Structural Analysis of ADP-Glucose Pyrophosphorylase from the Bacterium *Agrobacterium tumefaciens*[†,‡]

Jill R. Cupp-Vickery,[*,§] Robert Y. Igarashi,[||,⊥] Marco Perez,[||] Myesha Poland,[||] and Christopher R. Meyer[*,||]

Department of Physiology and Biophysics, University of California, Irvine, California 92697, and Department of Chemistry and Biochemistry, California State University, Fullerton, California 92834

Received September 24, 2007; Revised Manuscript Received February 14, 2008

ABSTRACT: ADP-glucose pyrophosphorylase (ADPGlc PPase) catalyzes the conversion of glucose 1-phosphate and ATP to ADP-glucose and pyrophosphate. As a key step in glucan synthesis, the ADPGlc PPases are highly regulated by allosteric activators and inhibitors in accord with the carbon metabolism pathways of the organism. Crystals of *Agrobacterium tumefaciens* ADPGlc PPase were obtained using lithium sulfate as a precipitant. A complete anomalous selenomethionyl derivative X-ray diffraction data set was collected with unit cell dimensions $a = 85.38$ Å, $b = 93.79$ Å, and $c = 140.29$ Å ($\alpha = \beta = \gamma = 90°$) and space group I_{222}. The *A. tumefaciens* ADPGlc PPase model was refined to 2.1 Å with an $R_{factor} = 22\%$ and $R_{free} = 26.6\%$. The model consists of two domains: an N-terminal $\alpha\beta\alpha$ sandwich and a C-terminal parallel β-helix. ATP and glucose 1-phosphate were successfully modeled in the proposed active site, and site-directed mutagenesis of conserved glycines in this region (G20, G21, and G23) resulted in substantial loss of activity. The interface between the N- and the C-terminal domains harbors a strong sulfate-binding site, and kinetic studies revealed that sulfate is a competitive inhibitor for the allosteric activator fructose 6-phosphate. These results suggest that the interface between the N- and C-terminal domains binds the allosteric regulator, and fructose 6-phosphate was modeled into this region. The *A. tumefaciens* ADPGlc PPase/fructose 6-phosphate structural model along with sequence alignment analysis was used to design mutagenesis experiments to expand the activator specificity to include fructose 1,6-bisphosphate. The H379R and H379K enzymes were found to be activated by fructose 1,6-bisphosphate.

Glycogen and starch serve as important energy and carbon storage compounds for nearly all living organisms. In bacteria and plants, ADP-glucose pyrophosphorylase (ADPGlc PPase,[1] *glgC* gene product; EC 2.7.7.27) catalyzes the rate-limiting step of the glucan biosynthesis pathway by converting glucose 1-phosphate (Glc1P) and ATP to ADP-glucose and pyrophosphate (*1, 2*). The product ADP-glucose serves as the glucosyl donor for glucan elongation and branching via the actions of glycogen/starch synthase(s) (EC 2.4.1.21) and branching enzyme (EC 2.4.1.18) (*1, 2*). Most bacterial ADPGlc PPases function as homotetramers of an ~50 kDa subunit while plant enzymes are active as heterotetramers of $\alpha_2\beta_2$ composition (*1, 2*).

As a key enzyme in a pathway generating renewable and biodegradable carbon, ADPGlc PPase is an attractive target for protein engineering. The regulation of ADPGlc PPase is mediated by various metabolites serving as allosteric activators and inhibitors in accord with the carbon utilization pathways of the organisms (*1, 2*). These metabolites act together to provide a sensitive and powerful feed-forward and feedback regulation of these enzymes. The ADPGlc PPases have recently been classified into nine distinct regulatory groups (*1*). The bacterial enzymes from four of these groups, represented by the *Agrobacterium tumefaciens*, *Escherichia coli*, *Rhodobacter sphaeroides*, and the cyanobacteria *Anabaena* enzymes, have been studied in molecular detail. The *A. tumefaciens* enzyme is activated by fructose 6-phosphate (F6P) and pyruvate and inhibited by phosphate while the *E. coli* enzyme is activated by fructose 1,6-bisphosphate (FBP) and inhibited by AMP. The enzyme from *Rb. sphaeroides* is activated by FBP, F6P, and pyruvate and inhibited by phosphate. The *Anabaena* enzyme, in common with higher plant enzymes, is activated by 3-phosphoglycerate (3PGA) and inhibited by phosphate. These effectors are anionic in nature with negative charge contributed by phosphate and/or carboxyl groups. The anion sulfate has also been shown to be an inhibitor of several ADPGlc PPases, including the *A. tumefaciens* and potato tuber enzymes (*3, 4*).

The cloning of various *glgC* genes (*5–8*) has led to numerous mutagenesis studies to determine the location of the active and allosteric sites. Site-directed mutagenesis studies of the *E. coli* enzyme have provided evidence for the roles of Lys195, Tyr114, and Asp142 in the active site (*9–11*). The role of the conserved Lys195 position in

[†] This work was supported by National Science Foundation grants (0131465 to J.R.C.-V. and 0131729 and 0448676 to C.R.M.).

[‡] PDB accession code 3BRK.

* To whom correspondence should be addressed. J.R.C.-V.: tel, (949) 824-9338; fax, (949) 824-8540; e-mail, jvickery@uci.edu. C.R.M.: tel, (714) 278-4173; fax, (714) 278-5316; e-mail, cmeyer@fullerton.edu.

[§] University of California, Irvine.

[||] California State University, Fullerton.

[⊥] Current address: Department of Chemistry, University of Central Florida, Orlando, FL 32816-2366.

[1] Abbreviations: ADPGlc PPase, ADP-glucose pyrophosphorylase; DTE, dithioerythreitol; DTT, dithiothreitol; FBP, fructose 1,6-bisphosphate; F6P, fructose 6-phosphate; Glc1P, glucose 1-phosphate; 3PGA, 3-phosphoglycerate; rmsd, root mean squared deviation.

10.1021/bi701933q CCC: $40.75
Published on Web 03/21/2008

apparent binding of the substrate Glc1P was reinforced by a study with the potato tuber enzyme (*12*). Similarly, the conserved Asp142 position in the potato tuber enzyme was shown to be important for catalysis (*13*). Mutagenesis of Arg22 and Arg25 to alanine (part of a conserved glycine-rich region in the N-terminus) of the *A. tumefaciens* enzyme also caused decreased activity (*3*). These results were in agreement with the known active site domains from the same superfamily of sugar nucleotidyl transferases that includes the *E. coli N*-acetylglucosamine-1-phosphate uridyltransferase (GlmU) (*14, 15*) and the *Pseudomonas aeruginosa* and *E. coli* glucose-1-phosphate thymidylyltransferases (*16, 17*).

In contrast to the active site, the location of the allosteric site(s) among the ADPGlc PPases is less certain. Mutagenesis of the *E. coli* enzyme indicated that Lys39 is involved in apparent FBP binding (*18*) while mutation of conserved N-terminal arginine residues of the *A. tumefaciens* enzyme resulted in desensitization to the activator F6P (*3*). Interestingly, analysis of chimeric enzymes comprised of the *E. coli* and *A. tumefaciens* enzymes supported a role for the C-terminus in FBP activation (*19*). Two C-terminal lysines of the *Anabaena* enzyme have been implicated in activation by 3PGA (*20, 21*). Analogous lysine residues have also been identified in the potato tuber enzyme, which also play a role in 3PGA activation (*22*). Further, the techniques of random mutagenesis (*23*) and directed evolution (*24*) have revealed a number of other residues in both the N- and C-terminus important for regulation for the potato tuber ADPGlc PPase subunits.

It is difficult to interpret the findings of these mutagenesis experiments with diverse ADPGlc PPases in the absence of three-dimensional structures from both the plant and bacterial enzyme families. The three-dimensional structure of the α-subunit of the potato tuber ADPGlc PPase was recently reported (*4*). The potato tuber α-subunit X-ray model revealed the enzyme to be comprised of an N-terminal domain (with an α/β nucleotide diphospho-sugar transferase fold) and a C-terminal domain adopting a left-handed parallel β-helix. The N-terminal domain contains the previously identified active site residues. Geiger and co-workers hypothesized that the allosteric site(s) of the potato tuber enzyme involved interactions between the N- and C-terminal domains (*4*). However, the α_4 potato tuber enzyme used for crystallography studies displays altered regulatory properties compared to the native $\alpha_2\beta_2$ enzyme, with decreased sensitivity to 3PGA activation and increased sensitivity to phosphate inhibition (*25*).

To date, the bacterial ADPGlc PPases have not been amenable to X-ray crystallography, and a high-resolution atomic structure has not been available. Even with the potato tuber structure, many questions remain about structure–function relationships in the ADPGlc PPase family, particularly with regard to allosteric regulation. Accurate modeling of the allosteric site(s) of the bacterial enzymes is problematic due to the different allosteric specificity of the bacterial enzymes as well as a general lack of homology between the potato subunit and bacterial enzymes, especially in the C-terminus. The catalytic subunit (α) of the potato tuber shares only about 30% identity to the *E. coli*, *A. tumefaciens*, and *Rb. sphaeroides* enzymes. In contrast, the *A. tumefaciens* enzyme shares greater than 50% sequence identity with these enzymes. In addition, because the native potato tuber enzyme is a heterotetramer ($\alpha_2\beta_2$), precise functional interpretation of the potato tuber α_4 homotetramer structure is uncertain.

We describe here the first X-ray structure of a native bacterial ADPGlc PPase. The *A. tumefaciens* enzyme is comprised of two domains: an N-terminal αβα sandwich domain and a C-terminal parallel β-helix. The active site and allosteric activator site were modeled, and site-directed mutagenesis studies on residues in or near the proposed active and allosteric activator sites are presented. The structure, modeling, and mutagenesis studies are discussed in relation to the available data from a variety of ADPGlc PPases.

MATERIALS AND METHODS

Purification. Wild-type and altered *A. tumefaciens* ADPGlc PPases used for kinetic studies were overexpressed in *glgC*$^-$ *E. coli* cells [strain TGC31 (*26*)]. The recombinant ADPGlc PPases were purified as previously described using phenyl-Sepharose, DEAE, and UNO-6 (Bio-Rad) chromatography (*12*). An additional size-exclusion chromatography step (Superdex-200, Pharmacia) was used as previously described (*27*) to remove trace contaminants from the H379K enzyme.

For structural work, the *A. tumefaciens* ADPGlc PPase used was a selenomethionyl-labeled protein produced by expression in the *E. coli* methionine auxotroph strain DL41. Selenomethionyl ADPGlc PPase was purified using the same method as for the wild-type enzyme with the addition of 5 mM dithiothreitol to prevent oxidation of the selenium (*28*).

Enzyme Assays and Kinetic Analysis. During purification, enzyme assays were performed in the pyrophosphorolysis direction (*29*) in the presence of 80 mM Tris-HCl (pH 8), 0.5 mg/mL bovine serum albumin, 8 mM Mg^{2+}, 5 mM NaF, 1 mM ADP-glucose, 2 mM [^{32}P]PP_i (500–2000 cpm/nmol, NEN-DuPont), and 2 mM F6P in a final volume of 250 μL. Assays were performed at different enzyme concentrations to ensure steady-state conditions. The assays were initiated by the addition of enzyme (typically 1−100 ng in 10 μL) freshly diluted in 50 mM HEPES (pH 7.5) with 1 mg/mL bovine serum albumin and 0.5 mM DTE. A unit of activity is defined as the amount of enzyme catalyzing the production of 1 μmol of [^{32}P]ATP per minute at 37 °C.

