Escherichia coli glycogen synthase (EcGS, EC 2.4.1.21) is a retaining glycosyltransferase (GT) that transfers glucose from adenosine diphosphate glucose to a glucan chain acceptor with retention of configuration at the anomeric carbon. EcGS belongs to the GT-B structural superfamily. Here we report several EcGS X-ray structures that together shed considerable light on the structure and function of these enzymes. The structure of the wild-type enzyme bound to ADP and glucose revealed a 15.2° overall domain-domain closure and provided for the first time the structure of the catalytically active, closed conformation of a glycan synthase. The main chain carbonyl group of His-161, previously thought to be catalytic is found on the α-face of the glucose and plays an electrostatic role in the active site and as a glucose ring locator. This is also consistent with the structure of the EcGS(E377A)-ADP-HEPPSO complex where the glucose moiety is either absent or disordered in the active site.

The biosynthesis of starch in plants and glycogen in bacteria is accomplished in three steps (1, 2). The first step is the production of adenosine diphosphate glucose (ADP-Glc) building blocks from glucose 1-phosphate and ATP in a reaction catalyzed by ADP-glucose pyrophosphorylase (3, 4). The second step produces linear α,1,4-linked chains of glucose units and is catalyzed by various starch (SS)3 or glycogen synthases (GS) (2, 5), whereas the final step produces α,1,6-linked glucan branches in the polymer and is catalyzed by branching enzymes (6, 7). Bacterial GSs and plant SSs share significant (30–34%) sequence homology over most of their length, consistent with a similar overall fold and enzymatic mechanism (8).

GSs and Ss are members of the GT-B superfamily of glycosyl transferases (8). GT-B superfamily members share a common double Rossmann-fold domain architecture (8, 9). Glycosyl transferases and hydrolyzes are categorized as either retaining or inverting enzymes based on the stereochemical result of the reaction (8). Although the inverting enzymes usually undergo reaction via an oxocarbenium ion-like transition state and are reasonably well understood mechanistically (10, 11), the mechanism(s) employed by the retaining enzymes are not as well understood. Although some retaining enzymes clearly use a double displacement mechanism, where the attack of an enzyme nucleophile results in a covalent intermediate (12), many of these enzymes do not appear to have such a nucleophile in their active sites. The simplest alternative is an S1-like mechanism, requiring an oxocarbenium cation intermediate (13). However, a series of experiments have indicated such a cation to be quite unstable in solution (14–16), which has led to the proposal of an S1,datum-like (substitution nucleophilic internal return) mechanism (10, 13). In this mechanism attack of the nucleophile occurs on the same face as the departure of the leaving group, and at nearly the same time, thus avoiding a discrete cation intermediate. Interaction between nucleophile and leaving group is a hallmark of this mechanism. It is postulated that attack and departure occur “asynchronously” (10). This mechanism has been suggested for several enzymes, including glycogen phosphorylase (GT-B, retaining) (17, 18), the first structure determined with the double Rossmann-fold domain architecture. Several structures of enzymes from the GT-B family have been determined where their C-domains contain the primary donor-binding site and their N-domains contain the primary acceptor binding site (19–25). The active site of all of these enzymes lies in the cleft between the two domains and therefore the correct relative orientation of the two domains is required for correct active site geometry and catalytic competence (20, 26, 27). Nowhere is this more critical than for GS enzymes, because they are characterized by a very large domain motion, much larger than that seen in any of the other GT-B retaining enzymes of known structure. The active site is therefore formed only upon closure of the two domains around the substrates. To date, there are two structures of GS, one from Agrobacterium tumefaciens (28) (AtGS) and one from Pyrococcus abyssi (29) (PaGS). However, both of these enzymes were crystallized in their catalytically inactive “open” conformations. To better understand both the structure and mechanism of GS, the GT-B enzyme family and retaining glycosyltransferases in general, we have determined a series of structures of GS from Escherichia coli. These structures include...
both the open, and for the first time the ADP/glucose-bound, closed form of the enzyme. Together, these structures give the clearest picture to date of the structure and function of GS and starch synthase enzymes.

**EXPERIMENTAL PROCEDURES**

**Protein Expression and Crystallization**—The wild-type EcGS (wtGS), mutant *E. coli* GS E377A, and double mutant C7S,C409S (dmGS) were overexpressed in *E. coli* and purified as described elsewhere (30). The purified proteins were then buffer-exchanged to 20 mM triethanolamine-HCl (pH 7.5), 5 mM dithiothreitol and concentrated to 5–8 mg/ml. All crystallizations were carried out at 4 °C by the hanging drop vapor-diffusion method. All wtGS crystals were obtained by cocrystallization of 7.8 mg/ml protein with 3 mM ADP-Glc. Four independent data sets were collected from four different crystals of wtGS with ADP-Glc (wtGSa, wtGSb, wtGSc, and wtGSD). The wtGsa crystal (0.01 × 0.2 × 0.2 millimeter in size) was obtained after a 4-week incubation in a solution of 40% (w/v) PEG 4000, 0.2 mM sodium acetate, and 0.1 mM HEPPSO (pH 8.1) (Fig. 1A). The crystals of wtGSb and wtGSc grew in 40% (w/v) PEG 4000, 0.2 mM sodium tartrate, and 0.1 mM HEPPSO (pH 7.7), and reached a maximum size 0.25 × 0.25 × 0.25 and 0.2 × 0.2 × 0.2 millimeters, respectively, over 12 weeks. The 0.2 × 0.2 × 0.2-millimeter wtGSc crystal was also obtained after a 12-week incubation under similar conditions, but at pH 7.6. *E. coli* GS E377A + ADP-Glc complex crystals were obtained from a buffer containing 7.8 mg/ml protein, 3 mM ADP-Glc, 3 mM maltotriose, 0.2 mM dithiothreitol and concentrated to 7.8 mg/ml protein, 3 mM ADP-Glc, and reached a maximum size 0.25 × 0.25 × 0.25 millimeter in size) was obtained after a 4-week incubation in a solution of 40% (w/v) PEG 4000, 0.2 mM sodium acetate, and 0.1 mM HEPPSO (pH 7.5). After 16 weeks incubation, the crystals grew to 0.2 × 0.2 × 0.2 millimeter and were used for data collection.

