Crystal Structure of an Archaeal Glycogen Synthase

INSIGHTS INTO OLIGOMERIZATION AND SUBSTRATE BINDING OF EUKARYOTIC GLYCOCEN SYNTHASES

Cristina Horcajada§, Joan J. Guinovart§, Ignacio Fita*, and Juan C. Ferrer*†‡

From the §Departament de Bioquímica i Biologia Molecular, Universitat de Barcelona, the †Institut de Recerca Biomèdica, Parc Científic de Barcelona, Universitat de Barcelona, and the *Institut de Biologia Molecular de Barcelona, Consejo Superior de Investigaciones Científicas-Parc Científic de Barcelona, 08028 Barcelona, Spain

Glycogen and starch synthases are retaining glycosyltransferases that catalyze the transfer of glucosyl residues to the non-reducing end of a growing α-1,4-glucan chain, a central process of the carbon/energy metabolism present in almost all living organisms. The crystal structure of the glycogen synthase from Pyrococcus abyssi, the smallest known member of this family of enzymes, revealed that its subunits possess a fold common to other glycosyltransferases, a pair of βαβ Rossmann fold-type domains with the catalytic site at their interface. Nevertheless, the archaeal enzyme presents an unprecedented homotrimeric molecular arrangement both in solution, as determined by analytical ultracentrifugation, and in the crystal. The C-domains are not involved in intersubunit interactions of the trimeric molecule, thus allowing for movements, likely required for catalysis, across the narrow hinge that connects the N- and C-domains. The radial disposition of the subunits conveys on the molecule a distinct triangular shape, clearly visible with negative staining electron microscopy, in which the upper and lower faces present a sharp asymmetry. Comparison of bacterial and eukaryotic glycogen synthases, which use, respectively, ADP or UDP glucose as donor substrates, with the archaeal enzyme, which can utilize both molecules, allowed us to propose the residues that determine glucosyl donor specificity.

Glycogen (starch in plants) is a polymer of α-1,4- and α-1,6-linked glucose units that provides a readily available source of energy in living organisms of the three domains: archaea, bacteria, and eukarya. Polymerization is performed by specific glycosyltransferases (GT) of the GT-3 and GT-5 families, named glycogen (GS) or starch synthases, which catalyze the formation of α-1,4-glycosidic bonds using UDP- or ADP-glucose as the glucosyl donor. The α-1,6-linked branches are introduced in the synthetic direction and eliminated in the catabolic pathway by various transglycosidases of the GH-13 family. Glycogen can be degraded phosphorylolytically by glycogen phosphorylases (GT-35) or hydrolytically by α-amylases (GH-13) (1). Few organisms are devoid of glycogen-active enzymes, and in bacteria it has been suggested that the absence of glycogen metabolism is a marker of host-dependent (parasitic or symbiotic) behavior. Thus, the capacity to synthesize and utilize glycogen appears to be central for the survival of “free-living” bacteria (1). In higher eukaryotes, glycogen and glycogen-metabolizing enzymes are present in almost every cell type, again pointing to the crucial role of this metabolic pathway in the carbon/energy metabolism of the cell.

Animal and fungal GS are grouped in family GT-3 on the basis of a high degree of sequence identity (45–50%) (2, 3). GT-3 enzymes have two other characteristics in common: they use UDP glucose as the glucosyl donor, and they are tightly regulated by reversible phosphorylation and by allosteric effectors, mainly glucose-6-phosphate. The structural information available for these enzymes is limited (to date, no enzyme of the GT-3 family has been structurally solved) and conflicting with regard to their oligomerization state. Different studies conclude that mammalian GS can exist either as dimers, trimers, or tetrarsers (4–6), while the fungal GS from Neurospora crassa has been reported to be a trimer (7).

In contrast to animal/fungal GS, bacterial GS and plant starch synthase, which are classified in the GT-5 family, are non-regulated enzymes and use ADP glucose exclusively as the source of glucose moieties. The first three-dimensional structure of an enzyme of this family has recently been reported (8). The bacterial GS from Agrobacterium tumefaciens (AAGS) is a homodimeric protein in which each subunit presents the characteristic GT-B fold, composed of two Rossmann fold domains (9).