The post-UNO-6 column recombinant enzymes (post Superdex enzyme for H379R) were suitable for kinetic assays run in the ADPG synthesis direction (*30*). Assays were initiated with the addition of enzyme (in 10 μL diluted as described above, yielding ca. 2–20 nmol of product per assay). In all cases, enzyme dilutions assayed were in a linear range (enzyme concentration versus rate) during the time course of the assay. Data points represent the average of at least two determinations that differed by less than 10%. Standard assays for activity measurements included 100 mM HEPES (pH 8), 0.5 mg/mL bovine serum albumin, 1 mM [^{14}C]Glc1P (1000–3000 cpm/nmol, NEN-DuPont), and 0.2 unit of inorganic pyrophosphatase (Sigma) together with varying amounts of ATP, Mg^{2+}, effectors, enzyme, and water in a total volume of 200 μL as indicated. Under our assay conditions for the recombinant *A. tumefaciens* enzyme, we did not observe significant activation by FBP. Saturation plots for substrates and effectors were analyzed with the use of a computer program (*31*) (using the Levenberg–Marquardt algorithm for regression) by fitting the data to a modified Michaelis–Menten equation accounting for the Hill number

as well as V_{max} and $S_{0.5}$, $A_{0.5}$, or $I_{0.5}$. The $S_{0.5}$, $A_{0.5}$, or $I_{0.5}$ values are defined here as the concentrations of substrate, activator, or inhibitor that give 50% maximal activity, activation, or inhibition, respectively. The kinetic parameters for the substrate Glc1P and cofactor Mg^{2+} were found to be similar to wild type (*3*) for the H379R and H379K enzymes.

Crystallization and Data Collection. Crystals of *A. tumefaciens* ADPGlc PPase were grown using the sitting-drop vapor diffusion method as previously described (*28*). Samples containing 5 mg/mL were mixed with an equal volume of 1.5 M lithium sulfate and 100 mM HEPES (pH 7.5) as the mother liquor. Crystals grew within 1 week and were transferred briefly to mother liquor containing 20% (v/v) glycerol for cryoprotection prior to immersion in liquid nitrogen for storage and data collection. Diffraction data were collected at the Stanford Synchrotron Radiation Laboratory (SSRL), beamline 9-1, at –170 °C. A complete set of data was obtained from a single crystal at the peak anomalous wavelength (0.979 Å) for the selenomethionyl *A. tumefaciens* ADPGlc PPase. Data were processed with MOSFLM (*32*) and scaled using SCALA (*33*) from the CCP4 program suite (*34*). A summary of the crystallographic data from the selenomethionyl *A. tumefaciens* ADPGlc PPase is presented in Table 1.

Refinement and Model Building. The anomalous Patterson map from the selenomethionyl *A. tumefaciens* ADPGlc PPase crystal contains 11 peaks above 6σ, which correlated well with the 9 expected selenomethionines in the protein. The program suites of SOLVE/RESOLVE (*35*) were used to solve the anomalous Patterson, determine phases, and perform initial automated building. Electron density map fitting was carried out using TOM (*36*), and Refmac5 (*34*) was used for all model refinement and map calculations. At a working R_{factor} of 26%, automatic water picking was performed using CNS, and the water structure was verified using $2F_o - F_c$ and $F_o - F_c$ electron density maps. Atomic coordinates for *A. tumefaciens* ADPGlc PPase structure have been deposited with the Protein Data Bank under the identifier code 3BRK.

Secondary structure assignments were determined with PROMOTIF (*37*), and the geometry of the models was analyzed using PRO-CHECK (*38*). The TOP program in the CCP4 suite was used to calculate superpositions and root mean squared deviations (rmsd). For superposition and rmsd calculations main-chain atoms were used while excluding side chains. Buried area calculations were performed using the CNS program suite.

For modeling the substrates in the catalytic site, the Glc1P thymidylyltransferases from *E. coli* (*17*; PDB ID 1MC3) and *Ps. aeruginosa* (*16*; PDB ID 1G0R) were aligned to the N-terminal catalytic domain (residues 1–308) of *A. tumefaciens* ADPGlc PPase. The *E. coli* Glc1P thymidylyltransferase structure contains TTP/Mg^{2+} bound in the active site (PDB ID 1MC3), which following alignment was transferred to the *A. tumefaciens* ADPGlc PPase model. TTP was replaced with ATP by alignment of all like atoms between the two molecules. The *Ps. aeruginosa* Glc1P thymidylyltransferase structure (PDB ID 1G0R) contains Glc1P bound in the active site that was moved to the *A. tumefaciens* ADPGlc PPase model following alignment. F6P was modeled in the proposed activator site by manual alignment of the phosphate group with the bound sulfate group in the model.

Table 1: X-ray Data Collection, Refinement, and Model Statistics

X-ray Data Collection	
wavelength (Å)	0.979
resolution limits (Å)	79−2.1
measurements	462576
unique reflections	33214
R_{sym} (%), overall/outer shell[a,b]	6.9/30.8
I/σ, overall/outer shell[b,c]	7.2/2.0
completeness (%), overall/outer shell[c]	99.9/100
multiplicity, overall/outer shell[c]	13.9 (14.3)
space group	I_{222}
unit cell, *a*, *b*, *c* (Å)	85.38, 93.79, 140.29
unit cell, α, β, γ (deg)	90
Phasing Statisics	
phasing power acentric (av/max)[d]	1.77/3.94
phasing power centric (av/max)[d]	2.21/4.10
R_{Cullis}[e]	0.64
FOM/solvent density modification[f]	0.317/0.923
Model Statistics	
molecules/asymmetric unit	1
R_{factor}[g] (R_{free})[h] for $\|F_o\| > 0$	22.0 (26.6)
no. of residues (actual/expected)[i]	390/420
no. of water molecules[j]	250
rmsd from target[k]	
bonds (Å)	0.014
angles (deg)	1.4
Ramachandran plot[l]	88/11/1/0
average *B*-factor	29.2
average Wilson *B*-factor	27.5
secondary structure	
β-strand residues (%)	26.9
α-helical residues (%)	20.3
3_{10} helical residues (%)	5.1

[a] $R_{sym} = \sum|I - \langle I\rangle|/\sum I$, where I is the observed intensity and $\langle I\rangle$ is the statistically weighted average intensity of multiple observations of symmetry-related reflections. [b] The outer shell corresponds to 2.10–2.15 Å. [c] Intensity signal-to-noise ratio. [d] Phasing power = $\langle|F_h(calc)|/$ phase-integrated lack of closure$]\rangle$, where F_h is the heavy atom structure amplitude. [e] $R_{Cullis} = \langle$phase-integrated lack of closure$\rangle/\langle|F_{PH} - F_P|\rangle$. [f] FOM = $\langle\cos \Delta\alpha_j\rangle$, where $\Delta\alpha_j$ is the phase angle error for the *j*th phase angle. [g] $R_{factor} = \sum||F_o| - |F_c||/\sum|F_o|$, where F_o and F_c are observed and calculated structure factor amplitudes. [h] R_{free} is the R_{factor} calculated from 10% of the native data set, randomly selected and not used in the structure refinement. [i] Number of residues reported for a monomer. [j] Total water molecules for the asymmetric unit. [k] Root mean square deviation of bond lengths and bond angles from ideal geometry. [l] Percentage of residues in most favored/additionally allowed/generously allowed/disallowed regions of the Ramachandran plot, according to PROCHECK (*40*).

Mutagenesis and Sequence Analysis. Sequence information and primer design were managed using VectorNTI (version 8.0; Invitrogen). Alignments of the sequences were performed with the AlignX module of VectorNTI. Site-directed mutagenesis was performed using the QuikChange kit (Stratagene). The following forward primers (and their complements) were used for conversion of glycine to alanine in the amino terminus and histidine to arginine and lysine at position 379 (base changes are in bold): G20A, 5′-C TAT GTC CTC GCA G**C**C GGA AGA GGA A-3′; G21A, 5′-C TAT GTC CTC GCA GGC G**C**A AGA GGA A-3′; G23A, 5′-GCA GGC GGA AGA **GCA** AGC CGT CTG AA-3′; H379R, 5′- GTC GTC ATC GAC **CGT** GGC GTG GTC ATT CCG-3′; H379K, 5′-GTC GTC ATC GAC **AAA** GGC GTG GTC ATT CCG-3′.

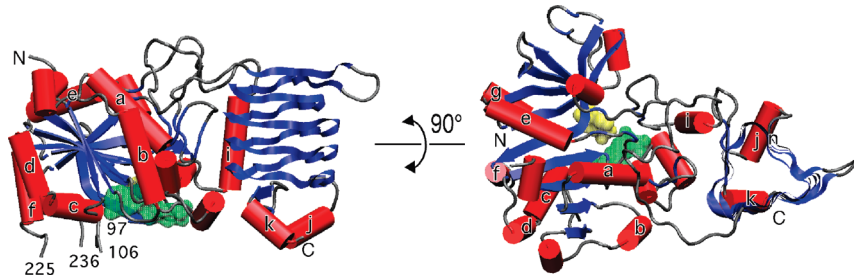

FIGURE 1: Structure of *A. tumefaciens* ADPGlc PPase with ATP and Glc1P modeled in the active site. Two views, rotated 90° relative to one another, illustrating the main secondary structural elements are presented. The α-helices (a−k) and β-sheets of *A. tumefaciens* ADPGlc PPase are shown in red and blue, respectively. The modeled ATP and Glc1P are shown as space filling in green and yellow, respectively. Figures 1−6 and 8 were created using VMD (*60*).

RESULTS AND DISCUSSION

Structure Determination. As previously described, *A. tumefaciens* ADPGlc PPase was crystallized, and a single data set was collected to 2.1 Å resolution at the Stanford Synchrotron Radiation Laboratory (*28*). The final model consists of 395 residues (419 residues total possible) with residues 2–6, 99–105, 226–235, and 419–420 omitted due to poor electron density in these regions, yielding an R_{factor} = 22.0% and R_{free} = 26.6%. Detailed refinement statistics and geometric properties of the final model are summarized in Table 1, and the structure has been deposited in the Protein Data Bank with identifier 3BRK.