Crystals of dmGS were obtained in the presence of 0.54 mM ADP-Glc in 40% (w/v) PEG 4000, 0.1 mM Tris (pH 7.5), and 0.2 mM sodium tartrate buffer. The crystal used for data collection had been incubated for 16 weeks before reaching its maximum size at 0.2 × 0.2 × 0.3 millimeter.

All crystals were harvested into the mother liquid with 15% glycerol as cryoprotectant and flash-frozen for 100 K data collection at the Advance Photon Source at Argonne National Laboratories (Argonne, IL). Diffraction data were reduced and scaled with DENZO and SCALEPACK (31, 32). The dmGS crystal belongs to the H32 space group with unit cell parameters *a* = *b* = 232.96 Å, *c* = 85.54 Å, *α* = *β* = 90°, *γ* = 120°. All other GS crystals belong to space group I41 with unit cell parameters *a* = *b* = 126 ± 1.0 Å, *c* = 152 ± 1.0 Å, *α* = *β* = *γ* = 90°. In all cases, only one molecule was found in each asymmetric unit. Other crystal parameters and detailed data collection statistics are listed in Table 1.

**Structure Determination and Refinement**—The structures of dmGS and wtGsb were solved by the Molecular Replacement method with MOLREP (Collaborative Computational Project, CCP4) using the AtGS structure as a search model, with non-conserved amino acid residues mutated to alanine, except for glycine. Initially, plausible rotation and translation functions for the N terminus (1–240) were identified, and the C terminus (271–456) was subsequently incorporated to search for its counterpart. All other EcGS complex structures were solved by the Molecular Replacement method with MOLREP using the wtGsb structure as model. Placement of the model was optimized by rigid-body refinement. Refinement and map calculations were carried out with the program REFMAC5 (CCP4). All model building was performed using TURBO-FRODO. After convergence, the ligands HEPPSO, ADP, and glucose were incorporated using refinement dictionaries calculated from the ideal structure provided by the CCP4 dictionary. Water molecules were added using the ARP/wARP option within REFMAC5 and inspected visually prior to deposition. In all data sets, 5% of the observations were flagged as “free” (33) and used to monitor Rfree. All final models display good stereochemistry as evaluated with the program PROCHECK (CCP4 with less than 0.3% of the residues in disallowed regions of the Ramachandran plots). All figures were generated with PyMOL (34).

**Multiple Sequence Alignment**—Multiple sequence alignment was conducted with DNASTAR. GS sequences used were: EcGS (PA06U8), AtGS (AAD03474), and PaGS (NP_125769). Sequences of granule-bound starch synthases from various organisms were those of barley (AA717109), maize (P04713), potato (CAA41359), and rice (P19395). Soluble starch synthases used were maize SSI (AAB99957), potato SSI (P39568), wheat SSI (Q43654), wheat SSIa (BAE48798), maize SSIa (AA577569), potato SSI (CAA61241), potato SSIII (Q43486), and wheat SSIIII (AAH87999), and *Chlamydomonas reinhardtii* SS.

**RESULTS**

**Overall Molecular Organization**—A total of four distinct structures are reported here: an apo, open form; a closed form, wtGS bound to ADP, glucose, and HEPPSO; wtGS bound to ADP, an unidentified glucose derivative; and HEPPSO and *E. coli* GS E377A bound to ADP and HEPPSO. The double mutant dmGS (C7S,C408S) was used as a wild-type EcGS alternative in crystallization trials because EcGS had shown a tendency to aggregate, and it was thought that a cysteine-less form might be more suitable for crystallization. Of the three Cys residues present in EcGS, Cys-379 was maintained because its replacement by Ser affects the apparent affinity for ADP-Glc. The variant dmGS exhibits comparable specific activity and apparent affinity for ADP-Glc to wtGS (30). C7S and C408S are not close to the active site and are separately buried in the protein structure (Fig. 1A), suggesting that this double mutation had little impact on the enzyme structure.

