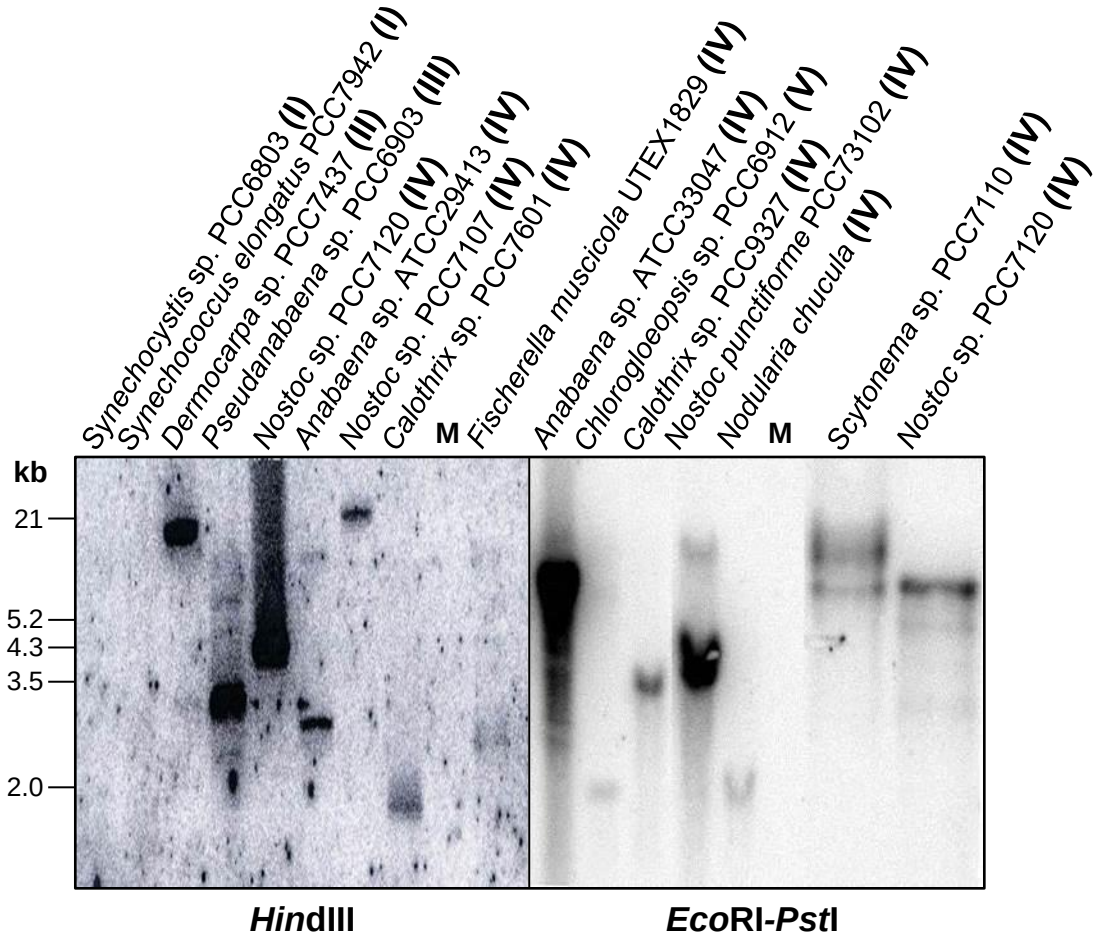
Synechocystis sp. PCC6803 (I)
Synechococcus elongatus PCC7942 (I)
Dermocarpa sp. PCC7437 (III)
Pseudanabaena sp. PCC6903 (III)
Nostoc sp. PCC7120 (IV)
Anabaena sp. ATCC29413 (IV)
Nostoc sp. PCC7107 (IV)
Calothrix sp. PCC7601 (IV)
M
Fischerella muscicola UTEX1829 (IV)
Anabaena sp. ATCC33047 (IV)
Chlorogloeopsis sp. PCC6912 (V)
Calothrix sp. PCC9327 (IV)
Nostoc punctiforme PCC73102 (IV)
Nodularia chucula (IV)
M
Scytonema sp. PCC7110 (IV)
Nostoc sp. PCC7120 (IV)
kb
21
5.2
4.3
3.5
2.0
HindIII
EcoRI-PstI