Structure of the Monomer. A schematic diagram of the X-ray crystal structure of the *A. tumefaciens* ADPGlc PPase monomer depicting the secondary structure elements is shown in Figure 1. The overall structure of the protein is partitioned into two subdomains. The N-terminal domain contains the catalytic site and is composed of residues 7–308. The location of the catalytic site is depicted in Figure 1 using transparent space-filling models of the substrates. The crystal structure we report here for *A. tumefaciens* ADPGlc PPase is substrate free, and Glc1P and ATP were modeled in the active site using structurally related sugar nucleotidyl transferases as described in Materials and Methods. The overall structure of the N-terminal domain contains 30% of the residues in an α-helical structure, 24% of the residues in a β-strand conformation, and 46% in a random coil. The secondary structural elements of the N-terminal domain form a seven-strand β-sheet with six of the β-strands oriented parallel to each other, while the β-strand of residues 200−212 is positioned antiparallel to the other strands and runs through the middle of the β-sheet. The seven-strand central β-sheet of the N-terminal domain is surrounded by 12 α-helices (α-helices a−l) that are amphipathic in nature and have their hydrophobic faces interacting with the hydrophobic side chains of the β-sheet. Residues 255–308, although considered part of the N-terminal domain because some side chains project into the active site, do not participate in formation of the seven-strand 12-helix structure. Rather these residues are in a random coil structure with a 3-turn α-helix spanning residues 272−281 (α-helix m). These residues of the N-terminal domain make the closest approach to the C-terminal domain, with the α-helix (m) having one hydrophobic face directly interacting with the N-terminal domain and the other hydrophobic face lying against the C-terminal domain. This region may be important in communication between the N-terminal and C-terminal domains.

The C-terminal residues 309−420 form a left-handed parallel β-helix domain. The β-helix domain is comprised of four β-strands stacked to form a small left-handed parallel β-helix with a short β-strand (residues 310−312) topping the β-helix and two small β-strands (residues 400−402 and 407−410) mixed with two short α-helices capping the bottom of the β-helix. Residues 314−319 do not participate in formation of the parallel β-helix but form a large loop that projects out from the β-helix surface. The overall three-dimensional fold of this domain approximates a triangular coil with sides composed of short parallel β-strands. The central core is completely filled with hydrophobic side chains, with no solvent present, and the outer surface with hydrophilic side chains. In *A. tumefaciens* ADPGlc PPase one of the three outside faces of the β-helix is hydrophobic and packs against the catalytic domain (α-helix m) with the other two faces interacting with solvent. Figure 2 shows the side-chain packing in the left-handed parallel β-helix of *A. tumefaciens* ADPGlc PPase.

Oligomerization. The majority of bacterial ADPGlc PPase enzymes are believed to function as homotetramers. The *A. tumefaciens* ADPGlc PPase crystallized in the I_{222} space group with only one molecule in the asymmetric unit. This crystallization lattice indicates that all four monomers of the *A. tumefaciens* ADPGlc PPase tetramer crystallized in an identical structure, and the tetramer must be generated by the crystallographic space group symmetry. Using the CNS program package to generate the nearest neighbors based on the I_{222} space group symmetry, a tetramer with the structure shown in Figure 3b was generated. In the tetramer the N-terminal catalytic domains interact with each other and pack in the inside of the tetramer, while the C-terminal β-helix domains stack along the outside. Each monomer in the tetramer makes contact with all other monomers within the tetramer. These interactions may play a role in allosteric communication among the molecules of the tetramer.

The largest buried surface area that is formed in the tetramer occurs at the stacking of the β-helix domains, ~2800 Å^2. The dimer formed by the stacking of the β-helix domains is shown in Figure 3a. The tops of the two β-helices come together to form perfect β-sheet backbone interactions, resulting in a 10-turn β-helix. However, since the C-terminal portions of both β-helices are capped by α-helices, the two β-helices must stack antiparallel to each other. Stacking antiparallel also allows for optimum packing of the hydrophobic side chains in the center of the β-helix. The top surface of the β-helix contains a phenylalanine (Phe311), alanine (Ala322), and valine (Val327) which form a hydro-

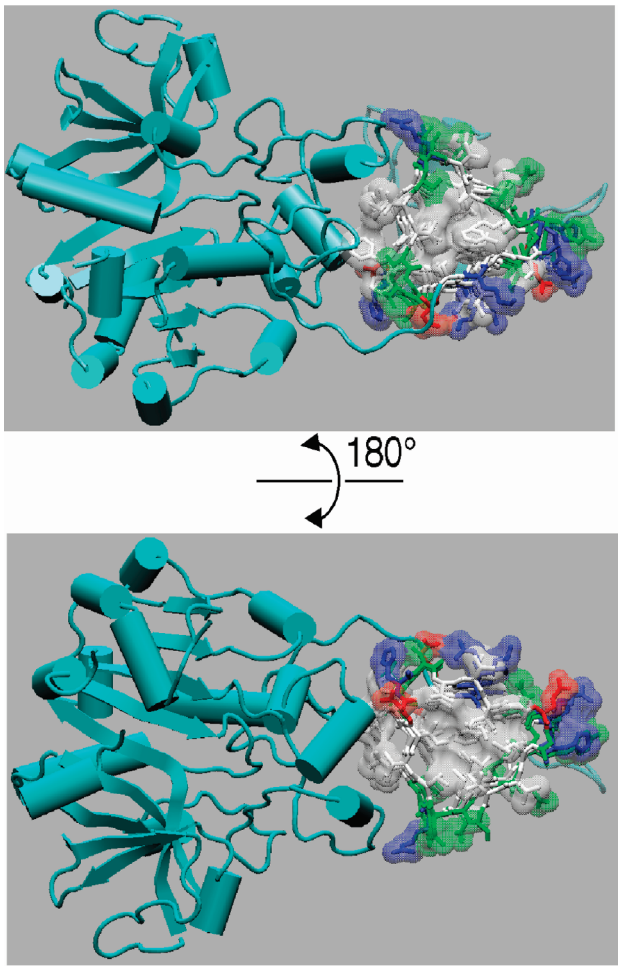

FIGURE 2: Side-chain packing in the parallel β-helix. The N-terminal catalytic domain is shown as cyan and is illustrated to show secondary structure. The C-terminal parallel β-helix domain is shown as space filling with color coding based on residue type: nonpolar, white; polar, green; basic, blue; acidic, red. In order to visualize both faces of the parallel β-helix, two views are shown rotated 180° from each other.

phobic surface of ~1500 Å that is exposed to solvent in the monomer. Stacking of the β-helices in the dimer completely buries this hydrophobic surface, and positioning of the β-helices antiparallel to each other allows for optimum packing of the side chains.

Comparison of A. tumefaciens and Potato Tuber ADPGlc PPase. As indicated earlier, the crystal structure of the recombinant potato tuber ADPGlc PPase α-subunit was recently reported (*4*; PDB ID 1YP2). In native potato tuber ADPGlc PPase, the α- and β-subunits share 53% identity. While only the α-subunit is capable of catalysis alone, recent studies have indicated that the β-subunit can bind substrates and effectors and the heterotetramer's catalytic and allosteric properties appear to be a product of synergy between the two subunits (*39–41*).

A sequence alignment of *A. tumefaciens* and potato tuber ADPGlc PPase based on their X-ray structures is given in Figure 4, and graphical representations of the sequence identity are shown in Figure 4b. *A. tumefaciens* and potato tuber ADPGlc PPase share 31% overall amino acid sequence identity; however, the N- and C-terminal domains of the proteins display very different sequence homology. The N-terminal catalytic domain displays the highest sequence identity of the two domains at 36%. As shown in Figure 4b, the residues of the N-terminal domain that are similar between *A. tumefaciens* and potato tuber ADPGlc PPase are residues that form the core of the domain and are clustered around the active site. The parallel β-helix of the C-terminal domain displays only 19% sequence similarity between the *A. tumefaciens* and potato tuber enzyme.

The two proteins display similar secondary structure, and structural alignment of the backbone structure of *A. tumefaciens* and potato tuber ADPGlc PPase yields an rmsd of 1.7 Å^2. However, when the alignment of the *A. tumefaciens* and potato tuber enzymes is constrained to residues of the individual domains, it is apparent that the N- and C-terminal domains are shifted relative to each other. Figure 5 shows the backbone alignment comparison for the *A. tumefaciens* and potato tuber enzymes where only residues of the N-terminal domain (Figure 5a) or the C-terminal domain (Figure 5b) are aligned. The N-terminal catalytic domains of *A. tumefaciens* and potato tuber ADPGlc PPase display similar secondary structure with an rmsd of 1.5 Å^2. When the N-terminal domain is aligned, however, the parallel β-helix domains of the two structures are offset by as much as 3 Å, with the β-helix of the potato tuber enzyme packing closer to the N-terminal domain than seen in the *A. tumefaciens* enzyme (Figure 5a). However, if only the residues of the C-terminal domain are used in the alignment, it is apparent that the parallel β-helices of the two enzymes are nearly identical in structure. The C-terminal domains of both structures form a 5-turn left-handed parallel β-helix, which displays an rmsd of only 0.84 Å^2 for residues that can be aligned between the two structures. It is not clear what precise roles the packing interface of the two domains plays in catalysis, regulation, or stability, and future mutagenesis studies will be needed to address these questions.

A. tumefaciens and potato tuber ADPGlc PPase both apparently function as tetramers, and the functional unit of both enzymes could be identified in their crystal form. The α1α2 and α3α4 dimers are similar between the two structures, being formed by stacking of the β-helix structure to create a long β-helix (see Figure 3a for tetramer nomenclature). However, in the *A. tumefaciens* tetramer the α1α2 and α3α4 dimers pack tighter than in the potato tuber structure. The potato tuber α3α4 dimer is shifted outward about 7 Å compared to the *A. tumefaciens* structure. Formation of the tetramer in the potato tuber enzyme involves two disulfide bonds between Cys12 of α1−α3 and α2−α4. The subunits of the *A. tumefaciens* tetramer are not covalently linked, and it may be the constraints of the potato tuber disulfide bonds that push the α1α2 and α3α4 dimers apart.