Although the primary sequence identity of archaeal and bacterial GS (20–40% similarity) has led to their classification in the GT-5 family, the archaeal enzymes share several structural features with their animal/fungal counterparts of the GT-3 family (10, 11). Furthermore, archaeal GS use both ADP- and UDP-glucose as glucosyl donors with similar efficiency (12–14). These observations are not unexpected given the higher phylogenetic proximity of eukaryotic to archaeal enzymes in comparison with their bacterial counterparts (15).

Here we report the crystallographic structure of the archaeal GS from Pyrococcus abyssi, which is the smallest known enzyme of the glycogen/starch synthase superfamily (GT-3 and GT-5).

EXPERIMENTAL PROCEDURES

Protein Expression, Crystallization, and Data Collection—A His-tagged form of the full-length PaGS, which was expressed in Escherichia coli as inclusion bodies, was purified and refolded as described (14). Small needle-shaped crystals were obtained using the vapor diffusion method with 100 mM sodium citrate, pH 5.6, containing 20% polyethylene glycol 4000 and 20% dioxane. Microseeding was performed by crushing these crystals with the seed bead kit (Sigma) and adding them to new drops containing 15% polyethylene glycol 4000 and 20% dioxane.
Crystal Structure of an Archaeal Glycogen Synthase

The larger needles produced were further enlarged by macroseeding in the same buffer until \(0.01 \times 0.01 \times 1.0\) mm\(^3\) in size. Finally, the crystals were frozen in liquid nitrogen in the presence of a cryoprotectant solution (the same mother liquor except for 20% glycerol). A complete X-ray diffraction data set was collected at 100 K (beamline ID4-EH2, European Synchrotron Radiation Facility, Grenoble) and processed using DENZO/SCALEPACK (16) and TRUNCATE4.1 (17).

The cDNA encoding the C-domain of PaGS (amino acids Gly-218 to Ser-413) was amplified by PCR using specific primers and cloned into the plasmid pET28a+ (Novagen), which adds a His tag and a thrombin cleavage site at the N terminus. A culture of E. coli BL21-CodonPlus (DE3)-RIL (Stratagene) transformed with the constructed plasmid was grown in Luria Bertani medium supplemented with 34 \(\mu\)g ml\(^{-1}\) chloramphenicol and 10 \(\mu\)g ml\(^{-1}\) kanamycin. After 24 h at 37 °C, cells were harvested and resuspended in a buffer containing 500 mM NaCl, 10 mM imidazole, 1 mM phenylmethylsulfonyl fluoride, and 1 mM dithiothreitol in 20 mM sodium phosphate, pH 7.4. After sonication, the suspension was centrifuged for 20 min at 10,000 \(\times\) g and loaded onto a HisTrap nickel-chelating column (Amersham Biosciences). The recombinant protein was eluted with a linear gradient of 10–500 mM imidazole, the positive fractions were pooled, and the buffer was replaced by 10 mM CaCl\(_2\)/50 mM Tris-HCl, pH 8. The His tag was cleaved by incubating the sample with a thrombin-agarose suspension (Sigma) for 12 h at room temperature. To separate the cleaved and uncleaved portions, the solution was reloaded onto a new HisTrap nickel-chelating column, and the unbound fractions were collected. After loading onto a Superdex 200 HR10/30 gel filtration column (Amersham Biosciences), the protein was eluted with a buffer containing 500 mM NaCl and 1 mM dithiothreitol in 20 mM Tris-HCl, pH 8, as a single peak and was brought to a final concentration of 10 mg ml\(^{-1}\). The purified protein ran as a single band of \(\sim 25\) kDa on an SDS-polyacrylamide gel.

Crystals were obtained using the sitting drop vapor diffusion method at 4 °C, with 19% polyethylene glycol 8000 and 50 mM Li\(_2\)SO\(_4\) as precipitants, buffered in 100 mM sodium acetate, pH 4.6. The crystals, which grew to \(0.1 \times 0.1 \times 0.1\) mm\(^3\) in size in 3–4 days, had a cubic morphology and were monoclinic (space group P2\(_1\)). A complete X-ray diffraction data set was collected at 100 K from a single crystal that had been frozen in liquid nitrogen in a cryogenic solution (20% glycerol), (beamline BM16, European Synchrotron Radiation Facility). Data were processed and scaled using MOSFLM6.2.0 and SCALA2.7.5 (17).