Fig. 6

Figure

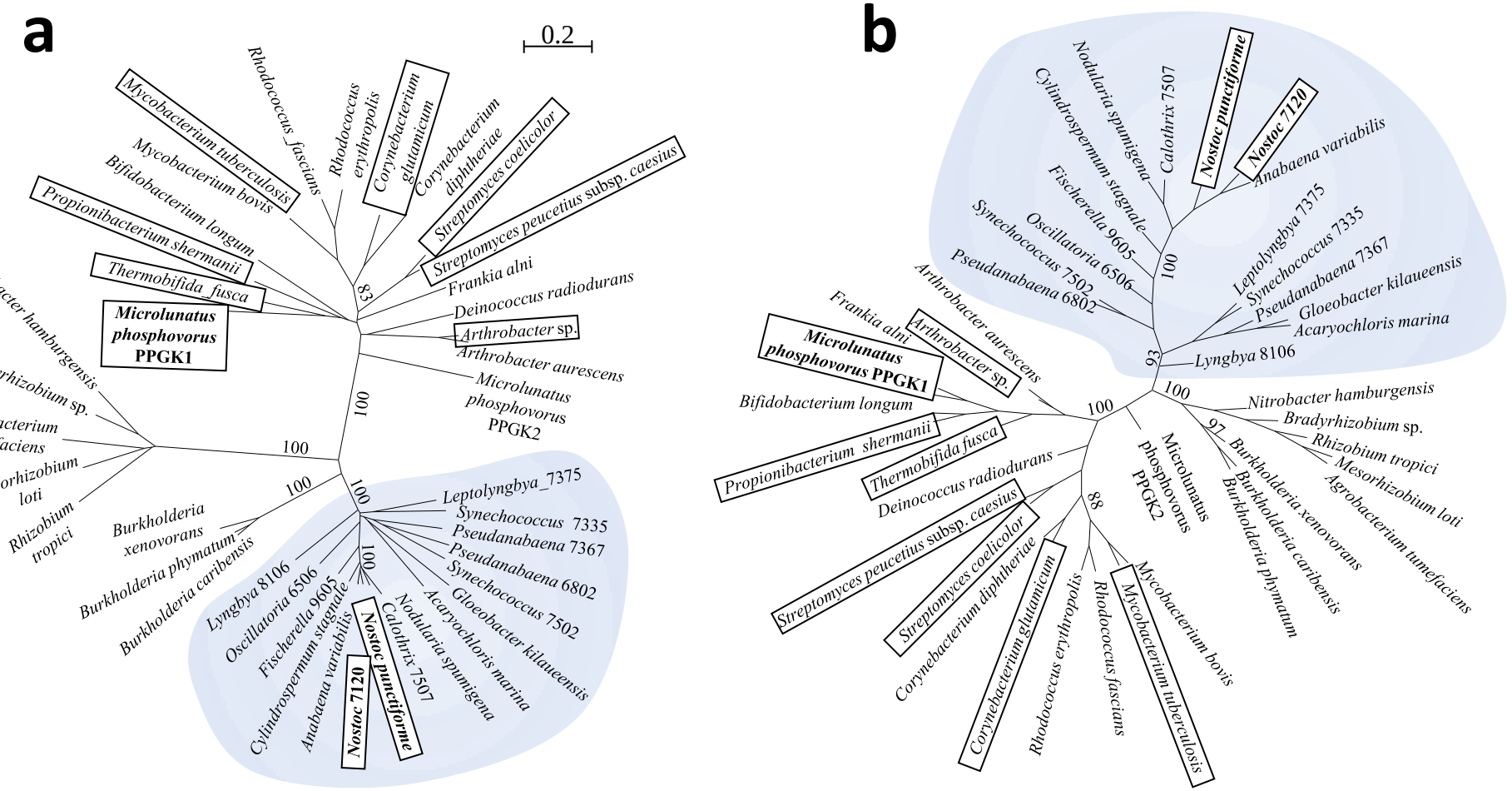
a
0.2
Mycobacterium tuberculosis
Rhodococcus_fascians
Rhodococcus erythropolis
Corynebacterium glutamicum
Corynebacterium diphtheriae
Streptomyces coelicolor
Streptomyces peucetius subsp. caesius
Mycobacterium bovis
Bifidobacterium longum
Propionibacterium shermanii
Thermobifida_fusca
Nitrobacter hamburgensis
Microlunatus phosphovorus PPGK1
Frankia alni
Deinococcus radiodurans
Arthrobacter sp.
Arthrobacter aurescens
Microlunatus phosphovorus PPGK2
Bradyrhizobium sp.
Agrobacterium tumefaciens
Mesorhizobium loti
Rhizobium tropici
Burkholderia xenovorans
Burkholderia phymatum
Burkholderia caribensis
Leptolyngbya_7375
Synechococcus 7335
Pseudanabaena 7367
Pseudanabaena 6802
Synechococcus 7502
Gloeobacter kilaueensis
Acaryochloris marina
Nodularia spumigena
Calothrix 7507
Nostoc punctiforme
Nostoc 7120
Anabaena variabilis
Cylindrospermum stagnale
Fischerella 9605
Oscillatoria 6506
Lyngbya 8106
83
100
100
100
100
b
Nodularia spumigena
Calothrix 7507
Nostoc punctiforme
Nostoc 7120
Anabaena variabilis
Cylindrospermum stagnale
Fischerella 9605
Oscillatoria 6506
Synechococcus 7502
Pseudanabaena 6802
Leptolyngbya 7375
Synechococcus 7335
Pseudanabaena 7367
Gloeobacter kilaueensis
Acaryochloris marina
Lyngbya 8106
100
93
Arthrobacter aurescens
Frankia alni
Arthrobacter sp.
Microlunatus phosphovorus PPGK1
Bifidobacterium longum
Propionibacterium shermanii
Thermobifida fusca
Deinococcus radiodurans
Streptomyces peucetius subsp. caesius
Streptomyces coelicolor
Corynebacterium diphtheriae
Corynebacterium glutamicum
Rhodococcus erythropolis
Rhodococcus fascians
Mycobacterium bovis
Mycobacterium tuberculosis
Microlunatus phosphovorus PPGK2
100
88
Nitrobacter hamburgensis
Bradyrhizobium sp.
97
Rhizobium tropici
Mesorhizobium loti
Agrobacterium tumefaciens
Burkholderia xenovorans
Burkholderia phymatum
Burkholderia caribensis

Fig. 7

Journal: Applied Microbiology and Biotechnology

Title:

Two strictly polyphosphate-dependent gluco(manno)kinases from diazotrophic *cyanobacteria* with potential to phosphorylate hexoses from polyphosphates

Authors: Tomás Albi · Aurelio Serrano

Instituto de Bioquímica Vegetal y Fotosíntesis, Centro de Investigaciones Científicas Isla Cartuja, CSIC-Universidad de Sevilla, Spain