EcGS is a 52.8-kDa protein with a total of 477 residues and is monomeric in all of our crystals. In contrast, the oligomerization state for the archael GS and AtGS was dimer and trimer, respectively. An early report indicated that active EcGS could be isolated in a variety of oligomerizations states (35). The density map of apo-dmGS shows the first 475 residues. The His<sub>8</sub> tag attached to the C terminus of wtGS was not traceable. A typical twin-Rossmann GT-B fold is exhibited by all the EcGS structures; the N- and C-terminal domains are similar in size (1–241 and 250–457) and are composed of a “sandwich” of parallel β-sheets between α-helices. The two domains are structurally homologous (r.m.s. deviation 1.9 Å), although there is no sequence homology between them. The extended interdomain linker peptidase-(242–254) connects the N- and C-terminal halves. The helical tail α18-

**Crystal Structure of Closed Conformation of E. coli GS**
Crystall Structure of Closed Conformation of E. coli GS

TABLE 1
Data collection and refinement statistics

|                      | dmGS(C7S;C408S) | wtGSa | wtGSb | wtGSc | wtGSD | E377A |
|----------------------|------------------|-------|-------|-------|-------|-------|
| Space group          | H 32             | 141   | 141   | 141   | 141   | 141   |
| Unit cell            |                  |       |       |       |       |       |
| a, b (Å)             | 233.0            | 125.8 | 126.5 | 127.1 | 126.2 | 126.1 |
| c (Å)                | 85.5             | 153.3 | 151.9 | 152.0 | 151.6 | 152.1 |
| α,β(°)               | 90.0             | 90.0  | 90.0  | 90.0  | 90.0  | 90.0  |
| γ (°)                | 120.0            | 90.0  | 90.0  | 90.0  | 90.0  | 90.0  |
| Data processing statistics |                  |       |       |       |       |       |
| Beam line            | COMCATE-ID32     | BIOCARS-ID14 | DNDCAE-BM5 | DNDCAE-BM5 | DNDCAE-BM5 | SBCCAE-BM19 |
| Wavelength (Å)       | 1.01             | 1.01  | 0.99  | 0.99  | 0.99  | 1.03  |
| Resolution (Å)       | 50.0-3.0(3.2-3.0) | 50-2.8 (2.9-2.8) | 97.1-2.2(2.4-2.2) | 97.1-2.3(2.4-2.3) | 97.1-2.4(2.5-2.4) | 50.0-2.3(2.4-2.3) |
| Total reflections     | 21871 (2146)     | 27528 (1765) | 57263 (4243) | 53217 (3870) | 44756 (3333) | 30260 (2994) |
| Multiplicity         | 9.4 (9.5)        | 7.5 (3.5) | 3.15 (3.1) | 2.5 (2.6) | 2.7 (2.8) | 3.5 (3.6) |
| Completeness (%)     | 99.9 (100)       | 98.5 (87.6) | 99.8 (99.3) | 99.2 (99.1) | 97.7 (90.8) | 92.5 (88.7) |
| Rmerge(%)            | 7.6 (34.6)       | 10.9 (36.0) | 5.7 (28.9) | 4.21 (20.6) | 6.12 (32.7) | 7.8 (42.0) |
| I/σ                  | 28.3 (1.9)       | 18.07 (2.0) | 11.29 (2.0) | 10.61 (2.0) | 10.34 (1.8) | 13.7 (2.0) |
| Refinement statistics |                  |       |       |       |       |       |
| R factor(%)          | 20.9             | 17.5  | 20.3  | 16.3  | 18.2  | 17.4  |
| Rfree(%)             | 26.3             | 19.9  | 22.8  | 17.9  | 20.2  | 19.7  |
| r.m.s. deviation from ideal values |      |       |       |       |       |       |
| Bond length (Å)      | 0.012            | 0.009 | 0.014 | 0.008 | 0.011 | 0.009 |
| Bond angle (°)       | 1.47             | 1.32  | 1.59  | 1.25  | 1.45  | 1.25  |
| Number of atoms      |                 |       |       |       |       |       |
| Protein atoms        | 3721             | 3735  | 3735  | 3735  | 3735  | 3749  |
| Water molecules      | 20               | 239   | 214   | 334   | 200   | 298   |
| Other components      | ADP, DGM(0.5)    | ADPGl (0.6) | ADPGl (0.8) | ADPGl (0.8) | ADPGl (0.8) | ADP |
|                      | HEPPSO, PEG      | HEPPSO (0.8) | HEPPSO, PEG | HEPPSO, PEG | HEPPSO, PEG | HEPPSO |
| Ramachandran statistics (%) |    |       |       |       |       |       |
| Most favoured regions| 84.7             | 90.0  | 90.0  | 90.0  | 90.7  | 88.7  |
| Additional allowed regions | 14.8         | 9.8   | 9.8   | 9.8   | 9.1   | 11.1  |
| Generally allowed regions | 0.2          | 0.2   | 0.2   | 0.2   | 0.2   | 0.2   |
| Disallowed regions   | 0.2              | 0.0   | 0.0   | 0.0   | 0.0   | 0.0   |
| Average B (Å²)       | 50.65            | 49.02 | 37.97 | 38.00 | 41.93 | 44.49 |
| PDB code             | 3DJJ             | 2QYY  | 2QZS  | 2R4T  | 2R4U  | 3COP  |

Values in parentheses are for the highest resolution shell in data section and for occupancy in refinement section.

aACT, acetate ion; PEG, polyethylene glycol chain.

(458–476) crosses over from the C-terminal domain to pack against the N-terminal domain (Fig. 1B). These results are consistent with previously reported AtGS (28) and PaGS (29) structures.

There are two major forms of EcGS: an open form with the two domains spread relatively far apart, and a closed form in which the two domains are very close together. Although the apo-dmGS structure adopts the open form, the rest of our structures represent the catalytically active, closed form of the enzyme (Fig. 1C). The N- and C-terminal halves of apo-dmGS are separated by a deep cleft with most interdomain interactions occurring between the N-terminal helix α7-(215–220), the preceding loop (212–215), and the C-terminal helix α14-(398–403). Val-248 in the linker peptide also makes one interdomain hydrogen bond with Phe-460 in the domain-spanning helix α18 (Fig. 1B).