**Structure Determination and Refinement**—All attempts to determine the PaGS structure using heavy atom or SeMet derivatives were unsuccessful. After the ArGS structure became available (8), molecular replacement trials were attempted that included the use of the locked rotation and locked translation functions (18). Because this approach also failed, an alternative and cumbersome scheme that consisted of first solving one of the two PaGS domains was planned. For the crystals of the PaGS C-terminal domain, molecular replacement gave an unambiguous solution for a polyalanine model derived from the C-domain of AtGS, with one domain in the asymmetric unit. Phases of the initial solution were improved by rigid body refinement (REFMAC5) (17) and density modification (17). Model completion was achieved by alternating steps of automatic refinement and manual rebuilding, using the graphic program O (19). Refinement was performed with all the diffraction data available (25–1.8 Å), except the 5% randomly selected data, to calculate the \(R_{	ext{free}}\). B-factor restraints were gradually loosened as refinement progressed. Solvent positions were assigned using ARP/wARP (17) and cross-checked manually. The density map also showed the presence of two acetate and four sulfate ions from the crystallization buffer. The stereochemical quality of the model, analyzed with PROCHECK (20), indicated that the main chain conformational angles were all located within favored regions of the Ramachandran diagram.

Molecular replacement with the full-length PaGS, using its C-domain as a starting partial model, was then attempted. A possible solution was obtained at 3.5 Å resolution, corresponding to three C-domains in the asymmetric unit. To complete each subunit of the trimer, a second search was performed using as a model the N-domain of AtGS with all the side chains converted to Ala. Graphic evaluation with the program O (19) confirmed that the connectivity between domains and the global packing were correct. Initial rigid body refinement was performed with REFMACS (17), assuming six rigid bodies (the N- and C-terminal domains of each monomer). Phases were improved by density modification using solvent flattening, histogram matching, and three-dimensional fold non-crystallographic symmetry averaging. Graphic model building was alternated with automatic refinement, excluding 5% of randomly selected data to calculate the \(R_{	ext{free}}\). TLS refinement was performed assuming two rigid bodies (N- and C-domains)/monomer. The model showed good stereochemistry with only 1 residue (Asp-128) in a disallowed region of the Ramachandran plot. Statistics for data collection and refinement are summarized in Table 1.

**Sedimentation Equilibrium Centrifugation**—Sedimentation equilibrium experiments were performed using an Optima XL-A analytical ultracentrifuge (Beckman). Protein solutions (1, 0.5, 0.01 mg ml\(^{-1}\)) were centrifuged at 10,000 \(\times\) g at 10 °C until equilibrium was achieved. The molecular weight of PaGS was estimated by fitting a theoretical curve to the observed radial distribution of the protein concentration in the centrifugation cell, using the program Origin v6.1 (OriginLab).

**Transmission Electron Microscopy**—Samples for transmission electron microscopy were prepared using the conventional negative staining procedure. A 20-µl drop of protein solution (0.2 mg ml\(^{-1}\)) was adsorbed onto a glow-ionized carbon-coated copper grid, washed with two drops of deionized water, and stained with two drops of freshly prepared 2% uranyl acetate solution. Samples were imaged at room temperature using a Jeol JEM 1010 transmission electron microscope, working at 80 KV. Images were collected at a \(\times300,000\) magnification.

**Sequence Alignments**—Multiple sequence alignments of GT-3 and GT-5 enzymes were performed with the program CLUSTALW v1.83 (21). Combined structural and multiple sequence alignments were performed using the program ESPript (22).

**RESULTS AND DISCUSSION**

**Monomer Folding**—The crystal structure of P. abyssi GS (PaGS) has been determined and refined to a resolution of 2.8 Å. The crystal asymmetric unit contains three monomers, each comprising residues from Met-1-Leu-437. Because conventional methods to solve the whole protein proved unsuccessful, the structure of the isolated C-terminal domain was first determined by molecular replacement from the equivalent domain of the bacterial AtGS (8). The atomic coordinates of this fragment (residues Gly-218-Ser-413 plus an N-terminal extension of 3 residues), refined at 1.8 Å resolution, were then used to obtain a partial solution for the full-length protein. Structure determination was completed by taking advantage of the non-crystallographic 3-fold symmetry of the molecule (see below).