Correspondence to: aurelio@ibvf.csic.es, +34 954 489 524

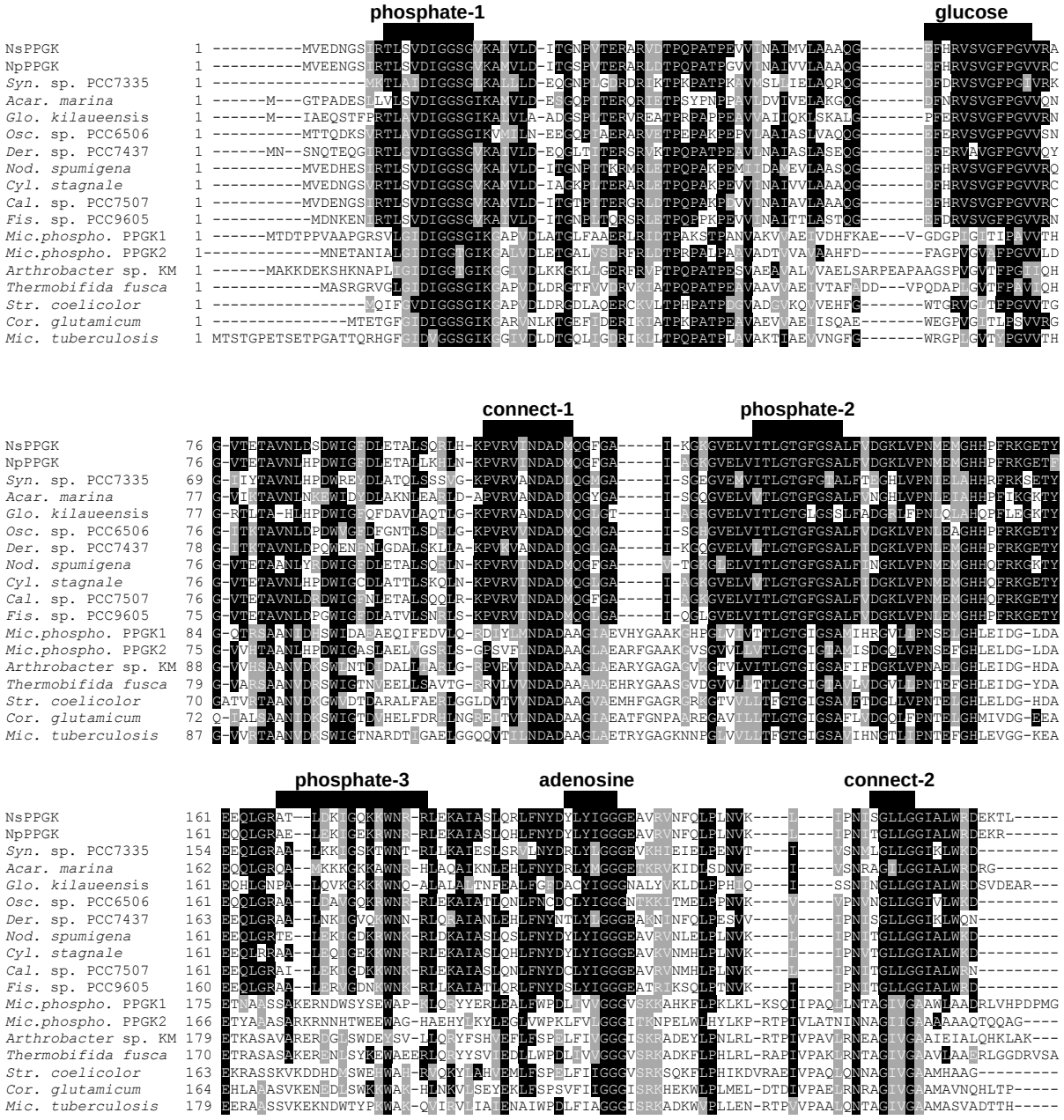

Fig. S1. Multiple sequences alignment and primary structure features of cyanobacterial PPGKs and other bacterial polyP/ATP-gluco(manno)kinases. The amino acid sequences (from top to bottom) of NsPPGK, NpPPGK, 9 putative cyanobacterial orthologs, and characterized homologs from the *Actinobacteria Microlunatus phosphovorous* NM-1 (the characterized strictly polyP-dependent PPGK1 and its uncharacterized paralog PPGK2), *Arthrobacter* sp. strain KM, *Thermobifida fusca* YX, *Streptomyces coelicolor* A3, *Corynebacterium glutamicum* and *Mycobacterium tuberculosis* H37Ra are shown (see Table S3 for details). Seven conserved motifs involved in binding of substrates - the so-called phosphate-1, glucose, phosphate-2, phosphate-3, connect-1, connect-2 and adenosine regions - are highlighted with black dashes. The specific heptapeptide in the glucose region of the poly(P)/ATP-glucomannokinase from *Arthrobacter* sp. strain KM is underlined. The sequence alignment was carried out using ClustalW (Larkin et al. 2007) and formatted with Boxshade version 3.21. Residues that are highly identical among the sequences are given a black background, and those that are similar among the sequences are given a gray background.