Since the reported potato tuber structure is of an *E. coli* expressed α_4 homotetramer and not the native $\alpha_2\beta_2$ heterotetramer, it is not clear if the native enzyme adopts the oligomeric structure reported by Geiger and co-workers for the α4 homotetramer or a structure more similar to the *A. tumefaciens* enzyme. The homotetrameric potato enzyme displays altered regulatory properties compared to the native enzyme with the homotetramer displaying a greater than 20-fold reduction in apparent affinity for the activator 3PGA and an ~8-fold increased apparent affinity for the inhibitor phosphate (*25*). The *A. tumefaciens* enzyme used for the studies reported here is the native enzyme that displays wild-type activity and regulation. However, one limitation of the

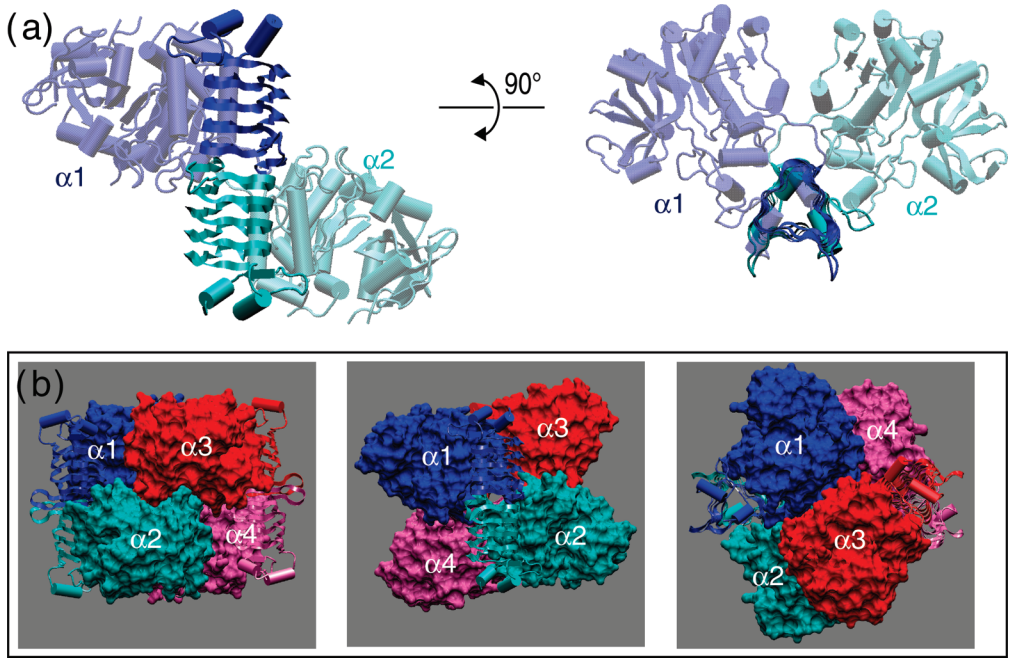

FIGURE 3: The oligomeric structure of *A. tumefaciens* ADPGlc PPase. (a) The dimer formed when α1 (blue) and α2 (cyan) monomers come together by stacking of the parallel β-helix. Two views of the dimer are shown rotated 90°. (b) The tetramer of *A. tumefaciens* ADPGlc PPase found in the crystal lattice. The tetramer can be represented as a dimer of dimers, and the monomers that form dimers are shown in like colors: α1 and α2, blue and cyan; α3 and α4, red and pink. Three views of the tetramer are shown: relative to the left panel, the center panel is rotated 90° about the *y*-axis and the right panel is rotated 90° about the *x*-axis.

A. tumefaciens enzyme structure, like the potato tuber structure, is that the enzyme was crystallized in the sulfate-inhibited state. Further comparison of the oligomeric structure of the bacterial and eukaryotic enzymes will require the structure of the native enzyme from eukaryotes containing both the α- and β-subunits.

The Active Site and Catalysis. Although the kinetic mechanism for ADPGlc PPase has been determined to be sequential Iso Ordered Bi Bi with ATP binding first followed by Glc1P (*42–44*), and subsequent product release of PP_i and ADPGlc, the precise molecular details of catalysis are not fully understood. A metal divalent cation (Mg^{2+}) is also required for catalysis, presumably for correct binding of ATP as well as stabilization of the negatively charged transition state.

The crystal structure for *A. tumefaciens* ADPGlc PPase we report here is substrate free in the presence of the inhibitor sulfate. The active site is located in a large surface pocket in the N-terminal subdomain. The obvious choice to assist in substrate modeling would be the ADPGlc PPase structure from the potato tuber α-subunit (*4*). However, the potato tuber ADPGlc PPase α-subunit was also crystallized in the apoenzyme form with sulfate present, and attempts to soak in substrates or products resulted in very low occupancy structures. In addition, the potato tuber α-subunit ATP complex structure adopted a conformation that was not compatible with catalysis and was different from the structure reported for other sugar-nucleotide pyrophosphates with nucleotide triphosphates bound. These difficulties may have arisen due to crystallographic lattice constraints imposed by the apoenzyme crystals. Due to these shortcomings, we examined other reported structures to use in generating a model of the catalytically active enzyme.

ATP, Mg^{2+}, and Glc1P were modeled in the active site using reported structures of these compounds bound to Glc1P thymidylyltransferases. The Glc1P thymidylyltransferases consist of ~300 residues and lack the C-terminal parallel β-helix seen in the ADPGlc PPases; these enzymes display approximately 30% sequence identity with the N-terminal domain of *A. tumefaciens* ADPGlc PPase. Extensive structural studies have been performed with the bacterial Glc1P thymidylyltransferases from *E. coli* (*RffH*) and *Ps. aeruginosa* (*RmlA*). Both of these enzymes display structural similarity to the N-terminal domain of *A. tumefaciens* ADPGlc PPase with an rmsd of 1.7 $Å^2$, the same as seen with alignment of the *A. tumefaciens* and potato tuber ADPGlc PPase.

The *E. coli* Glc1P thymidylyltransferase structure with TTP/Mg^{2+} bound (PDB identifier 1MC3 was used to model the nucleotide and divalent metal cation-binding site. TTP was replaced by ATP in the model by alignment of atoms identical between the compounds. Glc1P was modeled in the active site of *A. tumefaciens* using the reported structure of *Ps. aeruginosa* Glc1P thymidylyltransferase with Glc1P bound (PDB identifier 1G0R).

The active site is located in the N-terminal domain in a cleft created when the two antiparallel β-sheets come together (Figure 6), with the side chains of Arg25 and Lys188 approaching each other to form one side of the active site. Figure 6d shows this region of the active site in our model and comparison to the *E. coli* Glc1P thymidylyltransferase TTP/Mg^{2+}complex structure. In the *A. tumefaciens* ADPGlc PPase ATP/Mg^{2+}/Glc1P model, ATP is in position to make extensive hydrogen bonds with the backbone and side chains of residues 20−25. The modeled ATP lies near a highly conserved glycine-rich loop consisting of residues 18−26 (see sequence alignment, Figure 7) and actually lies too close to this loop to be at an energy minimum. Two of the atoms on the ATP molecule are less than 2 Å from the backbone atoms of Gly20 and Gly21 (Figure 6c). In the *E. coli* Glc1P

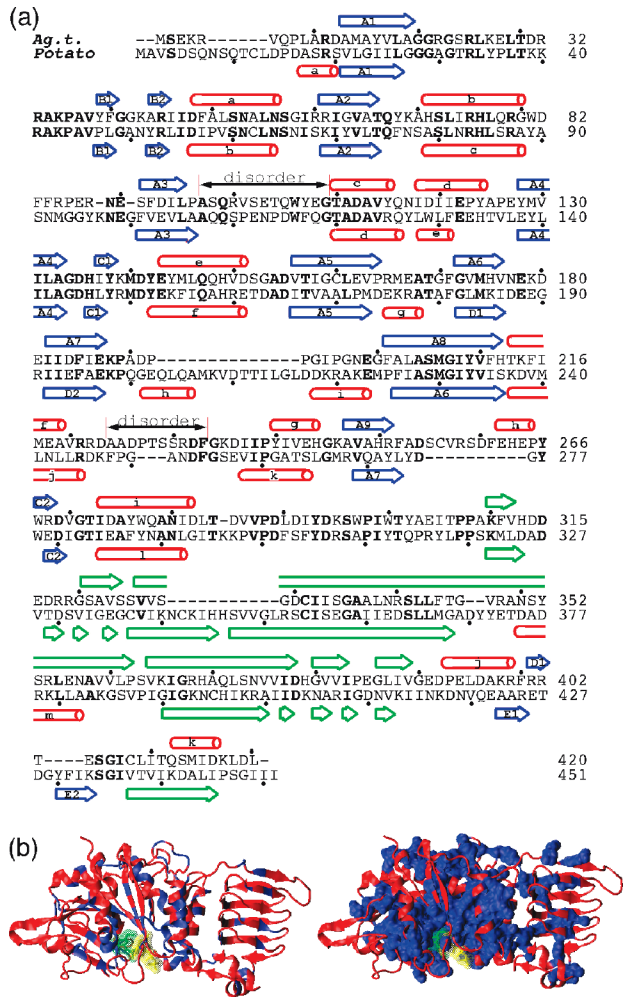

FIGURE 4: (a) Sequence alignment of *A. tumefaciens* and potato tuber ADPGlc PPase based on their crystal structures. Helices are indicated by red tubes, and β-strands are indicated by blue arrows. Residues participating in formation of the parallel β-helix are indicated by a green arrow. Conserved residues are shown in bold. (b) Illustration of the *A. tumefaciens* ADPGlc PPase structure color-coded based on sequence identity with the potato tuber enzyme; blue indicates identical residue, and red represents no similarity. The left and right panels are the same except the right panel shows all identical side chains as space filling.

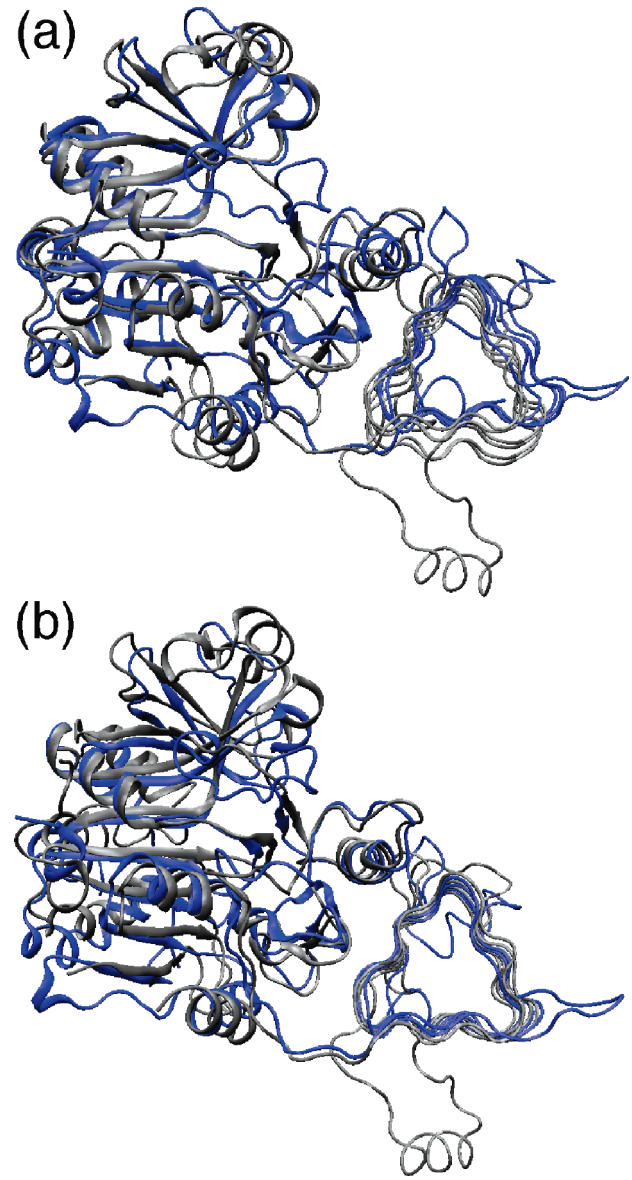

FIGURE 5: Comparison of the crystal structures of *A. tumefaciens* and potato tuber ADPGlc PPase. The *A. tumefaciens* structure is shown in blue, and the potato tuber enzyme structure is shown in gray. (a) Alignment of the two structures when only residues of the N-terminal catalytic domain are used for superposition. (b) Alignment of the two structures using the residues of the parallel β-helix domain for superposition.

thymidylyltransferase structure, used for nucleotide modeling, this region of the backbone is shifted about 3 Å out of the active site (Figure 6d). The glycine residues in this stretch of backbone would provide the flexibility necessary for movement of the backbone upon ATP binding. The difference in the ATP positioning may be due to the inhibited conformation of the *A. tumefaciens* enzyme in the presence of sulfate, a limitation similar to that seen with the potato tuber enzyme (*4*). It should also be noted that in the *A. tumefaciens* structure this region displays a high crystallographic *B*-factor indicating some disorder in the crystal. This is in accord with the predicted flexibility of this glycine-rich loop and has been observed in other proteins including cAMP-dependent protein kinase (*45*), Rm1A (*16*), GlmU (*15*), and the α-subunit of potato tuber ADPGlc PPase (*4*).