The Closed Form of EcGS Is Produced by Co-crystallization with ADP-Glc and HEPPSO—wtGS was co-crystallized with the donor substrate ADP-Glc. Crystals grew and were harvested and frozen at two time points, 4 and 12 weeks. Even at the earliest time point (4 weeks, wtGSa) ADP-Glc was not observed in the active site. Instead, ADP and a not precisely
identified glucose derivative were found. ADP-Glc was clearly not present because there was no electron density for the bond between the sugar ring and the terminal phosphate oxygen of ADP, and the distance between the glucose moiety and the phosphate oxygen was 3.5–3.6 Å, much too long to accommodate a covalent bond between the two atoms. Then again, no electron density was evident for an additional substituent on the C1 carbon, indicating that glucose was not present (Fig. 1D). Data from crystals that were harvested after 12 weeks also have electron density maps consistent with the presence of ADP and a glucose derivative with no covalent bond between them, but these crystals clearly show density for an additional oxygen atom attached to C1 of the glucose moiety, consistent with the presence of ADP and the hydrolysis product glucose in the active site (Fig. 1D). Three separate data sets were collected from three crystals after 12 weeks (wtGSb, wtGSc, and wtGSd) all showing the same glucose moiety in the active site. In all four of these structures, the two domains of GS have closed around the molecules in the active site, resulting in a 15.2° motion of the C-terminal domain relative to the N-terminal domain (Fig. 1C, all four structures are superimposable and only wtGSb is shown as representative). Most of the interdomain interactions seen in the open form are preserved in the closed form. Additionally, the closed form is stabilized by new interdomain interactions, including those between Asn-162 and Gln-304, and between Lys-15 and Glu-357.

Superposition of the C-terminal domain of the apo-dmGS and the ADP, glucose-bound wtGS structures revealed that displacement of the N-terminal domain is as much as 9.3 Å (Fig. 1C). However, the r.m.s. deviation between the N-terminal domains (the α-carbon atoms of residues 1–241) and the C-terminal domains (α-carbon atoms of residues 251–375 and 382–477) are only 0.68 and 0.61 Å, respectively, indicating a rigid global domain-domain movement. The program DynDom (36) revealed that a 15.2° domain-domain closure switches the open apo-GS structure to the closed form. The open form of GS has been reported for AtGS (28) and PaGS (29). However, the open form of apo-AtGS is 10.2° more open than the E. coli apo-dmGS structure, perhaps due to crystal packing or to species specificity or both. The structure of wtGS in complex with ADP, glucose, and HEPPSO shows the compact, closed form

- **FIGURE 1.** A, diagram of HEPPSO and atom numbering used in this paper. B, overall structure of E. coli dmGS. The N terminus (1–241) and C terminus (250–457) are colored yellow and blue, respectively. The interdomain peptide (242–254) and domain-spanning helix α18 (458–476) are colored magenta and orange, respectively. The residues involved in the interdomain interaction in the open form EcGS are shown in stick and are colored (N-terminal loop 212–215 and helix α7, lime; C-terminal α14 helix, light blue). ADP-base binding loop 354–356 is colored cyan. The mutated residues of dmGS (C7S, C408S) are shown in red as stick representation. C, superposition of the C-terminal domains from dmGS (cyan) and wtGSb (yellow). Ligands bound in wtGSb (HEPPSO, ADP, and glucose) are shown in black as stick representations. Structural elements essential for catalysis are colored red and their counterparts in the open form dmGS are colored blue. Some active site residues (His-161, Arg-300, and Lys-305) are shown in red as stick representations. D, 4-week incubated wtGSa (2.8-Å resolution) Fo – Fc electron density map contoured at 2.5 σ with DGM (left) and d-arabino-hex-1-enitol (middle). Right, 12-week incubated wtGSb (2.2-Å resolution) 2.5-σ contoured Fσ – Fc electron density map of glucose.
TABLE 2
Partial alignment of conserved residues in bacterial GS and plant SS
Residues mentioned in this paper and their conserved equivalents in SSs are in bold.