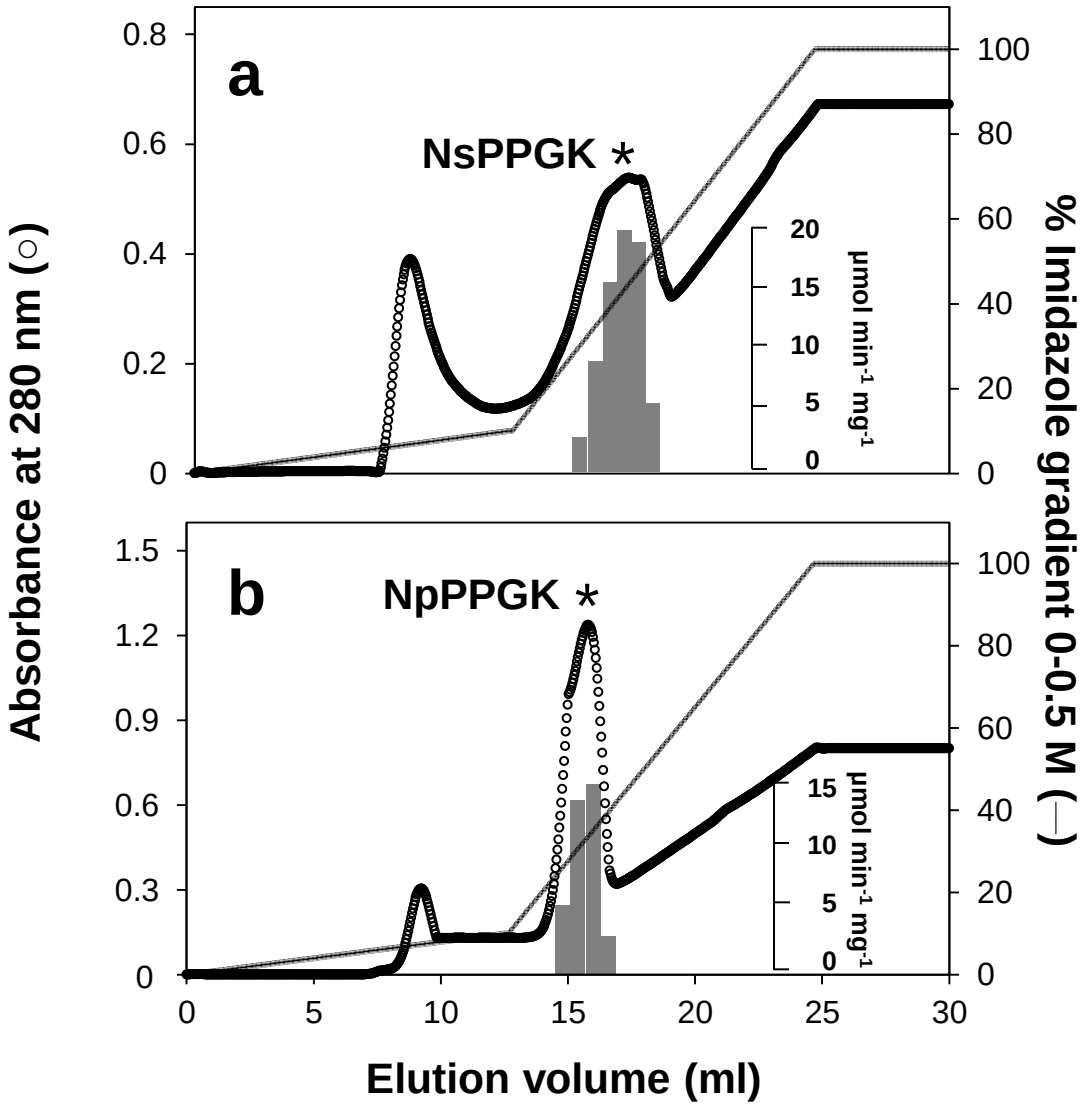

Fig. S2. Partial purification of recombinant cyanobacterial N_t-His PPGK by Ni-NTA metal-affinity chromatography. Protein elution profiles of the metal-chelate affinity chromatography of recombinant PPGKs from *Nostoc sp.* PCC 7120 (a) and *Nostoc punctiforme* PCC 73102 (b) are shown. Sonicated *E. coli* (BL21) cells were centrifuged and the crude supernatants containing the overproduced recombinant enzymes were loaded onto pre-equilibrated HisTrap FF Crude 1 mL Ni-NTA columns. Partially purified PPGKs were eluted with a linear gradient of imidazole with a target concentration of 500 mM. The asterisk indicates the fraction peak of recombinant protein as determined by its polyphosphate-glucokinase activity (grey bars) and absorbance at 280 nm.