No studies have addressed the precise role of these glycine residues in any ADPGlc PPase. To probe the ATP/Mg^{2+} binding site, we performed site-directed mutagenesis to generate the G20A, G21A, and G23A enzymes. Table 2 lists the kinetic parameters for the G20A, G21A, and G23A enzymes and, for comparison, the kinetic data on the R22A and R25A enzymes that were previously reported (*3*). All mutants displayed reduced affinity for ATP compared to the wild type; the $S_{0.5}$ for ATP was increased approximately 2−8-fold in the mutant enzymes. All of these mutations displayed significant decreases in V_{max}, suggesting that even though ATP can bind, it may not be in a catalytically favorable location. The G20A and G21A enzymes were most severely impaired with respect to catalysis, displaying a 1210- and 8643-fold decrease in V_{max}, respectively. Glycine residues at positions 20 and 21 lie next to the ring structure of the modeled ATP in our *A. tumefaciens* ADPGlc PPase substrate complex model and must move to accommodate ATP/Mg^{2+}/G1P binding. The substitution of alanine in this region may have reduced flexibility of the backbone resulting in loss of activity and affinity for ATP. The three glycine residues in this stretch of backbone would appear to be crucial to provide

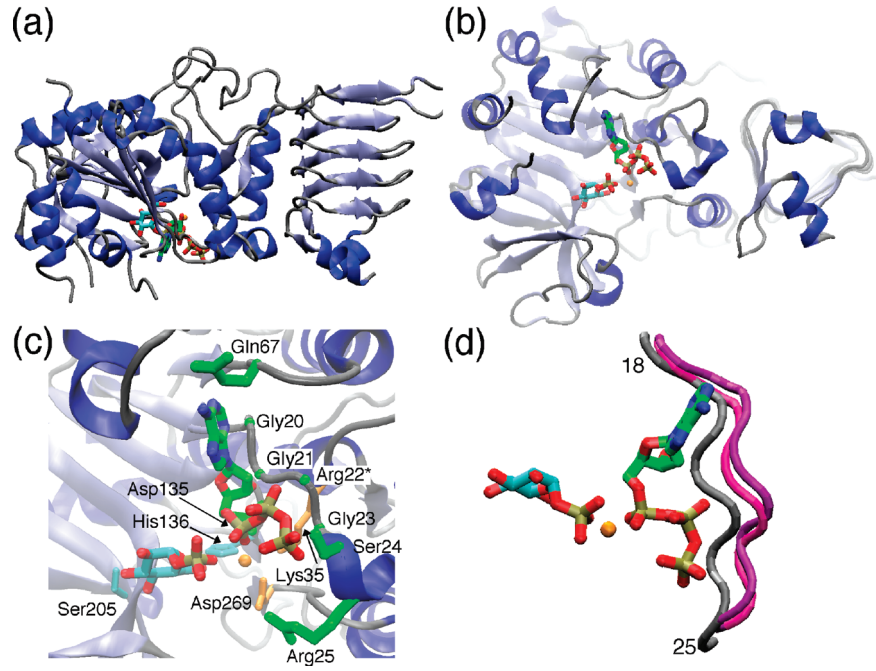

FIGURE 6: The active site region of *A. tumefaciens* ADPGlc PPase. (a) A schematic diagram of the *A. tumefaciens* ADPGlc PPase structure with ATP, Glc1P, and Mg^{2+} modeled in the catalytic pocket. ATP and Glc1P are colored based on atom type. Color coding: ATP carbon, green; Glc1P carbon, cyan; oxygen, red; nitrogen, blue; phosphate, brown; magnesium, gold. (b) Alternate view of the *A. tumefaciens* ADPGlc PPase ATP/Mg^{2+}/Glc1P complex. The structure is rotated approximately 90° relative to (a) to give a view where you are looking straight into the active site. Color coding is the same as in (a). (c) Close-up view of the active site showing residues that may interact with the substrates. Residues are color-coded based on the substrate they are within 4 Å of: ATP, green; Mg^{2+}, gold; Glc1P, cyan. Arg22 is modeled as a glycine in the X-ray structure due to disorder of the side chain. (d) The proximity of the modeled ATP/Mg^{2+}/Glc1P to the peptide backbone of residues 18–25. The *A. tumefaciens* ADPGlc PPase backbone is shown in gray, and the X-ray structures used for modeling, *E. coli* (PDB ID 1MC3) and *Ps. aeruginosa* (PDB ID 1G0R), are shown in pink and purple, respectively.

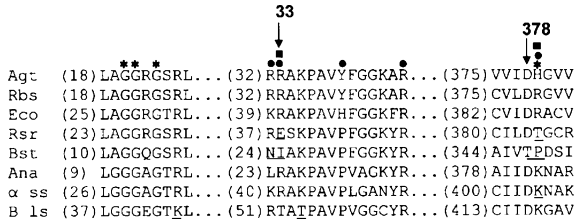

FIGURE 7: Comparison of the amino acid sequence of the *A. tumefaciens* (Agt) ADPGlc PPases with other bacterial ADPGlc PPases. Multiple alignment of selected regions of the N- and C-terminal domains from different activator classes represented by the sequences from *Rb. sphaeroides* (Rbs, (*8*)), *E. coli* (Eco (*5*)), and *Rs. rubrum* (Rsr, Genbank AF097739), *B. stearothermophilus* (Bst, (*26*)), *Anabaena* (Ana, (*6*)), and the potato tuber α-subunit (α ss [small subunit]) and *β*-subunit (*β* 1s [large subunit]), respectively (*4*)). The symbol (*) identifies the residues subject to site-directed mutagenesis in the glycine-rich region (G20A, G21A, and G23A) and adjacent to Asp378 (H379R, K). Residues that participate in a salt bridge (Arg33, Asp378) are indicated by arrows. The symbol (●) designates residues near the sulfate binding site of the *A. tumefaciens* enzyme while the symbol (■) points out residues involved in Sulfate Site 1 of the potato tuber α-subunit structure (*4*). The underlined residues in the Rsr and Bst sequences reveal differences associated with altered regulation. Underlined residues in the potato α ss sequence point out one of the C-terminal lysines involved in activation by 3-PGA (Lys382, (*21*)). Underlined residues in the *β* 1s enzyme in the glycine-rich and RRAKPAV regions highlight residues mutated to restore activity to the Ls (*46*).

the flexibility necessary for movement of the backbone upon actual ATP binding to an activated form of the enzyme for optimal catalysis.

In the *A. tumefaciens* Glc1P/Mg^{2+}/ATP model, only two other residues are within 4 Å of the ATP molecule, Asp135 and Gln67. The modeled Mg^{2+} ion is located near the ATP α-phosphate and is in position to have interactions with the side chains of Lys35, Asp269, and Asp135, as well as the oxygens of the phosphate at the α-phosphate of the ATP and Glc1P. Further, Glc1P is positioned perpendicular to the ATP molecule with the phosphate of Glc1P located 4 Å from the α-phosphate of the ATP. Glc1P is in position to make only three interactions with the protein, involving the side chain of His136 and the side chain and backbone of Ser205.

Our model is supported by a number of mutagenesis studies involving the residues involved in ATP/Mg^{2+} binding. In previous experiments we have shown by alanine scanning mutagenesis that both Arg22 and Arg25 are important for catalysis (*3*). The catalytic roles of Arg25 (contacts ATP) and Lys35 (coordinates Mg^{2+}) suggested by our model are in agreement with the recent findings of Ballicora et al. (*46*), in which the catalytic activity of the *β*-subunit (in a complex with an inactivated α-subunit) of the potato tuber enzyme was restored by mutations at the analogous positions (see sequence alignment in Figure 7). Both Asp135 and Gln67 are highly conserved among the bacterial ADPGlc PPases, and Asp142 in the *E. coli* enzyme [analogous to Asp135 of *A. tumefaciens* (*11*)] has been shown to be essential for catalysis. The catalytic role of Asp269, which coordinates Mg^{2+}, has recently been demonstrated by mutagenesis of the analogous residue in the *E. coli* enzyme (*47*). Lys188 is analogous to Lys195 in *E. coli*; mutagenesis of this residue was found to decrease apparent binding of Glc1P (*9*).

With respect to residues involved in binding Glc1P, mutagenesis of the position analogous to Ser205 in the *E. coli* enzyme resulted in substantial decreased apparent affinity

Table 2: Kinetic Parameters for *A. tumefaciens* ADPGlc PPase Active Site Mutants

	ATP[a]		ATP + F6P[b]		F6P[c]
	$S_{0.5}$ (mM)	V_{max} (units/mg)	$S_{0.5}$ (mM)	V_{max} (units/mg)	$A_{0.5}$ (mM)
WT[d]	0.21 ± 0.04	12 ± 1.0	0.051 ± 0.004	150 ± 4.1	0.13 ± 0.01
G20A	na[e]	na[f]	0.40 ± 0.03	0.11 ± 0.01	0.31 ± 0.15
G21A	0.58 ± 0.12	0.0014 ± 0.0002	0.43 ± 0.13	0.0078 ± 0.0005	0.5 ± 0.20
G23A	1.8 ± 0.19	0.19 ± 0.01	0.69 ± 0.15	4.1 ± 0.29	0.45 ± 0.02
R22A[d]	0.71 ± 0.17	5.8 ± 0.60	0.39 ± 0.061	60 ± 12.0	0.90 ± 10.0
R25A[d]	0.45 ± 0.05	0.03 ± 0.002	0.20 ± 0.02	2.3 ± 0.06	0.75 ± 0.04

[a] Saturation plots were performed in the presence of 10 mM Mg^{2+} and 1 mM Glc1P under standard conditions as described (*3*). [b] Saturation plots were performed as described (*3*) in the presence of 2 mM F6P. [c] Saturation plots were performed in the presence of 2 mM ATP, 5 mM Mg^{2+}, and 1 mM Glc1P under standard conditions as described (*3*). [d] Wild-type, R22A, and R25A data from ref *3*. [e] The G20A enzyme displayed nonsaturation kinetics. [f] The G20A enzyme displayed nonsaturation kinetics. Specific activity at 8 mM ATP was 0.02 unit/mg.