|    | E. coli GS | AtGS | barley GBSS | maize GBSS | rice GBSS | potato SS | potato SSII | potato SSIIa | wheat SS | wheat SSII | Chlamydomonas reinhardtii SS |
|----|------------|------|-------------|------------|-----------|------------|-------------|--------------|----------|------------|-----------------------------|
| 9 E | 15 KT  GGL ADV | 95 Y | 137 DW H | 161 H N | 246 N G | 300 R L | 304Q | K G | 355 Y | 373 PSRF | E PCGLTQL | E. coli GS |
| 9 E | 15 KT  GGL ADV | 96 Y | 138 DW Q | 162 H N | 246 N G | 300 R L | 304Q | K G | 354 Y | 372 PSRF | E PCGLTQL | AtGS |
| 9 E | 14 K V GGL ADV | 89 Y | 128 DW H | 151 H R | 217 N G | 257 R F | 262Q | K G | 317 L | 335 PSYF | E PFGLTQL | PsA |
| 84 E | 90 KT  GGL GDV | 176 Y | 227 DW H | 257 H N | 346 N G | 401 R L | 405Q | K G | 457 F | 475 TSRF | E PCGLTQL | barley GBSS |
| 85 E | 91 KT  GGL GDV | 177 Y | 228 DW H | 258 H N | 347 N G | 402 R L | 406Q | K G | 459 F | 477 TSRF | E PCGLTQL | maize GBSS |
| 89 E | 95 KT  GGL GDV | 181 Y | 232 DW H | 262 H N | 351 N G | 406 R L | 410Q | K G | 461 F | 480 PSRF | E PCGLTQL | potato GBSSI |
| 91 E | 97 KT  GGL GDV | 183 Y | 234 DW H | 264 H N | 353 N G | 408 R L | 412Q | K G | 463 F | 481 PSRF | E PCGLTQL | rice GBSS |
| 140 E | 146 KT  GGL GDV | 232 Y | 277 DW H | 307 H N | 401 N G | 455 R L | 459Q | K G | 510 F | 528 PSRF | E PCGLTQL | potato SS |
| 139 E | 145 KT  GGL GDV | 232 Y | 277 DW H | 307 H N | 401 N G | 455 R L | 459Q | K G | 510 F | 528 PSRF | E PCGLNQL | wheat SS |
| 147 E | 153 KT  GGL GDV | 239 Y | 284 DW H | 314 H N | 408 N G | 462 R L | 466Q | K G | 517 F | 535 PSRF | E PCGLNQL | potato SSII |
| 316 E | 322 KT  GGL GDV | 403 Y | 446 DW H | 476 H N | 562 N G | 621 R L | 625Q | K G | 676 F | 694 PSRF | E PCGLNQL | wheat SSIIa |
| 249 E | 255 KT  GGL GDV | 333 Y | 376 DW H | 406 H N | 492 N G | 551 R L | 555Q | K G | 605 F | 624 PSRF | E PCGLNQL | maize SSIIa |
| 284 E | 290 KT  GGL GDV | 371 Y | 414 DW H | 444 H N | 530 N G | 589 R L | 593Q | K G | 644 F | 662 PSRF | E PCGLNQL | potato SSII |
| 788 E | 794 KT  GGL GDV | 875 Y | 909 DW S | 938 H N | 986 N G | 1040 R L | 1044Q | K G | 1100 Y | 1118 PS F | E PCGLTQL | potato SSIII |
| 1188 E | 1194 KT  GGL GDV | 1275 Y | 1308 DW S | 1337 H N | 1385 N G | 1439 R L | 1443Q | K G | 1499 Y | 1517 PS F | E PCGLTQL | wheat SSIII |
| 156 E | 162 KT  GGL GDV | 259 Y | 294 DW H | 323 H N | 416 N G | 470 R L | 474Q | K G | 528 Y | 544 PSMF | E PCGLTQL | Chlamydomonas reinhardtii SS |

with a narrow cleft at the N- and C-domain interface, which serves as the binding pocket of the substrate. In addition to the global domain-domain rotation, superposition of the C-terminal domain of the apo-dmGS on to the ADP/glucose-bound wtGS also revealed local conformational rearrangements, most of which are small side chain movements on the N- and C-terminal domain interface. The only conspicuous local conformational change was found in the loop 376–381 (Fig. 1C). The entire loop flips down to avoid collision with the glucose moiety. Still in proximity, this loop makes extensive interactions with the glucose to be transferred. The conformational change of loop 376–381 also establishes the interdomain hydrogen bond between the conserved Ser-212 side chain and the Pro-378 backbone carbonyl group as well as packing between Tyr-215 and Pro-378, making it likely that domain closure and loop motion are correlated.

Substrate Binding in wtGS—ADP is similarly located in all four wtGS co-complexes with virtually identical interactions observed between the protein and the ADP molecule. ADP is bound in the wtGS interdomain cleft, mostly along the C-terminal domain wall and makes interactions with several C-terminal domain residues. The adenine moiety extensively interacts with the C-terminal 354–356 loop; the adenine ring stacks against the conserved Tyr-355 (Phe in animal GS and plant SS) (Table 2; Fig. 2A). The adenine N1 atom accepts a hydrogen bond from the backbone amide of His-356 (3.1 Å), and the adenine amide N6 contacts the Gly-354 backbone carboxyl (2.9 Å). The interaction of ADP with the N-terminal domain is restricted to Asp-21, Lys-15, and Gly-18, all of which are absolutely conserved in both glycogen and starch synthases. The fact that ADP makes interactions with both the C- and N-terminal domains of GS could lead to the conclusion that ADP binding may be enough to produce the closed conformation of GS. However, this is not the case because the AtGS-ADP complex is still in the open form (28). As will be explained below, the binding of HEPPSO in the glycan acceptor binding site appears also to be required for maintenance of the closed conformation.

The ribose of ADP is located at the domain interface and adopts a C2’-endo conformation relative to the adenine base (Fig. 2A). A water-mediated hydrogen bond between the ribose 2’-hydroxyl and the adenine N3 atom was found in all wtGS structures, presumably reinforcing the observed ribose orientation. The ribose O3’ hydroxyl group makes hydrogen bonds with both OD2 of Asp-21 (2.8 Å) and N-ε of Lys-15 (2.7 Å).