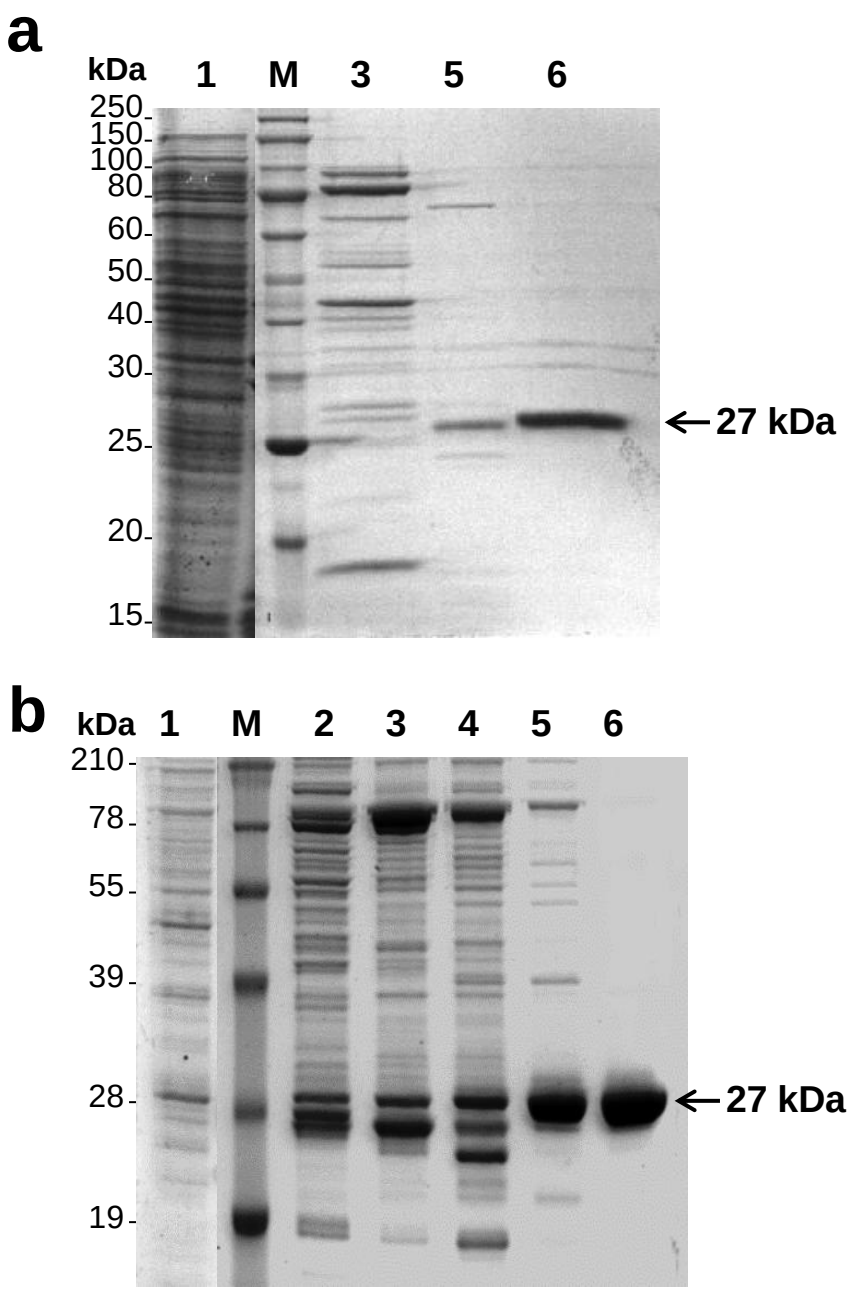

Fig. S3. Purification profiles of recombinant NsPPGK (a) and NpPPGK (b) preparations analyzed by Coomassie blue staining of SDS-PAGE 12% (w/v) gels. *Lane 1*: PPGK-overexpressing *E. coli* (BL21) crude extracts; *lanes 2-4,* fractions eluted at low concentration of imidazole with a linear gradient (20-50 mM); *lane 5*, partially purified PPGK preparation from HisTrap FF Crude column eluted at high concentration of imidazole (ca. 0.25 M); *lane 6,* purified PPGK after FPLC gel filtration. Numerals on the left indicate the molecular masses (kDa) of protein markers (M). Arrows indicate the bands of the recombinant PPGK proteins.

MATRIX SCIENCE **Mascot Search Results**

Probability Based Mowse Score

Match to: **gi|81772382**

polyphosphate glucokinase – [*Nostoc* sp. PCC 7120] –> NsPPGK

Sequence coverage of natural protein: **82 %**

Nominal mass (M_m): **25,919 (without the N-terminal tag of 12 aa)**

```
  1 mrgshhhhhh gsMVEDNGSI RTLSVDIGGS GVKALVLDIT GNPVTERARV
 51 DTPQPATPEV VINAIMVLAA AQGEFHRVSV GFPGVVRAGV TETAVNLDSD
101 WIGFDLETAL SQRLHKPVRV INDADMQGFG AIKGKGVELV ITLGTGFGSA
151 LFVDGKLVPN MEMGHHPFRK GETYEEQLGR ATLDKIGQKK WNRRLEKAIA
201 SLQRLFNYDY LYIGGGEAVR VNFQLPLNVK LIPNISGLLG GIALWRDEKT
251 L
```

MATRIX SCIENCE **Mascot Search Results**

Probability Based Mowse Score

Match to: **gi|23125685**

Transcriptional regulator/sugar kinase –[*Nostoc punctiforme* PCC 73102] –> NpPPGK

Sequence coverage of natural protein: **55 %**

Nominal mass (M_m): **25,816 (without the N-terminal tag of 12 aa)**

```
  1 mrgshhhhhh gsMVEENGSI RTLSVDIGGS GVKAMVLDIT GSPVTERARL
 51 DTPQPATPGV VINAIVVLAA AQGEFHRVSV GFPGVVRCGV TETAVNLHPD
101 WIGFDLETAL LKHLNKPVRV INDADMQGFG AIAGKGVELV ITLGTGFGSA
151 LFVDGKLVPN MEMGHHPFRK GETFEQQLGR AELEKIGEKR WNRRLEKAIA
201 SLQHLFNYDY LYIGGGEAVR VNFQLPLNVK LIPNITGLLG GIALWRDEKR
```

Fig. S4. Sequence and domain structure validation of cyanobacterial PPGKs by tryptic-peptide fingerprinting and MALDI-TOF mass spectrometry analysis. The Pfam domain structures of the two natural PPGKs are shown, as well as the sequences of the corresponding purified recombinant proteins in which the amino acid residues are bold-coloured accordingly, the experimentally identified peptides are underlined and the N-terminal His-tags are in lowercase. Identified peptides cover about 80 and 55 % of the predicted protein sequences of natural NsPPGK and NpPPGK, respectively.