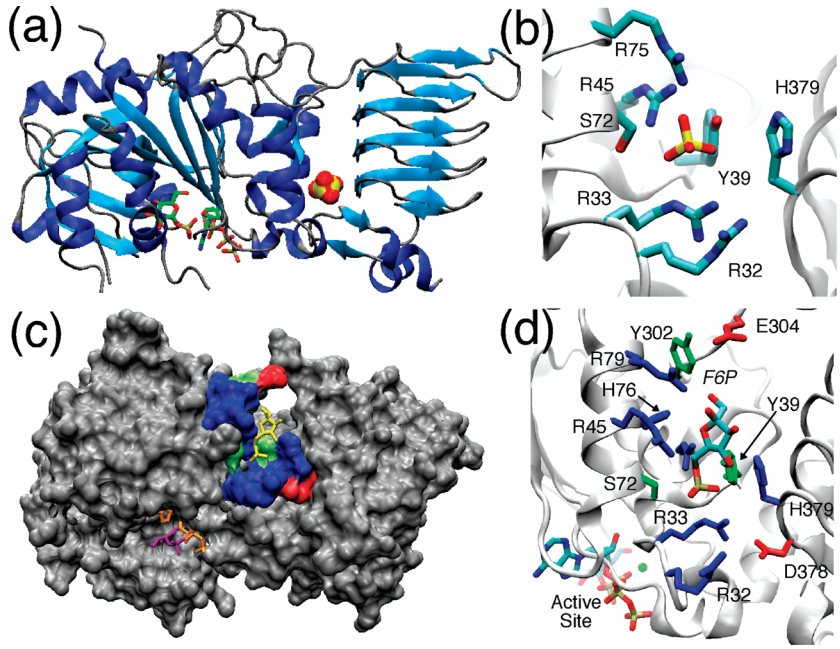

FIGURE 8: The activator-binding site of *A. tumefaciens* ADPGlc PPase. (a) The sulfate-binding site of *A. tumefaciens* ADPGlc PPase shown using the space-filling model of the sulfate atom which is colored based on atom type: sulfate, yellow; oxygen, red. The catalytic site is depicted by modeled ATP/Glc1P which are colored by atom type: carbon, green; oxygen, red; nitrogen, blue; phosphate, brown. (b) The sulfate-binding site positioned between the catalytic domain and the parallel β-helix. Side chains that are in position to participate in interactions with the sulfate are shown. Coloring is based on atom type with carbon, cyan; oxygen, red; nitrogen, blue; and sulfate, yellow. (c) Space-filling representation of the *A. tumefaciens* ADPGlc PPase structure with F6P, shown as yellow licorice, modeled in the proposed activator-binding site. The residues surrounding the F6P are colored by residue type: basic, blue; acidic, red; polar, green. To show the location of the active site, ATP and Glc1P are shown in orange and purple, respectively. (d) View of the proposed F6P binding site showing side chains that may interact with F6P. The F6P molecule is colored by atom type: carbon, cyan; oxygen, red; phosphate, brown. Side chains are colored by residue type as described in (c).

for Glc1P (*47*). Also, His136 is highly conserved among the ADPGlc PPases, and mutagenesis of the analogous residue in the *E. coli* enzyme to glutamine resulted in an ~50% decrease in activity compared to wild type (*48*).

The *A. tumefaciens* ADPGlc PPase ATP/Mg^{2+}/Glc1P model constructed provides insight into the sequential ligand binding observed in ADPGlc PPases, with ATP and Mg^{2+} binding prior to Glc1P. ADPGlc PPases only function under conditions of high Glc1P and ATP, when carbon storage is beneficial to the cell. Since Glc1P must bind after ATP/Mg^{2+}, the enzyme must have evolved to minimize Glc1P binding in the absence of the other substrates. It appears that, to reduce Glc1P binding in the absence of ATP, the enzyme limits the number of interactions with Glc1P. As seen in our model, Glc1P has only three hydrogen bonds with the protein; therefore, enthalpy of binding Glc1P in the absence of ATP would be low. ATP and Mg^{2+} in the active site provide additional interactions with bound Glc1P that may stabilize binding in the active site. In addition, steric constraints of the active site may dictate that ATP/Mg^{2+} binds prior to Glc1P. In our model the width of the ATP/ Mg^{2+}/ Glc1P complex, measured from the tip of Glc1P across to the ATP where it approaches Gly20, is 13.2 Å. The active site width at this region prior to backbone movement of residues 20–25 is only 14.3 Å. Therefore, binding of ATP first may allow conformational changes to occur around residues 20–25 and an opening of the active site to permit Glc1P to bind. If Glc1P bound first, steric constraints would prohibit ATP from entering the active site. The *A. tumefaciens* ADPGlc PPase ATP/Mg^{2+}/Glc1P model presented provides a basis for rational mutagenesis on amino acids in or near the active site to test these hypotheses.

The Allosteric Binding Site of Bacterial ADPGlc PPases. As the initial step in carbon storage, the reaction catalyzed by ADPGlc PPase is tightly regulated. ADPGlc PPases from different species display different effector specificity in

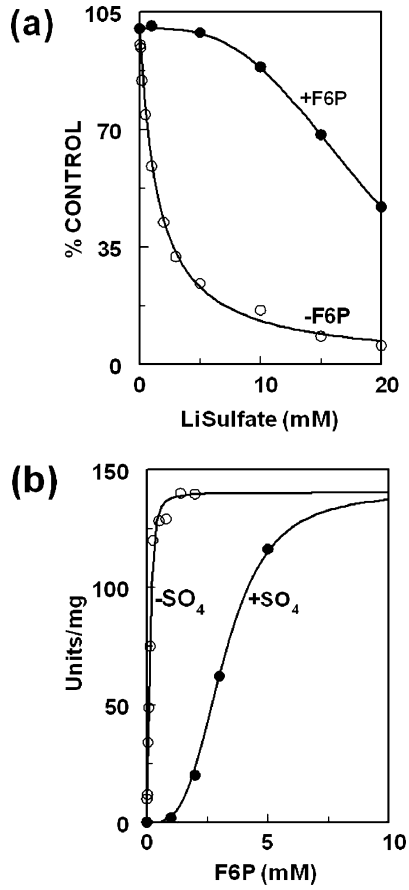

FIGURE 9: Effect of sulfate on *A. tumefaciens* ADPG PPase. (a) Sulfate saturation plot in the absence of activator (−F6P, apparent $I_{0.5\ SO_4} = 1.5$ mM, Hill number = 1) and presence of 2 mM F6P (+F6P, apparent $I_{0.5\ SO_4} = 19.2$ mM, Hill number = 3.1). (b) F6P saturation plot in the absence of sulfate ($-SO_4$, $A_{0.5\ F6P} = 0.13$ mM, Hill number = 2.0) and presence of 25 mM sulfate ($+SO_4$, $A_{0.5\ F6P} = 3.2$ mM, Hill number = 3.4). Lithium sulfate was used for the assays shown; similar effects were seen with sodium and ammonium sulfate.

accord with their carbon metabolism pathways, and these allosteric modulators have been shown to affect both the $S_{0.5}$ values for the substrates as well as the V_{max} of the reaction (*1, 2*). The major activator of *A. tumefaciens* ADPGlc PPase is F6P. Under our assay conditions, F6P binds to *A. tumefaciens* ADPGlc PPase with an apparent affinity ($A_{0.5}$value) of ~130 μM and acts to increase the apparent affinity for ATP by decreasing the $S_{0.5}$ value ~4-fold and increases the V_{max} by ~12-fold (*3*).

The *A. tumefaciens* enzyme was crystallized in the presence of a high concentration of lithium sulfate. In the *A. tumefaciens* X-ray model, a sulfate ion is bound in the cleft between the N- and C-terminal domains and is coordinated by the side chains Arg33, Tyr39, Arg45, and Ser72, with Arg32, Arg75, and His379 within 5 Å of the sulfate ion (Figure 8a,b). This sulfate ion appears highly ordered in the crystal and is modeled at 100% occupancy. Because sulfate and phosphate share similar chemical properties, sulfate may be binding in the phosphate position of the F6P binding pocket for the *A. tumefaciens* enzyme. This is supported by kinetic studies with F6P and sulfate indicating a competitive relationship between the two (Figure 9). The I_{50} value for sulfate increases ~13-fold in the presence of 2 mM F6P. Similarly, the $A_{0.5}$ value for F6P increases ~25-fold in the presence of 25 mM sulfate.

As outlined below, available mutagenesis and alignment data suggest that the residues involved in sulfate coordination in our model are also involved in activator binding. Alanine scanning mutagenesis of arginines in the N-terminus of the *A. tumefaciens* ADPGlc PPase revealed that the R33A and R45A enzymes were desensitized to F6P activation and the R32A enzyme exhibited decreased affinity for F6P. Furthermore, a parallel study with the similar *Rb. sphaeroides* ADPGlc PPase (see alignment in Figure 7) revealed that the R33A enzyme was no longer activated by sugar phosphates and did not significantly bind F6P as measured by affinity capillary electrophoresis (*49*). It should also be noted that ADPGlc PPases that are not activated by sugar phosphates, such as the *Rs. rubrum* and *Bacillus stearothermophilus* enzymes, have substitutions at the position corresponding to Arg33 of the *A. tumefaciens* enzyme (Figure 7). The substitution of glutamate at this position in the *Rs. rubrum* enzyme could result in repulsion of the negatively charged sugar phosphates. Another observation that suggests that the domain interface may be an allosteric site is the fact that the Glc1P thymidylyltransferase enzymes, which display high structural similarity with the catalytic domain but lack the entire parallel β-helix domain, also lack allosteric regulation (*1, 16*).

The proposed *A. tumefaciens* allosteric site has both similarities and differences with the putative potato tuber regulatory site (*4*). In the α-subunit potato tuber ADPGlc PPase structure, two sulfate ions were found in the region between domains, and it was hypothesized that these could be part of the allosteric site responsible for the activator 3PGA binding (*4*). The *A. tumefaciens* sulfate site is most similar to the potato tuber sulfate site 1, which includes residues Arg41, Arg53, Asp403, and Lys404 corresponding to Arg33, Arg45, Asp378, and His379 in the *A. tumefaciens* enzyme. In our model, Arg33 is also involved in a salt bridge with Asp378, which is analogous to the salt bridge identified in the potato tuber model between Arg41 and Asp403. The two sites differ in that the sulfate in the potato tuber enzyme is also associated with Lys441 while the *A. tumefaciens* sulfate region also includes Arg32, Tyr39, Ser72, and Arg75. It should be noted that the Lys404 and Lys441 positions have been demonstrated to be important for the apparent binding of the activator 3PGA in both the potato tuber and *Anabaena* enzymes (*20–22*). Lys441 is not conserved in bacterial ADPGlc PPases that are not activated by 3PGA. Our A. *tumefaciens* X-ray structure contains no sulfate at the position corresponding to sulfate site 2 in the potato tuber α-subunit model.