The ADP phosphate groups are tucked back with respect to the adenine base, allowing Gly-18 to make a hydrogen bond with O5’ of ADP (3.2 Å) and O3B of the distal phosphate group (2.7 Å) (Fig. 2A). The proximal phosphate is near the conserved loop 374–381 and both phosphate O1A and O2A make hydrogen bonds with the main chain amide of Leu-381 (3.3 Å and 3.0 Å, respectively). The O2A atom of the proximal phosphate is also hydrogen-bonded to the Thr-382 backbone amide (3.2 Å). The distal phosphate group is close to loop 299–306 and interacts extensively with Arg-300 and Lys-305 through ionic interaction, direct hydrogen bonds, and water-mediated hydrogen bonds (Fig. 2B and C). Arg-300 makes a hydrogen bond with O1B and O2B of the distal phosphate with its N-ε (2.8 Å) and N-ε (2.8 Å) atoms, respectively. A water molecule mediates an additional interaction between the Arg-300 backbone amide and the distal phosphate O2B atom. Lys-305 makes a 2.8-Å hydrogen bond to the distal phosphate O1B through its N-ε group, which is also in contact with the proximal phosphate O3A atom (3.0 Å). The interaction between Lys-305 and the proximal phosphate also includes two water-mediated hydrogen bonds between its N-ε atom and the proximal phosphate O3A and O2A (Fig. 2B).
The different density maps of wtGS structures all showed considerable gaps between the distal phosphate oxygen and the anomeric carbon of the glucose moiety of ADP, indicating that the phosphate-glucosyl bond had been broken. A glucose molecule is found in the active sites of wtGSb (refined to 2.2-Å resolution), wtGSc (refined to 2.3-Å resolution), and wtGSd (refined to 2.4-Å resolution). Because the glucose positions in wtGSb, wtGSc, and wtGSd are almost identical, the highest resolution structure, wtGSb (2.2 Å), will be used to describe the interaction between glucose and protein. The Glc is at the domain interface and extensively interacts with the C-terminal 373–380 loop and N-terminal residues His-161 and Asn-162 (Fig. 2A). The 2-hydroxyl group forms hydrogen bonds with the side chain of Asn-162 (3.2 Å) and Gln-304 (3.2 Å). The 3-hydroxyl group forms a 2.5-Å hydrogen bond with the carboxyl oxygen of Glu-377 and the backbone amide of Cys-379 (3.0 Å) and Gly-380 (2.9 Å). The 4-hydroxyl group makes a hydrogen bond with O1A of the proximal phosphate (2.4 Å), and with the Gly-380 backbone amide (2.9 Å) (Fig. 2A). The 6-hydroxyl group hydrogen bonds to the interdomain peptide residue Asn-246 side chain carbonyl group (2.9 Å) and His-161 side chain (2.8 Å). The backbone carbonyl group of His-161 is only 2.9 Å from the anomeric carbon C1, directly below the 1-phosphate oxygen of ADP, and thereby adopting the position a catalytic nucleophile would be expected to occupy (Fig. 2B). In all the wtGS structures the anomeric C1 is 3.5–3.6 Å from oxygen O3B of the distal phosphate, which is almost twice a regular carbon-oxygen bond length, clearly showing that the glucose phosphate bond is broken. On the other hand, the 3.0-Å distance between the glucose 1-hydroxyl and O3B of the distal phosphate means that there is a hydrogen bond between glucose O1 and the leaving group phosphate.

All GS complex crystals grew in the presence of the buffer HEPPSO. Clear electron density was consistently observed for HEPPSO in the interdomain cleft in all cases. HEPPSO lies along the N-terminal side of the crevice wall with its hydroxy-ethyl end near the glucose. The HEPPSO O5 was also close to the KTGGL motif loop and makes a 2.9-Å hydrogen bond with the Leu-19 backbone amide. The sulfonic acid end of

FIGURE 2. A, stereo views of ADP and glucose (shown in yellow and in atom colors) bound in the active site of EcGS. Hydrogen bonds between ADP, glucose, and protein are shown as broken lines. Carbon atoms of N-terminal and C-terminal residues are colored pink and blue, respectively. B, stereo view of the active site rotated about the vertical axis by about 90° relative to A. Critical interactions between substrate and GS are shown as broken lines. C, structural comparison (aligned by TURBO-FRODO (50)) using the Ca carbons of the C-terminal domain) of bound ADP in the open form of AtGS (yellow, PDB 1rzu) and closed form of EcGS (blue). The residues from AtGS and EcGS are colored yellow and blue, respectively. The interactions between ADP and the protein are shown as broken lines. D, structural comparison of loop 13–20 in dmGS (yellow) and wtGSb (red) and the equivalent loop 14–24 in the UDP/imidazole/Glu-6-P-bound OtsA (green, PDB 1gz5) and UDP-Glc-bound OtsA (blue, PDB 1uqu). The coil preceding and the helix subsequent to the superimposable loop 14–24 are also shown. When OtsA is in complex with UDP-Glc, the region between Asp-14 and Ser-19 is disordered and loop 14–24 is only partially structured. Gly-22 in OtsA and Gly-18 in EcGS interact with the phosphate oxygen and are shown as sticks. Ligands other than UDP (OtsA) and ADP (GS) are not shown for clarity. Structural alignment by TURBO-FRODO using the C-terminal domains Ca carbons of the two EcGS proteins and the active site residues for OtsA).
HEPPSO points to the protein surface and interacts with N-terminal residues along the cleft wall. Conserved Trp-138 forms hydrogen bonds with both the branched hydroxyl O4 and sulfonic acid oxygen O1. The conserved residue Asp-137 also interacts with the branching hydroxyl O4 through a water-mediated hydrogen bond (Fig. 2B).