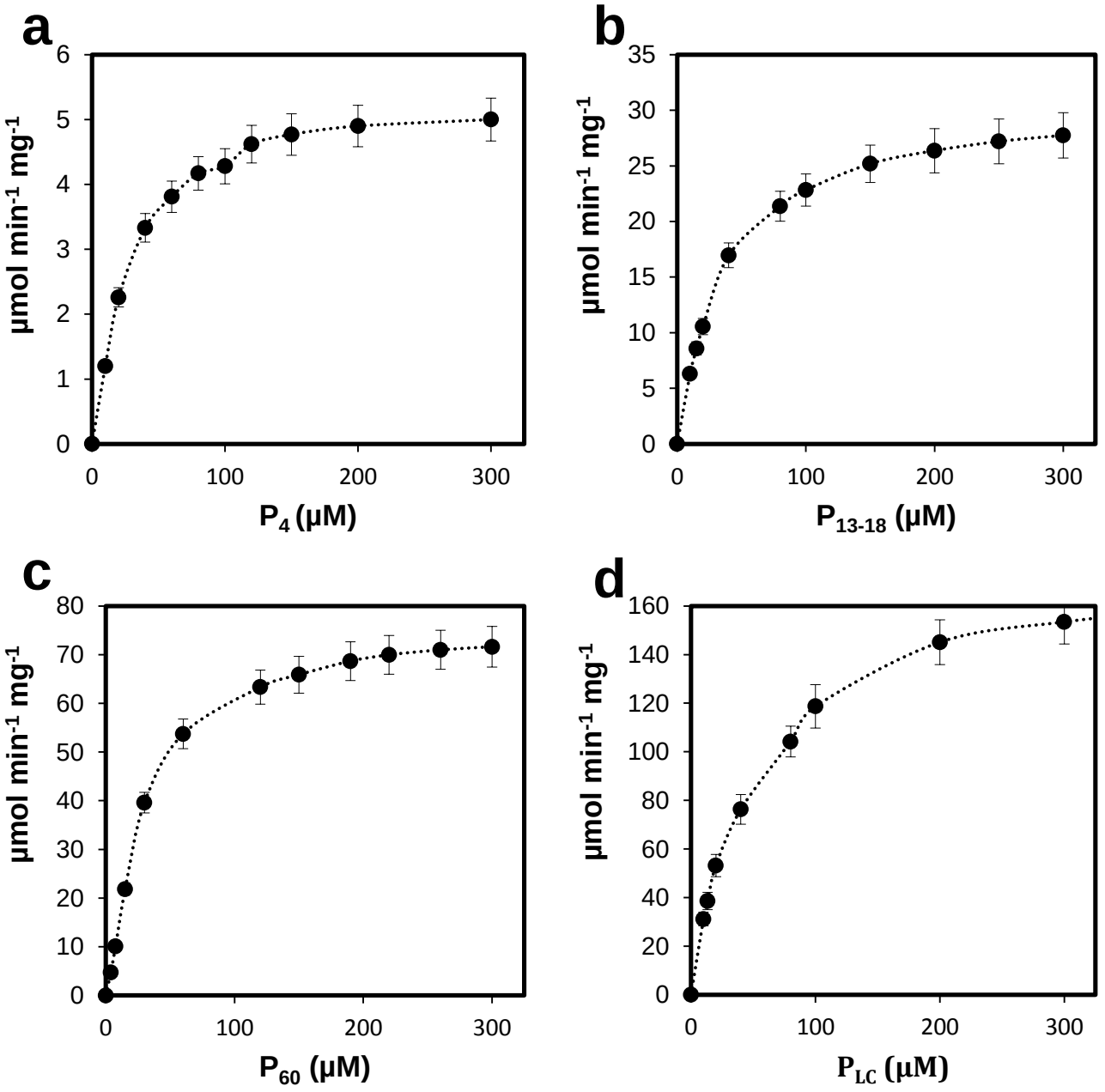

Fig. S5. Determination of the kinetic parameters of the recombinant NsPPGK. Increasing concentrations of polyPs of different chain lengths were used in the range up to 300 µM to phosphorylate glucose (5 mM). Concentration dependence curves of NsPPGK activity with P_4 (a), $P_{13\text{-}18}$ (b), P_{60} (c), and P_{LC} (d) as phosphoryl donor substrates are shown. Enzymatic reactions were carried out at optimal conditions as described in the Materials and Methods section. Points represent mean activity values $\pm$ S.E. of three independent determinations. Kinetic parameters were determined by nonlinear curve fitting from the Michaelis-Menten plot using the spreadsheet Anemona.xlt (Hernández *et al.* 1998).

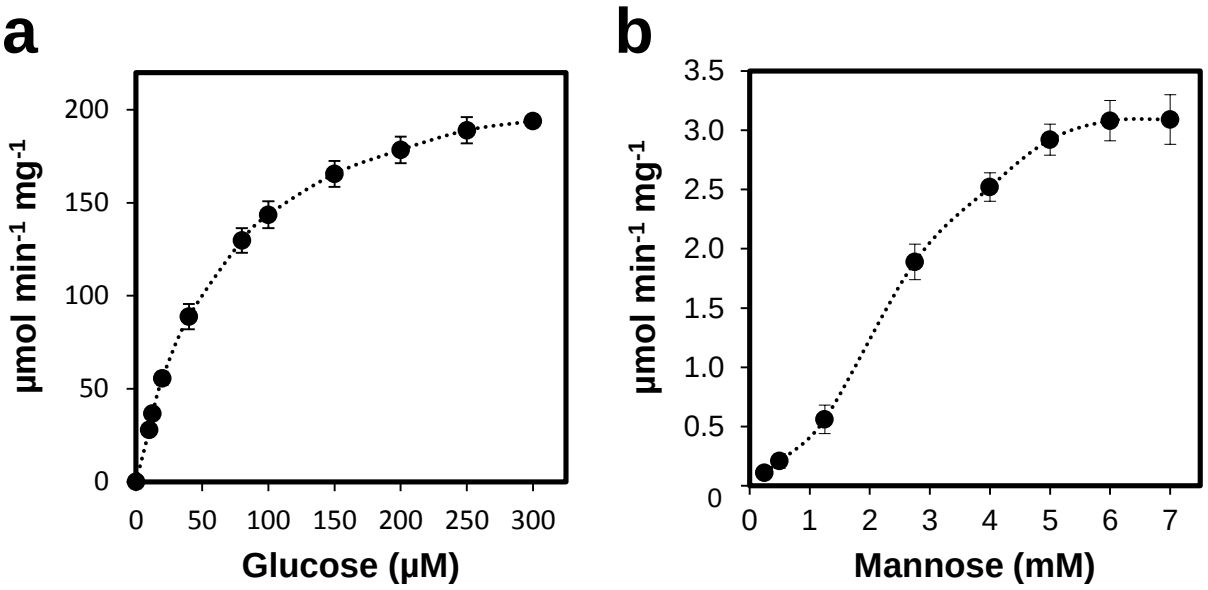

Fig. S6. Determination of the kinetic parameters of the recombinant NpPPGK with different monosaccharide substrates. Concentration dependence phosphorylation activity curves with glucose (a) and mannose (b) of purified NpPPGK, using 1 mM P_{LC} as phosphoryl donor, are shown. Enzymatic reactions were carried out at optimal conditions as described in the Materials and Methods section. Points represent mean activity values ± S.E. of three independent determinations. Kinetic parameters were determined by nonlinear curve fitting from the Michaelis-Menten plot using the spreadsheet Anemona.xlt (Hernández *et al.* 1998).

Online Resource Table S1. List of cyanobacterial strains used in this work

Strain [a]	Description, alternative names	N_2-fixation conditions	Group [b]
Synechococcus elongatus PCC7942	Unicellular rod-shaped, also called *Anacystis nidulans*	No	I
Synechocystis sp. PCC6803	Unicellular globular-shaped	No	I
Dermocarpa sp. PCC7437	Colonial, also called *Stanieria cyanosphaera*	Microaerobic	II
Pseudanabaena sp. PCC6903	Filamentous, non-heterocystous	Microaerobic	III
Anabaena sp. ATCC29413	Filamentous, heterocystous, also called *Anabaena variabilis*	Aerobic	IV
Anabaena sp. ATCC33047	Filamentous, heterocystous	Aerobic	IV
Calothrix sp. PCC7601	Filamentous, heterocystous, also called *Fremyella diplosiphon*	Unable to fix N_2 (mutant strain)	IV
Calothrix sp. PCC9327	Filamentous, heterocystous, also called *Tolypothrix* sp.	Aerobic	IV
Nodularia chucula	Filamentous, heterocystous	Aerobic	IV
Nostoc sp. PCC6719	Filamentous, heterocystous	Aerobic	IV
Nostoc sp. PCC7107	Filamentous, heterocystous	Aerobic	IV
Nostoc sp. PCC7120	Filamentous, heterocystous, also called *Anabaena*	Aerobic	IV
Nostoc punctiforme PCC73102	Filamentous, heterocystous	Aerobic	IV
Scytonema sp. PCC7110	Filamentous, heterocystous	Aerobic	IV
Chlorogloeopsis sp. PCC6912	Branched filamentous, heterocystous, also called *Chlorogloeopsis fritschii*	Aerobic	V
Fischerella muscicola UTEX1829	Filamentous, heterocystous	Aerobic	V

[a] Culture Collection abbreviations are as described in Table 4.