On the basis of the assumption that the sulfate ion is occupying the phosphate-binding site of F6P, we have modeled F6P in the *A. tumefaciens* ADPGlc PPase structure (Figure 8c,d). An interesting feature of the proposed F6P binding site is the large number of basic residues projecting into the binding pocket. The F6P binding pocket is approximately 10 Å (w) × 14 Å (h) × 7 Å (d). Within this region there are seven basic side chains: Arg32, Arg33, Arg45, His71, His76, Arg79, and His 379 (Figure 8c,d). Crystals were grown at pH 7.5 so all arginine residues should carry a positive charge. Only two acidic residues are near the F6P binding pocket, with Glu304 positioned at the top of the cleft and Asp378 at the bottom of the cleft. As previously mentioned, Asp378 forms a salt bridge with Arg33. The positive charge patch created by the seven basic

Table 3: Kinetic Parameters for *A. tumefaciens* ADPGlc PPase H379 Allosteric Site Mutants

	no activator		fructose 6-phosphate				fructose 1,6-bisphosphate			
	$S_{0.5}$[a] (mM)	V_{max} (units/mg)[b]	$A_{0.5}$ (mM)	$S_{0.5}$[a] (mM)	V_{max} (units/mg)[b]	FA[c]	$A_{0.5}$ (mM)	$S_{0.5}$[a] (mM)	V_{max} (units/mg)[b]	FA[c]
WT[d]	0.21 ± 0.04	12 ± 1.0	0.13 ± 0.01	0.051 ± 0.004	150 ± 4.3	12	NA	0.21 ± 0.04	12 ± 1.0	1
H379R	0.35 ± 0.06	2.6 ± 0.2	0.034 ± 0.003	0.073 ± 0.004	100 ± 1.8	39	0.074 ± 0.007	0.11 ± 0.017	57 ± 2.5	22
H379K	0.10 ± 0.02	19 ± 0.7	0.029 ± 0.004	0.073 ± 0.016	100 ± 4.3	5.3	0.043 ± 0.007	0.076 ± 0.005	58 ± 0.7	3

[a] All $S_{0.5}$ values are for the substrate ATP. [b] A unit of activity is defined here as the amount of enzyme producing 1 μmol/min ADPGlc as indicated in Materials and Methods. Standard assay conditions are as described in Table 2. [c] Fold activation is defined as the velocity in the presence of saturating F6P or FBP (V_A) divided by the velocity in the absence of activator (V_O), V_A/V_O. [d] Wild-type data from ref *3*.

residues seems likely to attract the negatively charged phosphate of F6P and provide hydrogen bonds with the oxygens of the sugar.

The proposed F6P activator-binding site is close to the ATP binding site with the closest atomic approach of the modeled F6P and ATP at ~14 Å. The peptide backbone and side chains involved in formation of the F6P binding site and the ATP binding site are involved in several interactions. A model of F6P binding influencing ATP binding can easily be drawn. The peptide backbone of residues 19−23 that abuts the ATP and has been proposed to undergo conformational change during catalysis (*4*) makes hydrogen bonds with the side chain of Thr30 and the backbone atoms of Lys35, Gln67, and Tyr68. In addition, the close proximity of Arg33 of the proposed activator site displays structural proximity to the glycine-rich loop and the catalytically important Lys35 (*4, 46*). Communication between the substrate binding site and the activator site may occur via these interactions.

Evidence for such communication between the two sites is clearly shown by the effects of mutations that affect binding. Mutations in residues that form the ATP binding site and decrease the affinity for ATP also alter activator-binding affinity. Alanine scanning mutagenesis of the active site residues G20, G21, and G23 and that previously described for R22 and R25 (*3*) all resulted in a decrease in affinity for the activator F6P (Table 2). Mutations of the *E. coli* enzyme that disrupt the activator binding [K39E, A44T (*18, 50*)] or the substrate site [Y114F (*10*)] also have been shown to decrease the corresponding affinity for substrate or activator.

Toward Engineering the Allosteric Specificity of ADPGlc PPase. The modeling of the allosteric binding site of the *A. tumefaciens* ADPGlc PPase provides insight into the molecular basis for allosteric specificity among the other bacterial ADPGlc PPases and opens up the possibility of rational engineering of effector specificity. Unlike the *A. tumefaciens* ADPGlc PPase, the *E. coli* and *Rb. sphaeroides* enzymes are activated by FBP and have an arginine at the position corresponding to His379 of the *A. tumefaciens* enzyme. His379 is positioned within 5 Å of the sulfate in our model and is adjacent to Asp378 that forms a salt bridge with Arg33. To probe the role of this residue on allosteric specificity, we generated the H379R and H379K enzymes for analysis. Table 3 displays the kinetic parameters for these altered enzymes compared to wild type with respect to the substrate ATP and effectors F6P and FBP. The major effect of placing an arginine at this position was significant activation by FBP, with an ~3-fold increase in apparent affinity for ATP and an ~22-fold stimulation of V_{max}. Similar to the wild-type enzyme, the H379R enzyme was also activated by F6P. The higher fold activation of the H379R enzyme compared to wild type is due in part to the lower observed V_{max} in the absence of activator. The substitution of the similarly charged lysine at position 379 also resulted in significant activation by FBP. The H379K enzyme is different from H379R in that the unactivated enzyme displays a higher V_{max} and lower $S_{0.5}$ value for ATP. It should be noted that the $A_{0.5}$ values for the H379R and H379K enzymes are in the same range as those reported for the *E. coli* enzyme for FBP (*50*) and the *Rb. sphaeroides* enzyme for FBP and F6P (*8*). The positive charge on H379R and H379K may interact with the additional phosphate moiety of FBP and stabilize binding and/or downstream conformational effects. These results appear to identify the residue responsible for activation by FBP in the C-terminus of an *A. tumefaciens*−*E. coli* chimeric enzyme (*19*).

Concluding Remarks. We report here for the first time a high-resolution structure of a native bacterial ADPGlc PPase. One limitation of the structure, however, is the fact that the enzyme was crystallized in the sulfate-inhibited form. Our structural and biochemical data have extended our knowledge of the active site and identified an allosteric site. Our data support the hypothesis that regions from both the N- and C-terminal domains of this enzyme family form the regulatory binding site(s). The *A. tumefaciens* enzyme allosteric site appears capable of accommodating sulfate, phosphate, and F6P and gains the function of FBP binding/activation when His379 is converted to arginine or lysine.

Our results extend a structural perspective to previous results that have demonstrated the inherent plasticity of the allosteric site. This apparent flexibility has allowed the evolution of diverse forms of the enzyme to fit specific metabolic niches (*1, 2*). In general, the allosteric site(s) of the ADPGlc PPases appear(s) to have a limited ability to bind diverse molecules with phosphate and other negatively charged moieties, which can be enhanced by mutation. Downstream conformational changes are triggered by the binding of diverse effector molecules.

Our structure will allow for modeling of other bacterial ADPGlc PPases as well as *in silico* ligand docking. This study is an important step in rational engineering of bacterial ADPGlc PPase with a range of properties that could alter starch accumulation in transgenic crops. The use of modified ADPGlc PPases in transgenic plants has already led to significant increases in starch yields and/or seed weight and biomass in potato, maize, wheat, rice, and cassava (*52–57*). The recently solved structures of the glycogen synthase (*58*) and branching enzyme (*59*) further open up the possibility of engineering the entire starch biosynthetic pathway to make designer starches in high yields.

ACKNOWLEDGMENT

We thank Dr. Takashi Kuriki (Biochemical Research Laboratory, Ezaki Glico Co.) for the generous gift of the

TGC31 cell line. We also thank Chris Berger and Steve Igwe for valuable assistance with enzyme preparations and assays.