DISCUSSION

Comparison of the Open ADP-bound AtGS and the Closed ADP/Glucose-bound EcGS—GS differs from all other GT-B retaining enzymes whose structures are known by the very large interdomain motion accompanied by substrate binding. The EcGS active site is composed of residues on flexible loops between the β strand and α helix emanating from the domain boundaries. Comparison of the open and closed forms of EcGS suggests that only after the domain-domain closure are these substrate-binding residues brought together to construct a competent active site. In an effort to understand how protein conformation would affect the ligand binding, we compared the binding mode of ADP in the open form of the ADP-bound AtGS structure (28) with the closed form ADP-containing wtGS structure (Fig. 2C). Compared with ADP in the AtGS open form, the adenosine diphosphate group in the closed form of wtGS exhibits a twisted conformation at the ribose ring and the proximal phosphate. The ribose ring in the closed form of wtGS rotates ~52° and the ribose 3-hydroxyl group then falls within hydrogen bonding range of the side chain of Lys-15 and Asp-21, which were ~8 Å away in the AtGS open form. The proximal phosphate forms two hydrogen bonds to the backbone amid of Leu-381. The distal phosphate of ADP, which extensively interacts with C-terminal residues Arg-300 and Lys-305, did not show the same extent of twist as the ribose and the proximal phosphate. It seems that the difference in conformation of the ADP molecule between open and closed forms of GS is initiated at the ribose where the N-terminal residues approach and interact, suggesting that closure of the enzyme is necessary for the ADP-Glc substrate to adopt a catalytically competent conformation in the protein. Asp-21 replacement for Ala generates both a 12-fold decrease in $V_{max}$ and an 18-fold increase in $K_m$ for ADP-Glc (37), consistent with its importance for closure, correct positioning, and subsequent reaction of ADP-Glc.

The KTGGGL Motif—The KTGGGL motif is widely conserved among glycosyn synthases and starch synthases and was suggested to be involved in substrate binding (38, 39). This motif, in the loop 13–20 at the N- and C-terminal domain interface, is 5–7 Å away from the active site in the apo-dmGS open form (Fig. 2D). The domain-domain closure brings this N-terminal loop into the vicinity of the C-terminal domain to interact with both the substrate ADP-Glc and the acceptor molecule. Lys-15 and Gly-18 make hydrogen bonds with the ribose hydroxyl and β-phosphate oxygen, respectively. Lys-15 also participates in a domain-spanning hydrogen bonding network together with Glu-357 and Tyr-355, presumably further facilitating the cross-talk between N- and C-domains. The last three residues of the KTGGGL motif are highly conserved in many other GT-B retaining enzymes, including GP, MalP, and trehalose-6-phosphate synthase (OtsA). OtsA transfers glucose from UDP-Glc to glucose 6-phosphate (Glc-6-P) to form α,α1,1-trehalose-6-phosphate. The OtsA GGL motif is on loop 13–20, structurally equivalent to EcGS loop 13–20 (Fig. 2D). In the UDP-Glc-bound OtsA structure, loop 14–24 is partially disordered (14–19), whereas the rest of the loop (20–24) is 7–9 Å away from the substrate UDP-Glc (20), reminiscent of the comparable distance between GS loop 13–20 in the open apo-dmGS structure and the ADP-Glc (modeled by alignment of the EcGSb structure). In the presence of the substrate analogue UDP/imidazole and the acceptor Glc-6-P, the entire loop 14–24 of OtsA becomes structured and moves to a position adjacent to the UDP molecule. In the OtsA complex, the GGL motif interacts with the substrate donor and locks the substrate in position for the glucosyl transfer, with its Gly-22 (Gly-18 equivalent) hydrogen-bonding O5* and the β-phosphate oxygen of UDP (21). OtsA manages to encapsulate the active site with the GGL motif via local conformational changes as evidenced by the small N-terminal domain displacement (~1–1.5 Å). In contrast, EcGS undergoes a large global domain-domain rotation (9.5-Å N-terminal displacement) to move the GGL motif into the interdomain catalytic center.

The very large interdomain motion that appears to occur upon binding both ligands to GS is different from that seen in other GT-B fold retaining enzymes whose structures are known, as none of these enzymes exhibit such dramatic interdomain rearrangement (19, 20, 22, 25–27). Although the active sites become comparatively more open or closed upon binding ligands in these structures, most of the motion comes from the loops surrounding the active site. The explanation for this large interdomain motion is unclear, but it does allow independent binding of both donor ADP-Glc and acceptor oligosaccharide without interference as the structure indicates that the active site would be completely inaccessible to ADP-Glc binding if bound to a glycan in the closed conformation. This is consistent with kinetic studies of human polymorphonuclear leukocyte (1), rat liver (2), and Neurospora crassa (3) glycogen synthases that show that these enzymes have a rapid equilibrium random bi-bi mechanism (4, 5, 40, 41). Although the allosteric nature of these enzymes may alter their kinetics relative to the E. coli enzyme, the sequence homology between them and the E. coli enzyme, especially in the active site, indicate that a similar catalytic mechanism may be employed. The kinetic mechanism for a bacterial glycogen synthase has not been determined.