Similar to what has been described for the bacterial AtGS (8) and other enzymes of the GT-B fold superfamily (9), PaGS subunits fold into a pair of Rossmann fold domains separated by a deep cleft in which the catalytic center is located (Fig. 1A). The N-domain of PaGS (residues 1–217 and 414–437) presents a root mean square deviation (r.m.s.d.) of 2.4 Å for the Ca atoms of the 219 residues (79%) having equivalents in the AtGS structure. In the C-domain (residues 218–413), the r.m.s.d.
for 183 structurally equivalent residues (86%) between PaGS and AtGS is 1.3 Å. Sequence identities between the two proteins are 21 and 22% for the N- and C-domains, respectively. The narrow hinge that connects the N- and C-terminal domains is formed by only two chain segments (Fig. 1A), which comprise highly conserved residues among the members of the GT-3 and GT-5 families (Fig. 2). In PaGS, these residues, which present extended conformations, correspond to Asn-217-Gly-218-Ile-219, connecting the N- and C-terminal domains, and Phe-412-Ser-413-Trp-414, between the terminal helices $H_9$-$H_{10}$-$H_{11}$, respectively.

The structure of the isolated C-domain of PaGS was essentially identical to that of the C-terminal domain in the full-length protein (r.m.s.d. of 0.71 Å) (Fig. 1B), with only three noticeable differences. The loop

| TABLE 1 Structural data |
|--------------------------|
| **Data collection**       | C-domain | PaGS |
| Space group              | $P_2_1$  | $C_2$ |
| Cell dimensions (Å, °)   | a = 32, b = 52, c = 42, $\beta = 105$ | a = 203, b = 74, c = 148, $\beta = 131$ |
| Resolution (Å)$^a$       | 50–1.8 (1.9–1.8) | 30–2.8 (2.9–2.8) |
| Unique reflections       | 15,701 | 41,544 |
| Completeness (%)$^a$     | 97.6 (85.9) | 99.9 (99.8) |
| Mosaicity (%)$^a$        | 0.4 | 0.6 |
| $R_{merge}$ (%)$^{ab}$   | 4.3 (24.8) | 6.6 (45) |
| Multiplicity$^a$         | 2.4 (2.0) | 3.7 (3.7) |
| ($I/I_0$)$^a$            | 15.1 (3.5) | 14.1 (1.8) |

| Refinement statistics     |       |       |
| Resolution range (Å)      | 40–1.8 | 20–2.8 |
| Reflections (%)$^a$       | 14581 (833) | 38891 (2822) |
| R Factor ($R_{free}$) (%)$^a$ | 18.3 (23.5) | 20.0 (26.6) |
| Average B (Å²)            | 15.64 | 49.0 |
| Number of atoms           | 1866 | 10569 |
| Protein atoms             | 1669 | 10449 |
| Acetate molecules         | 2 | 0 |
| Sulfate molecules         | 4 | 0 |
| Glucose molecules         | 0 | 2 |
| UDP molecules             | 0 | 1 |
| Dioxane molecules         | 0 | 9 |
| Glycerol molecules        | 0 | 7 |
| Water molecules           | 198 | 0 |

| R.m.s.d.                  |       |       |
| Bond lengths (Å)          | 0.008 | 0.012 |
| Bond angles (°)           | 0.98 | 1.4 |

| Ramachandran statistics (%)$^{ac}$ |
|-------------------------------------|
| Most favored regions               | 91.8 | 88.7 |
| Additional allowed regions         | 8.2 | 10.8 |
| Generously allowed regions         | 0.0 | 0.3 |
| Disallowed regions                 | 0.0 | 0.2 |

$^a$ Values in parentheses are for the highest resolution shell.

$^b$ $R_{merge} = \frac{\sum_{hkl}(I_hkli_hkli_hk)}{\sum_{hkl}I_hkli_hk} - \frac{(\langle hkl \rangle)}{\sum_{hkl}I_hkli_hk}$.

$^c$ As evaluated by PROCHECK (20).