[b] With reference to the taxonomic classification of Rippka et al. (1979).

Online Resource Table S2. Primers for cloning the *ppgK* genes from *Nostoc* sp. PCC7120 and *Nostoc punctiforme* PCC73102

Gene		Primers (new restriction site, underlined)
ppgK (N.7120)	F (*Bam*HI)	5'-GCCGGATCCATGGTGGAAGATAACGGC-3'
ppgK (N.7120)	R (*Pst*I)	5'-TCACTGCAGCTATAGTGTTTTTTCATC-3'
ppgK (N.73102)	F (*Bam*HI)	5'-GCCGGATCCATGGTTGAAGAAAATGGATCG-3'
ppgK (N.73102)	R (*Pst*I)	5'-TCACTGCAGTTACCTTTTTTCATCTCGCC-3'

Online Resource Table S3. PPGK sequences used in this study

Source organism	Protein or Gene ID entry [a]	Predicted protein length (aa)	Taxonomy [b]
Synechococcus sp. PCC7335*	B4WQL2	228	*Cyanobacteria*, *Chroococcales* (I)
Synechococcus sp. PCC7502*	K9SS42	238	*Cyanobacteria*, *Chroococcales* (I)
Gloeobacter kilaueensis JS1*	U5QNN6	241	*Cyanobacteria*, *Gloeobacterales* (I)
Pleurocapsa sp. PCC7319	2509711744	235	*Cyanobacteria*, *Pleurocapsales* (II)
Dermocarpa sp. PCC7437	2503800427	237	*Cyanobacteria*, *Pleurocapsales* (II)
Acaryochloris marina MBIC11017*	B0C5T1	238	*Cyanobacteria*, *Chroococcales* (II)
Acaryochloris sp. CCMEE5410	2514738042	238	*Cyanobacteria*, *Chroococcales* (II)
Oscillatoriales sp. JSC-12	2510096246	230	*Cyanobacteria*, *Oscillatoriales* (III)
Pseudanabaena biceps PCC7429	L8N8S2	234	*Cyanobacteria*, *Oscillatoriales* (III)
Pseudanabaena sp. PCC6802*	2506783054	234	*Cyanobacteria*, *Oscillatoriales* (III)
Coleofasciculus chthonoplastes PCC7420	B4W415	235	*Cyanobacteria*, *Oscillatoriales* (III)
Crinalium epipsammum PCC9333	2504685141	235	*Cyanobacteria*, *Oscillatoriales* (III)
Lyngbya majuscula 3L	2506483678	235	*Cyanobacteria*, *Oscillatoriales* (III)
Microcoleus chthonoplaste PCC7420	647572171	235	*Cyanobacteria*, *Oscillatoriales* (III)
Microcoleus sp. PCC7113	2509437012	235	*Cyanobacteria*, *Oscillatoriales* (III)
Microcoleus vaginatus PCC9802	2505167359	235	*Cyanobacteria*, *Oscillatoriales* (III)
Moorea producens 3L	F4Y1Y5	235	*Cyanobacteria*, *Oscillatoriales* (III)
Oscillatoria sp. PCC6407	2508875670	235	*Cyanobacteria*, *Oscillatoriales* (III)
Oscillatoria sp. PCC6506*	D8FXR8	235	*Cyanobacteria*, *Oscillatoriales* (III)
Oscillatoria sp. PCC7112	2504089037	235	*Cyanobacteria*, *Oscillatoriales* (III)
Oscillatoria acuminata PCC6304	2509419241	236	*Cyanobacteria*, *Oscillatoriales* (III)
Pseudanabaena sp. PCC7367*	K9SLD8	237	*Cyanobacteria*, *Oscillatoriales* (III)
Lyngbya aestuarii BL J	2578016402	238	*Cyanobacteria*, *Oscillatoriales* (III)
Lyngbya sp. PCC8106*	A0YLE5	238	*Cyanobacteria*, *Oscillatoriales* (III)
Oscillatoria sp. PCC10802	2509508264	239	*Cyanobacteria*, *Oscillatoriales* (III)
Leptolyngbya sp. 2LT21S03	2509751146	243	*Cyanobacteria*, *Oscillatoriales* (III)
Leptolyngbya sp. PCC7375*	K9EMV6	251	*Cyanobacteria*, *Oscillatoriales* (III)
Leptolyngbya sp. Heron Island J	2579003743	255	*Cyanobacteria*, *Oscillatoriales* (III)
Calothrix sp. PCC6303	K9V545	234	*Cyanobacteria*, *Microchaetaceae* (IV)
Calothrix desertica PCC7102	2510030452	235	*Cyanobacteria*, *Microchaetaceae* (IV)
Calothrix sp. PCC7103	WP_019497362	235	*Cyanobacteria*, *Microchaetaceae* (IV)
Calothrix sp. PCC7507*	K9PMJ8	235	*Cyanobacteria*, *Microchaetaceae* (IV)
Fremyella diplosiphon UTEX481	2501543286	235	*Cyanobacteria*, *Microchaetaceae* (IV)
Microchaete sp. PCC7126	2509783891	235	*Cyanobacteria*, *Microchaetaceae* (IV)
Raphidiopsis brookii D9	D4TRI7	232	*Cyanobacteria*, *Nostocaceae* (IV)
Anabaena sp. PCC7108	2506493476	235	*Cyanobacteria*, *Nostocaceae* (IV)
Anabaena cylindrica PCC7122	2504134473	235	*Cyanobacteria*, *Nostocaceae* (IV)
Cylindrospermum stagnale PCC7417*	K9WXY9	235	*Cyanobacteria*, *Nostocaceae* (IV)
Nodularia spumigena CCY9414*	A0ZFN0	235	*Cyanobacteria*, *Nostocaceae* (IV)
Nostoc azollae (strain 0708)	D7E4T8	235	*Cyanobacteria*, *Nostocaceae* (IV)