REFERENCES

1. Ballicora, M. A., Iglesias, A. A., and Preiss, J. (2003) ADP-Glucose pyrophosphorylase, a regulatory enzyme for bacterial glycogen synthesis. *Microbiol. Mol. Biol. Rev. 67*, 213–225.
2. Ballicora, M. A., Iglesias, A. A., and Preiss, J. (2004) ADP-glucose pyrophosphorylase: a regulatory enzyme for plant starch biosynthesis. *Photosynth. Res. 79*, 1–24.
3. Gomez-Casati, D., Igarashi, R. Y., Berger, C., Brandt, M., Iglesias, A. A., and Meyer, C. R. (2001) Identification of functionally important amino terminal arginines of *Agrobacterium tumefaciens* ADPglucose pyrophosphorylase by alanine-scanning mutagenesis. *Biochemistry 40*, 10169–10178.
4. Jin, X., Ballicora, M. A., Preiss, J., and Geiger, J. H. (2005) Crystal structure of potato tuber ADP-glucose pyrophosphorylase. *EMBO J. 24*, 694–704.
5. Okita, T. W., Rodriguez, R. L., and Preiss, J. (1981) Biosynthesis of bacterial glycogen. Cloning of the glycogen biosynthetic enzyme structural genes of *Escherichia coli*. *J. Biol. Chem. 256*, 6944–6952.
6. Charng, Y.-Y., Kakefuda, G., Iglesias, A. A., Buikema, W. J., and Preiss, J. (1992) Molecular cloning and expression of the gene encoding ADP-glucose pyrophosphorylase from the cyanobacterium *Anabaena* sp. strain PCC 7120. *Plant Mol. Biol. 20*, 37–47.
7. Uttaro, A. D., Ugalde, R. A., Preiss, J., and Iglesias, A. A. (1998) Cloning and expression of the *glgC* gene from *Agrobacterium tumefaciens*: Purification and characterization of the ADPglucose synthetase. *Arch. Biochem. Biophys. 357*, 13–21.
8. Igarashi, R., and Meyer, C. R. (2000) Cloning and sequencing of glycogen metabolism genes from *Rhodobacter sphaeroides* 2.4.1. Expression and characterization of recombinant ADPglucose pyrophosphorylase. *Arch. Biochem. Biophys. 376*, 47–58.
9. Hill, M. A., Kaufmann, K., Otero, J., and Preiss, J. (1991) Biosynthesis of bacterial glycogen. Mutagenesis of a catalytic site residue of ADPglucose pyrophsophorylase from *E. coli*. *J. Biol. Chem. 266*, 12455–12460.
10. Kumar, A., Ghosh, P., Lee, Y. M., Hill, M. A., and Preiss, J. (1989) Biosynthesis of bacterial glycogen. Use of site-directed mutagenesis to probe the role of tyrosine 114 in the catalytic mechanism of ADP-glucose synthetase from *E. coli*. *J. Biol. Chem. 263*, 14634–14639.
11. Frueauf, J. B., Ballicora, M. A., and Preiss, J. (2001) Aspartate residue 142 is important for catalysis by ADP-glucose pyrophosphorylase from *Escherichia coli*. *J. Biol. Chem. 276*, 46319–46325.
12. Fu, Y., Ballicora, M. A., and Preiss, J. (1998) Mutagenesis of the glucose-1-phosphate-binding site of potato tuber ADP-glucose pyrophosphorylase. *Plant Physiol. 117*, 989–996.
13. Freauf, J. B., Ballicora, M. A., and Preiss, J. (2003) ADP-glucose pyrophosphorylase from potato tuber: site-directed mutagenesis of homologous aspartic acid residues in the small and large subunits. *Plant J. 33*, 503–511.
14. Brown, K., Pompeo, F., Dixon, S., Mengin-Lecreulx, D., Cambillau, C., and Bourne, Y. (1999) Crystal structure of the bifunctional N-acetylglucosamine 1-phophate uridyltransferase from *Escherichia coli*: a paradigm for the related pyrophosphorylase superfamily. *EMBO J. 18*, 4096–4107.
15. Olsen, L. R., and Roderick, S. L. (2001) Structure of the *Escherichia coli* GlmU pyrophosphorylase and acetyltransferase active sites. *Biochemistry 40*, 1913–1921.
16. Blankenfeldt, W., Asuncion, M., Lam, J. S., and Naismith, J. H. (2000) The structural basis of the catalytic mechanism and regulation of glucose-1-phosphate thymidylyltransferase (RmlA). *EMBO J. 19*, 6652–6663.
17. Sivaraman, J., Sauve, V., Matte, A., and Cygler, M. (2002) Crystal structure of *Escherichia coli* glucose-1-phosphate thymidylyltransferase (RffH) complexed with dTTP and Mg^{2+}. *J. Biol. Chem. 277*, 44214–44219.
18. Gardiol, A., and Preiss, J. (1990) *Escherichia coli* E-39 ADPglucose synthetase has different activation kinetics from the wild-type allosteric enzyme. *Arch. Biochem. Biophys. 280*, 175–180.
19. Ballicora, M. A., Sesma, J. I., Iglesias, A. A., and Preiss, J. (2002) Characterization of chimeric ADPglucose pyrophosphorylases of *Escherichia coli* and *Agrobacterium tumefaciens*. Importance of the C-terminus on the selectivity for allosteric regulators. *Biochemistry 41*, 9431–9437.
20. Sheng, J., Charng, Y.-Y., and Preiss, J. (1996) Site-directed mutagenesis of lysine-382, the activator-binding site, of ADP-glucose pyrophosphorylase from *Anabaena* PCC 7120. *Biochemistry 35*, 3115–3121.
21. Charng, Y.-Y., Iglesias, A. A., and Preiss, J. (1994) Structure-function relationships of cyanobacterial ADP-glucose pyrophosphorylase. *J. Biol. Chem. 269*, 24107–24113.
22. Ballicora, M. A., Fu, Y., Nesbett, N. M., and Preiss, J. (1998) ADP-glucose pyrophosphorylase from potato tubers. Site-directed mutagenesis studies of the regulatory sites. *Plant Physiol. 118*, 265–274.
23. Kavakli, I. H., Park, J.-S., Slattery, C. J., Salamone, P. R., Frohlick, J., and Okita, T. W. (2001) Analysis of allosteric effector binding sites of potato ADP-glucose pyrophosphorylase through reverse genetics. *J. Biol. Chem. 276*, 40834–40840.
24. Salamone, P. R., Kavakli, I. H., Slattery, C. J., and Okita, T. W. (2002) Directed molecular evolution of ADP-glucose pyrophosphorylase. *Proc. Natl. Acad. Sci. U.S.A. 99*, 1070–1075.
25. Ballicora, M. A., Laughlin, M. J., Fu, Y., Okita, T. W., Barry, G. F., and Preiss, J. (1995) Adenosine 5′-diphosphate-glucose pyrophosphorylase from potato tuber. Significnace of the N terminus of the small subunit for catalytic properties and heat stability. *Plant Physiol. 109*, 245–251.
26. Takata, H., Takaha, T., Okada, S., Takagi, M., and Imanaka, T. (1997) Characterization of a gene cluster for glycogen biosynthesis and a heterotetrameric ADP-glucose pyrophosphorylase from *Bacillus stearothermophilus*. *J. Bacteriol. 179*, 4689–4698.
27. Meyer, C. R., Borra, M., Igarashi, R., Lin, Y.-S., and Springsteel, M. (1999) Purification and characterization of ADPglucose pyrophosphorylase from *Rhodobacter sphaeroides* 2.4.1. Evidence for the involvement of arginine residues in allosteric regulation. *Arch. Biochem. Biophys. 372*, 179–188.
28. Cupp-Vickery, J. R., Igarashi, R. Y., and Meyer, C. R. (2005) Preliminary crystallographic analysis of ADP-glucose pyrophosphorylase from *Agrobacterium tumefaciens*. *Acta Crystallogr. F 61*, 266–268.
29. Ghosh, H. P., and Preiss, J. (1966) Adenosine diphosphate glucose pyrophosphorylase. A regulatory enzyme in the biosynthesis of starch in spinach leaf chloroplasts. *J. Biol. Chem. 241*, 4491–4504.
30. Preiss, J., Shen, L., Greenberg, E., and Gentner, N. (1966) Biosynthesis of bacterial glycogen. IV. Activation and inhibition of the adenosine diphosphate glucose pyrophosphorylase of *Escherichia coli B*. *Biochemistry 5*, 1833–1845.
31. Brooks, S. P. J. (1992) A simple computer program with statistical tests for the analysis of enzyme kinetics. *BioTechniques 13*, 906–911.
32. Leslie, A. G. W. (1998)MOSFLM v.6.0, MRC Laboratory of Molecular Biology, Cambridge, U.K.
33. Weiss, M. (2001) *J. Appl. Crystallogr. 34*, 130–135.
34. Collaborative Computational Project, Number 4 (1994) The CCP4 suite: programs for protein crystallography, *Acta Crystallogr., Sect. D 50*, 760–763.
35. Perrakis, A., Morris, R. J. H., and Lamzin, V. S. (1999) Automated protein model building combined with iterative structure refinement. *Nat. Struct. Biol. 6*, 458–463.
36. Jones, T. A. (1985) Diffraction methods for biological macromolecules. Interactive computer graphics: FRODO. *Methods Enzymol. 115*, 157–171.
37. Hutchinsom, E. G., and Thornton, J. M. (1996) PROMOTIF—A program to identify structural motifs in proteins. *Protein Sci. 5*, 212–220.
38. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK—A programme to check the stereochemical quality of protein structures. *J. Appl. Crystallogr. 26*, 283–291.
39. Hwang, S. K., Hamada, S., and Okita, T. W. (2007) Catalytic implications of the higher plant ADP-glucose pyrophosphorylase large subunit. *Phytochemistry 68*, 464–477.
40. Hwang, S.-K., Hamada, S., and Okita, T. W. (2006) ATP binding site in the plant ADP-glucose pyrophosphorylase large subunit. *FEBS Lett. 580*, 6741–6748.
41. Hwang, S. K., Salamone, P. R., and Okita, T. W. (2005) Allosteric regulation of the higher plant ADP-glucose pyrophosphorylase is a product of synergy between the two subunits. *FEBS Lett. 579*, 983–990.
42. Paule, M. R., and Preiss, J. (1971) Biosynthesis of bacterial glycogen. X. The kinetic mechanism of adenosine diphosphoglucose pyrophosphorylase from *Rhodospirillum rubrum*. *J. Biol. Chem. 246*, 4602–4609.

43. Kleczhowski, L. A., Villand, P., Preiss, J., and Olsen, O.-A. (1993) Kinetic mechanism and regulation of ADP-glucose pyrophosphorylase from barley (*Hordeum vulgare*) leaves. *J. Biol. Chem. 268*, 6228–6233.
44. Haugen, T. H., and Preiss, J. (1979) Biosynthesis of bacterial glycogen. The nature of the binding of substrates and effectors to ADP-glucose synthase. *J. Biol. Chem. 254*, 127–136.
45. Hemmer, W., McGlone, M. M., Tsigelny, I., and Taylor, S. S. (1997) Role of the glycine triad in the ATP-binding site of cAMP-dependent protein kinase. *J. Biol. Chem. 272*, 16946–16954.
46. Ballicora, M. A., Dubay, J. R., Devillers, C. H., and Preiss, J. (2005) Resurrecting the ancestral enzymatic role of a modulatory subunit. *J. Biol. Chem. 280*, 10189–10195.
47. Bejar, C. M., Jin, X., Ballicora, M. A., and Preiss, J. (2006) Molecular architecture of the glucose-1-phosphate site in ADP-glucose pyrophosphorylases. *J. Biol. Chem. 281*, 40473–40484.
48. Hill, M. A., and Preiss, J. (1998) Functional analysis of conserved histidines in ADP-glucose pyrophosphorylase from *Escherichia coli*. *Biochem. Biophys. Res. Commun. 244*, 573–577.
49. Kaddis, J., Zurrita, C., Moran, J., Borra, M., Polder, N., Meyer, C. R., and Gomez, F. A. (2004) Estimation of binding constants for the substrate and activator of *Rhodobacter sphaeroides* ADP-glucose pyrophosphorylase using affinity capillary electrophoresis. *Anal. Biochem. 327*, 252–260.
50. Meyer, C. R., Ghosh, P., Nadler, S., and Preiss, J. (1993) Cloning, expression, and sequence of an allosteric mutant of ADP-glucose pyrophosphorylase from *Escherichia coli*. *Arch. Biochem. Biophys. 302*, 64–71.
51. Charng, Y. Y., Sheng, J., and Preiss, J. (1995) Mutagenesis of an amino acid residue in the activator-binding site of cyanobacterial ADP-glucose pyrophosphorylase causes alteration in activator specificity. *Arch. Biochem. Biophys. 318*, 476–480.
52. Stark, D. M., Timmerman, K. P., Barry, G. F., Preiss, J., and Kishore, G. M. (1992) Regulation of the amount of starch in plant tissues by ADP glucose pyrophosphorylase. *Science 258*, 287–292.
53. Giroux, M. J., Shaw, J., Barry, G., Cobb, B. G., Greene, T., Okita, T., and Hannah, L. C. (1996) A single gene mutation that increases maize seed weight. *Proc. Natl. Acad. Sci. U.S.A. 93*, 5824–5829.
54. Smidansky, E. D., Clancy, M., Meyer, F. D., Panning, S. P., Blake, N. K., Talbert, L. E., and Giroux, M. J. (2002) Enhanced ADP-glucose pyrophosphorylase activity in wheat endosperm increases seed yield. *Proc. Natl. Acad. Sci. U.S.A. 99*, 1724–1729.
55. Smidansky, E. D., Martin, J. M., Hannah, L. C., Fischer, A. M., and Giroux, M. J. (2003) Seed yield and plant biomass increases in rice are conferred by deregulation of endosperm ADP-glucose pyrophosphorylase. *Planta 216*, 656–664.
56. Sakulsingharoj, C., Choi, S.-B., Hwang, S.-K., Edwards, G. E., Bork, J., Meyer, C. R., Preiss, J., and Okita, T. W. (2004) Engineering starch biosynthesis for increasing rice seed weight: The role of cytoplasmic ADP-glucose pyrophosphorylase. *Plant Sci. 167*, 1323–1333.
57. Ihemere, U., Arias-Garzon, D., Lawrence, S., and Sayre, R. (2006) Genetic modification of cassava for enhanced starch production. *Plant Biotechnol. J. 4*, 453–465.
58. Buschiazzo, A., Ugalde, J. E., Guerin, M. E., Shepard, W., Ugalde, R. A., and Alzari, P. M. (2004) Crystal structure of glycogen synthase: homologous enzymes catalyze glycogen synthesis and degradation. *EMBO J. 23*, 3196–3205.
59. Abad, M. C., Binderup, K., Rios-Steiner, J., Arni, R. K., Preiss, J., and Geiger, J. H. (2002) The X-ray crystallographic structure of *Escherichia coli* branching enzyme. *J. Biol. Chem. 277*, 42164–42170.
60. Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD—Visual Molecular Dynamics. *J. Mol. Graphics 14*, 33–38.

BI701933Q