**FIGURE 1.** Structure of a PaGS monomer and the isolated C-terminal domain. A, general fold of a PaGS monomer shown in two orientations related by 90°. The color ramp from blue to red indicates N- to C-terminal direction. B, stereographic drawing of the superimposed structures of the isolated C-domain (blue) and the same domain in the full-length enzyme (orange).
FIGURE 2. Structurally weighted sequence alignment of GS from P. abyssi, A. tumefaciens, rabbit muscle and liver, C. elegans, N. crassa, and yeast Gsy2p. Structurally equivalent regions are boxed, invariant residues are shown in white letters on a red background, and conserved amino acids are shown in red. Secondary structural elements of PaGS (α-helices and β-strands) are shown above its sequence. Several amino acids are marked with triangles: catalytic site (red), charge-relay-like system (dark blue), trimerization hydrophobic tail (orange), and complementary cavity (green). Segments connecting the N- and C-terminal domains (pink), residues that interact with glucose moieties in subsites 1 and 2 (light blue). The region responsible for the dimerization of AtGS is underlined in green.
connecting the α16 and α17 helices and that between the β15 strand and the α12 helix (Fig. 2) show small differences that, in the latter case, are probably due to the presence of a buffer-derived acetate ion in the nucleotide binding site of the isolated C-terminal domain (not shown). This observation is in agreement with what has been described for AtGS, in which the equivalent loop slightly changes its main chain trace upon ADP binding (8). Finally, Cys-221 and Cys-350 form a disulfide bridge in the full-length protein that is not observed in the structure of the isolated PaGS C-domain.

Oligomeric Structure—Despite the structural similarities between the subunits of the archaeal and bacterial enzymes, PaGS presents a homotrimeric organization, whereas AtGS is a dimeric protein (8) (Fig. 3A). In PaGS, the last α-helix of the C-domain (α17) undergoes a kink at position Ser-413, a common feature of GT-B enzymes, whereas the remaining residues (Trp-414–Thr-426) crossover to the N-domain and recover the α-helix structure for 12 more residues (α18). However, helix α18 of PaGS is continued by an extended and protruding tail, composed mostly by hydrophobic residues (Gly-427-Leu-437), which interacts with a hydrophobic pocket in the N-domain of a neighbor subunit. The total contact area between two subunits is 1410 Å², of which ~820 Å² correspond to hydrophobic contacts. The main interactions are established with amino acids Lys-53-Ile-54-Arg-55, several residues from Arg-105 to Leu-115, and Tyr-141-Phe-142 (Figs. 2 and 3B), resulting in the three N-domains of the PaGS trimer being rigidly bound to each other. In contrast, the C-domains do not participate in intersubunit interactions (Fig. 3A), and consequently the trimeric organization of the PaGS molecule does not restrain movements across the hinge that connects the two domains (Fig. 1A). In fact, although two of the three N-domains are related by an accurate molecular 3-fold axis (the slight deviation of the third N-domain is attributed to crystallographic contacts), the C-domains present clear deviations from a 3-fold symmetry that result in variability in the aperture of the interdomain cleft among the three monomers of the molecule (Fig. 4 and Table 2). For a number of enzymes of the GT-B superfamily, it has been proposed that an inac-
Crystal Structure of an Archaeal Glycogen Synthase

The more open upper face) to the molecular 3-fold axis, gives conspicuous triangular shape with a side length of crystal structure of 40 Å (Fig. 5A). The arrangement of the subunits in the trimer, with the longest subunit axis almost perpendicular (~110° for the more open upper face) to the molecular 3-fold axis, gives PaGS a conspicuous triangular shape with a side length of ~130 Å and a thickness of 40 Å (Fig. 5A). The trimeric molecular organization found in the crystal structure of PaGS was fully supported both by sedimentation equilibrium ultracentrifugation, which provided a molecular mass of 158 kDa for the oligomer in solution (monomer MW, 51144 Da), and by negative staining electron microscopy (Fig. 5B).

Oligomerization of Glycogen Synthases—The structural sequence alignment of AtGS and PaGS indicates that the striking differences in molecular organization arise from only a few very specific changes in their sequences (Fig. 2). In particular, with respect to PaGS, the C-domain of AtGS has an insertion of ~10 residues (407–416) that dominates the interactions between the two AtGS subunits in the asymmetric unit. In turn, AtGS lacks the few hydrophobic residues at the C-terminal end that are critical for the trimeric oligomerization of PaGS (Fig. 3A).

Sequence alignments (not shown) indicate that most bacterial GS possess the dimerization insertion and consequently are expected to be dimers. In contrast, this segment is absent in most (apparently all) archaeal GS, which instead have the hydrophobic tail, indicating they likely are organized as trimers. For the larger eukaryotic GS of the GT-3 family, the analysis is more complex. The dimerization helix of AtGS is present, at most, only partly in the eukaryotic enzymes and the amino acid sequence of this segment is not conserved. In contrast, these enzymes hold several structural features that could be indicative of a trimeric oligomerization, in particular the hydrophobic peptide equivalent to the C-terminal tail of PaGS. Although the amino acid identity in this region between the archaeal and the eukaryotic enzymes is not high, most of the changes are conservative (Fig. 2).