Nostoc sp. PCC7107	K9Q5V7	235	*Cyanobacteria*, *Nostocaceae* (IV)
Scytonema hofmanni PCC7110	2551958472	236	*Cyanobacteria*, *Scytonemataceae* (IV)
Nostoc sp. PCC7524	2509810004	237	*Cyanobacteria*, *Nostocaceae* (IV)
Nostoc punctiforme PCC73102*	B2J3R4	238 (**NpPPGK**)[c]	*Cyanobacteria*, *Nostocaceae* (IV)
Anabaena variabilis ATCC29413*	Q3M5W7	239	*Cyanobacteria*, *Nostocaceae* (IV)
Nostoc sp. PCC7120*	Q8YX46	239 (**NsPPGK**)[c]	*Cyanobacteria*, *Nostocaceae* (IV)
Fischerella muscicola SAG1427-1	2550703822	234	*Cyanobacteria*, *Stigonematales* (V)
Fischerella sp. PCC9605*	WP_026733157	234	*Cyanobacteria*, *Stigonematales* (V)
Mastigocladopsis repens MORA, PCC10914	2517243485	235	*Cyanobacteria*, *Stigonematales* (V)
Fischerella sp. PCC9431	WP_026723397	235	*Cyanobacteria*, *Stigonematales* (V)
*Bifidobacterium longum**	D6DBE0	255	*Actinobacteria*, *Bifidobacteriaceae*
*Corynebacterium glutamicum**	Q6M4B1	250	*Actinobacteria*, *Corynebacteriaceae*
*Corynebacterium diphtheriae**	Q6NGU6	253	*Actinobacteria*, *Corynebacteriaceae*
*Frankia alni**	Q0RE01	289	*Actinobacteria, Frankiaceae*
Arthrobacter sp. KM*	A0JVB2	267	*Actinobacteria, Micrococcaceae*
*Arthrobacter aurescens**	A1R5H5	272	*Actinobacteria, Micrococcaceae*
*Mycobacterium bovis**	C1AFG1	265	*Actinobacteria, Mycobacteriaceae*
*Mycobacterium tuberculosis**	P9WIN1	265	*Actinobacteria, Mycobacteriaceae*
*Rhodococcus erythropolis**	C0ZYU0	273	*Actinobacteria, Nocardiaceae*
*Rhodococcus fascians**	Q8VM93	274	*Actinobacteria, Nocardiaceae*
Thermobifida fusca TM51	R9F6L8	262	*Actinobacteria, Nocardiopsaceae*
Thermobifida fusca YX*	Q47NX5	262	*Actinobacteria, Nocardiopsaceae*
Microlunatus phosphovorus NM-1*	F5XI06	253 (PPGK2)	*Actinobacteria, Propionibacteriaceae*
*Propionibacterium shermanii**	D7GI59	261	*Actinobacteria, Propionibacteriaceae*
Microlunatus phosphovorus NM-1*	F5XK61	266 (**PPGK1**)[c]	*Actinobacteria, Propionibacteriaceae*
Streptomyces peutecius subsp. *caesius**	S5DRF7	242	*Actinobacteria, Streptomycetaceae*
Streptomyces coelicolor A3(2)*	Q9ADE8	246	*Actinobacteria, Streptomycetaceae*
Candidatus Poribacteria sp. WGA-4E [d]	2265139082	231	*Poribacteria*
Agrobacterium tumefaciens C58*	A9CH74	225	α-*Proteobacteria*
Nitrobacter hamburgensis X14*	Q1QLK2	229	α-*Proteobacteria*
Mesorhizobium loti MAFF303099*	Q98EJ9	240	α-*Proteobacteria*
Bradyrhizobium sp. DFCI-1*	U1H9M5	244	α-*Proteobacteria*
*Rhizobium tropici**	L0LZW1	250	α-*Proteobacteria*
Burkholderia xenovorans LB400*	Q13JL2	266	β-*Proteobacteria*
Burkholderia caribensis MBA4*	W4NCD7	269	β-*Proteobacteria*
Burkholderia phymatum DSM17167 / STM815*	B2JFU9	270	β-*Proteobacteria*
Cystobacter fuscus DSM2262 [d]	2538040499	255	δ-*Proteobacteria*
*Deinococcus radiodurans**	Q9RW46	279	*Thermus*/*Deinococcus* group, *Deinococcaceae*

[a] Sequences (mostly putatives) are referred to their corresponding UniProtKB, GeneBank or IMG-JGI databases entry codes. Those sequences used for the molecular phylogenetic analyses are indicated with an asterisk.

[b] Taxonomy was established following both general bacteriological rules and the specific rules for *Cyanobacteria* (taxonomic sections denoted by roman numbers in parenthesis) of Rippka et al. (1979).

[c] Biochemically characterized strictly polyP-dependent PPGKs.

[d] These two sequences of non-photosynthetic bacteria cluster into the cyanobacterial assembly and are not included in the trees for clarity purposes; they may be the results of horizontal gene transfer events.

Online Resource Table S4. Purification of His-tagged NsPPGK from transformed *E. coli* cells

Step	Specific activity (μmol min^{-1} mg^{-1})	Protein (mg)	Recovery (%)	Purification (folds)
Crude supernatant	0.16	285.0	100	1
Ni-NTA	14.57	1.4	45	91
Amicon Ultra-3 kDa ultrafiltration	14.12	1.3	40	--
Superdex 200	31.40	0.5	28	196