The oligomerization status of GS from distinct sources is a long-standing debate that dates from the early 70s. Different authors have concluded that mammalian GS can exist either as dimers, trimers, or tetramers (4–7). It must be noted that in most of these studies the molecular mass of the oligomer was estimated by determining its sedimentation coefficient, either in buffered solutions or in sucrose gradients. From this coefficient, and assuming that the protein had a globular form, a molecular mass was calculated. Large errors can arise in this determination when the overall shape of the protein differs greatly from globular, such as the triangular form that these molecules would have if they exhibited an arrangement similar to PaGS. The two studies that report a trimeric oligomerization for rabbit muscle (23) and rat liver GS (24) were performed by sedimentation equilibrium centrifugation, the technique that was used in the present study to determine the molecular mass of PaGS in solution and that does not depend on the shape of the molecule. A much more recent report provides additional indirect evidence that muscle GS can exist as a trimer (25). In this study the authors show that in rabbit muscle fibers depleted of glycogen GS accumulates in spherical structures in which the enzyme exhibits a triangular lattice pattern, indicative of a repetitive pseudocrystalline structure, when observed in the transmission electron microscope. The calculated side length of the triangle is ~180 Å, somewhat larger than the 130 Å determined from the crystal structure of PaGS (Fig. 5A). This difference may simply reflect the larger molecular mass of the muscle enzyme. On the basis of the above evidence, we hypothesize that, at least in certain conditions, eukaryotic GS present a trimeric organization similar to that of PaGS.

Glucosyl Donor Binding Site—PaGS (14), as well as other archaeal GS (12, 13), can use ADP- or UDP-glucose as glucosyl donor substrates. In turn, plant starch synthase and bacterial GS (GT-5) are specific for ADP-glucose, whereas yeast or mammalian enzymes (GT-3) utilize exclusively UDP-glucose. In the open form of AtGS, ADP was found bound to a pocket in the C-domain wall of the catalytic crevice (8). Specificity for the adenine nucleotide was determined apparently by only two weak hydrogen bonds between nitrogens N1 and N6 of the heterocycle with the protein backbone (the carbonyl and amide groups of Gly-353 and Asn-355, respectively) and stacking interactions with Tyr-354 (Fig. 6A). The ADP phosphate groups hydrogen bonded to the side chains of Arg-299 and Lys-304, which in turn formed a salt bridge with Glu-376 (Fig. 6A).

After superimposing the AtGS and PaGS C-terminal domains, we used the ADP atomic coordinates in the AtGS-ADP complex (8) to position an ADP moiety in the PaGS donor site. The N6 and N1 nitrogens of the adenine base were located within hydrogen-bonding dis-

---

**TABLE 2**

Superposition of the N- and C-terminal domains of the PaGS subunits

| Subunits | N-terminal domain | C-terminal domain |
|----------|------------------|------------------|
|          | R.m.s.d. | Angle X | R.m.s.d. | Angle X |
| A–B      | 0.26    | 118.2   | 0.42    | 113.7   |
| A–C      | 0.27    | 120.1   | 0.45    | 121.9   |
| B–C      | 0.25    | 120.7   | 0.50    | 124.4   |
tance of the backbone carbonyl group of Glu-315 and the amide of Leu-317, respectively (Fig. 6B). This observation and the strict conservation of the residues that interact with the phosphate groups of the glucosyl donor (Arg-257, Lys-263, Glu-339 in PaGS) (Fig. 2) suggest that the ADP binding mode is essentially identical in both enzymes (Fig. 6, A and B).

A UDP molecule was manually modeled in the donor binding pocket of ArGS to optimize hydrogen-bonding interactions of the uracil base.
with the backbone amide groups of Lys-291 and Leu-317 and the carbonyl of Glu-315 (Fig. 6C). This model predicts that the nucleotide binding site of PaGS allows for specific hydrogen-bonding patterns of a uracil or an adenine with only small displacement between the locations of the two bases (Fig. 6, B and C). The added flexibility of the Met-316 side chain, which adopted two distinct conformations both in the crystal structures of PaGS and its isolated C-terminal domain (not shown), compared with the aromatic Tyr or Phe residues found in this position in all the ADP-glucose-specific GS (Fig. 2), would facilitate stacking interactions in alternative positions. In PaGS, hydrophobic interactions with the uracil base would also be contributed by the aliphatic part of Lys-291 side chain. In agreement with this model, residues equivalent to Lys-291 are generally small in GS specific for the larger ADP-glucose (an Ala in AtGS), whereas they are bulky, often a Phe, in eukaryotic GS specific for UDP-glucose (Fig. 2).

Catalytic Mechanism—Despite the very low sequence identity, a clear topological and structural resemblance was found between AtGS and other GT-B enzymes that also operate with retention of configuration of the transferred sugar, such as maltodextrin phosphorylase and the core of glycogen phosphorylase (both enzymes of the GT-35 family) (8). All the maltodextrin phosphorylase residues that make polar or hydrophobic interactions with glucosyl units in subsites −1, +1, and +2 (26) and that are also found in AtGS (8), are present in PaGS as well (Fig. 2), indicating that the binding mode of the +1 and +2 glucosyl residues of the glycogen acceptor substrate is highly similar in these enzymes. Furthermore, the strong structural similarities of the maltodextrin phosphorylase and glycogen phosphorylase active centers with those of E. coli trehalose synthase (OtsA, GT-20) and AtGS have led to the notion that all these enzymes may operate through a common catalytic mechanism (8, 27). However, the reaction mechanism of GS should differ from that proposed for glycogen phosphorylase (26, 28), in which the formation of a new α-1,4-glycosidic bond is initiated by a proton transfer from the 5′-phosphate group of the bound pyridoxal phosphate cofactor to the phosphate moiety of the donor substrate, glucose-1-P. Such transference is not feasible in the case of GS, because the phosphate group distal to the sugar moiety of ADP- or UDP-glucose, which is the structural equivalent of the 5′-phosphate of pyridoxal phosphate in glycogen phosphorylase, is likely to be deprotonated. Whereas the second ionization pKₐ of the pyridoxal phosphate is around neutrality, that of the first and unique ionization of the distal phosphate of ADP- or UDP-glucose is around 2.

The His residue in position 151 of PaGS is invariant in the GT-3 and GT-5 families (His-163 in AtGS) (Fig. 2) and is also present in glycogen phosphorylase (His-377) and trehalose synthase (His-154) (27). It has been proposed that the main chain carbonyl oxygen of this residue participates in catalysis (27, 29). However, the null activity of the H163A variant of AtGS (8) indicates that, at least in the GS family, the imidazole side chain has more than only a structural role. In fact, in the crystal of PaGS we have found a peculiar pattern of hydrogen-bonding interactions of the His-151 side chain with His-125, His-127, Thr-149, Thr-191, and Tyr-425 (Fig. 7), which is reminiscent of a charge-relay system and is strongly preserved in enzymes of the GT-3 and GT-5 families (Fig. 2). Such complex hydrogen-bonding network may have a critical role in catalysis, as it is also suggested by the observation that the H167A variant of the yeast Gsy2p, a residue isostructural to His-127 in PaGS, possesses only 1% of the activity shown by the wild-type enzyme (30). The absence of a similar charge-relay-like system in glycogen phosphorylase and maltodextrin phosphorylase and, furthermore, the observation that His-334 of the starch phosphorylase from Corynebacterium callunae (GT-35), which is the structural equivalent to His-151 in PaGS, is not essential for catalysis (31) are also indicative of the non-total equivalence between the glycogen phosphorylase and GS catalytic mechanisms.

CONCLUSION

Although the subunits of the archaeal GS from P. abyssi present the characteristic GT-B fold, the enzyme is a homotrimeric protein. The contacts between subunits, which only involve the N-terminal domains, do not restrict the movements of the corresponding C-domains, likely required for catalysis. Structurally weighted sequence alignments suggest that eukaryotic GS may possess a similar trimeric arrangement.

Acknowledgments—We thank A. Vagin and I. Navaza for valuable help with the attempts to solve the PaGS structure by molecular replacement. L. Ribas de Pouplana and C. López-Iglesias for assistance with the analytical centrifugation and electronic microscopy experiments, respectively, and P. Alzari for providing the AtGS coordinates before publication. We thank T. Yates for correcting the English version of the manuscript.

REFERENCES
1. Henrissat, B., Deleury, E., and Coutinho, P. M. (2002) Trends Genet. 18, 437–440
2. Campbell, J. A., Davies, G. J., Bulone, V., and Henrissat, B. (1997) Biochem. J. 326, 929–939
3. Coutinho, P. M., and Henrissat, B. (1999) http://afmb.cnrs-mrs.fr/CAZY/ (Carbohydrate-Active Enzymes server)
4. Jett, M. F., and Soderling, T. R. (1979) J. Biol. Chem. 254, 6793–6745
5. Takeda, Y., Brewer, H. R., Jr., and Larner, J. (1975) J. Biol. Chem. 250, 8943–8950
6. Wittsüsswanakul, D., and Kim, K. H. (1977) J. Biol. Chem. 252, 7812–7817
7. Takahara, H., and Matsuda, K. (1978) Biochim. Biophys. Acta 522, 363–374
8. Buschiazzo, A., Ugalde, J. E., Guerin, M. E., Shepard, W., Ugalde, R. A., and Alzari, P. M. (2002) EMBO J. 21, 3196–3205
9. Coutinho, P. M., Deleury, E., Davies, G. J., and Henrissat, B. (2003) J. Mol. Biol. 328, 307–317
10. Cid, E., Geremia, R. A., Guinovart, J. I., and Ferrer, J. C. (2002) FERS Lett. 528, 5–11
11. Cid, E., Gomis, R. R., Geremia, R. A., Guinovart, J. I., and Ferrer, J. C. (2000) J. Biol. Chem. 275, 33614–33621
12. Gruyer, S., Legin, E., Bliard, C., Ball, S., and Duchiron, F. (2002) Curr. Microbiol. 44, 206–211
13. Zee, C. J., MacDonell, S. W., and Pohl, N. L. (2003) J. Am. Chem. Soc. 125, 13666–13667
14. Horcajada, C., Cid, E., Guinovart, J. I., Verdugue, N., and Ferrer, J. C. (2003) Acta Crystallogr. Sect. D Biol. Crystallogr. 59, 2322–2324
15. Woese, C. R., Kandler, O., and Wheelis, M. L. (1990) Proc. Natl. Acad. Sci. U. S. A. 87, 4576–4579
16. Otwinowski, Z., and Minor, W. (1997) Methods Enzymol. 276, 307–326
17. Collaborative Computational Project Number 4 (1994) Acta Crystallogr. Sect. D Biol. Crystallogr. 50, 760–763
18. Tong, L., and Rossman, M. G. (1997) Methods Enzymol. 294, 591–611
19. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard (1991) Acta Crystallogr. Sect. A 47, 110–119
20. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) J. Appl.
Crystal Structure of an Archaeal Glycogen Synthase

21. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) Nucleic Acids Res. 22, 4673–4680
22. Gouet, P., Courcelle, E., Stuart, D. I., and Metoz, F. (1999) Bioinformatics 15, 305–308
23. Brown, N. E., and Larner, J. (1971) Biochim. Biophys. Acta 242, 69–80
24. Lin, D. C., and Segal, H. L. (1973) J. Biol. Chem. 248, 7007–7011
25. Prats, C., Cadefau, J. A., Cussó, R., Qvortrup, K., Nielsen, J. N., Wojtaszewski, J. F., Hardie, D. G., Stewart, G., Hansen, B. F., and Ploug, T. (2005) J. Biol. Chem. 280, 23165–23172
26. Geremia, S., Campagnolo, M., Schinzel, R., and Johnson, L. N. (2002) J. Mol. Biol. 322, 413–423
27. Watson, K. A., McCleverty, C., Geremia, S., Cottaz, S., Driguez, H., and Johnson, L. N. (1999) EMBO J. 18, 4619–4632
28. Gibson, R. P., Turkenburg, J. P., Charnock, S. J., Lloyd, R., and Davies, G. J. (2002) Chem. Biol. 9, 1337–1346
29. Mitchell, E. P., Withers, S. G., Ermert, P., Vasella, A. T., Garman, E. F., Oikonomakos, N. G., and Johnson, L. N. (1996) Biochemistry 35, 7341–7355
30. Pederson, B. A., Cheng, C., Wilson, W. A., and Roach, P. J. (2000) J. Biol. Chem. 275, 27753–27761
31. Schwarz, A., Pierfederici, F. M., and Nidetzky, B. (2005) Biochem. J. 387, 437–445