The Role of Glu-377—Glu-377 is absolutely conserved in all GS and SS (Table 2) enzymes and is also conserved in MalP, GP, trehalose phosphorylase, and OtsA (Asp). It is the first residue in the E-X7-E motif of some retaining glycosyltransferases, such as eukaryotic glycogen synthases (GT3), α-glucosyltransferase (GT4), starch synthases, and bacterial glycogen synthases (GT5). Substitution of EcGS Glu-377 by Gln or Ala (30) results in a drop in activity of more than 4 orders of magnitude, with little change in apparent affinity for the substrates. Replacement of the equivalent MalP residue Glu-637 by Gln (42), or the equivalent GTB Glu-303 by Ala (19) all result in a drastic reduction in enzyme activity. Mutating Glu-510 (the Glu-377 equivalent) completely inactivated the human muscle glycogen synthase (43). Glu-377 and its equivalents were therefore originally thought to be either the catalytic nucleophile or the general
acid/base catalyst for these glycosyltransferases (18, 19, 30, 43, 44). However, virtually all of the GT-B retaining enzyme crystal structures that have so far been determined (MalP, OtsA, GS, and WaaG, AGT is discussed below) have shown that Glu-377 is located on the H9251 -face of the glucose to be transferred and is relatively far from the C1 glucose position (5.4 Å in EcGS), which seems inconsistent with Glu-377 playing a direct role in nucleophilic addition (Fig. 3A). Although the Glu-377 side chain is not far from the H9252 -phosphate (3.9 Å in EcGS), positively charged Arg-300 and Lys-305 directly interact with the substrate phosphate group (Fig. 2, B and C), and are therefore more likely to play the role of proton transfer agent. However, Glu-377 does directly interact with Lys-305 (Fig. 3A) stabilizing the positive charge on this residue. This is likely to be important for catalysis as the equivalent interaction is preserved in almost all retaining GT-B enzymes whose structures are known. In the one exception to this, AGT, the short 377 loop is replaced by a much longer loop and the EcGS Gly-380 residue is replaced by a Glu (Glu-306 in AGT), wherein it occupies a location identical to EcGS Glu-377 and makes an equivalent salt bridge with its Lys-305 equivalent (Lys-209 in AGT). Therefore it appears to be critical for catalysis that a negative charge be located proximal to Lys-305 or its equivalent in retaining GT-B enzymes. However, this cannot be its only role in GS, because the effect on catalysis of the mutant E377A is larger than that of K305A by an order of magnitude (45). Our structures suggest that its other role is in the proper positioning of the glucose moiety in the active site. A similar situation can be observed for the equivalent acidic residue in MalP, OtsA, WaaG, and GP where the side chain carboxyl group also makes a direct interaction with the glucose moiety (21–23, 46, 47). Again only AGT is different, as Glu-306 does not make a hydrogen bond to the sugar donor in this structure.

To shed further light on this issue we have determined the structure of the E377A mutant ADP-Glc-HEP-PSO complex. The structure is very similar to the other closed conformation GS complexes, but there is no density for the glucose moiety of ADP-Glc in its binding site. Instead there is diffuse density around Arg-300, which has in fact flipped out of the active site (Fig. 3A). However, this diffuse density is not interpretable. The glucose moiety appears to be either disordered or absent in this structure consistent with the idea that loss of Glu-377 results in the inability of the enzyme to properly position the donor glucose moiety, consistent with its role as glucose locator. On the other hand, it is not surprising that loss of this interaction does not have such a detrimental effect on the Km for ADP-Glc as GS makes a large number of interactions with this large ligand and the loss of one interaction is not expected to result in significant loss of binding.
energy (30, 37). This is also consistent with the E377A EcGS-ADP-Glc-HEPPSO complex structure as the ADP portion of ADP-Glc is identically positioned, well ordered, and fully occupied in the active site of the mutant. We therefore conclude that Glu-377 likely plays a major role in the proper placement of only the glucose moiety of the substrate, whereas the rest of the substrate binds tightly even in the absence of this interaction. It is also interesting to note that despite the loss of its interaction with the side chain of Glu-377, Lys-305 is positioned identically to all of our other EcGS structures, indicating that Glu-377 is not required for positioning this residue, although it likely has a significant impact on its electrostatics as described above. The fact that the E377A EcGS-ADP-Glc-HEPPSO complex structure adopts the closed conformation despite the apparent disorder or loss of the glucose moiety indicates that the glucose moiety may not be critical for maintenance of the closed conformation.

Accepter Analogue HEPPSO—HEPPSO was added as a buffer in GS crystallization trials, but turned out to bind in the active sites of all four closed GS structures, suggesting that HEPPSO is critical for capturing the closed form of EcGS. Although HEPPSO itself directly interacts only with the N-terminal domain of GS, its indirect interactions to the C terminus via the KTGGGL motif and ADP-Glc are presumably critical in bringing the two domains together. Superposition of wtGS-HEPPSO and MalP-maltopentaose complex structures (23) reveals that HEPPSO superimposes on the maltopentaose position in the MalP structure and probably acts as an oligosaccharide acceptor analogue for EcGS (Fig. 3B). The HEPPSO branching hydroxyl group O4 makes a hydrogen bond to the conserved Asp-137 and Trp-138. Asp-137 was suggested by mutagenesis and modeling studies to be involved in recruiting and locking the glucan chain acceptor, making it a good enough leaving group to allow the hydrolysis negative charge on the phosphate, or is even the proton donor, making it a good enough leaving group to allow the hydrolysis reaction we see in the active site to occur. If true, this would be consistent with the concept that direct interaction between acceptor glycans and leaving group phosphate is an important aspect of the catalytic mechanism. Because none of the other retaining GT-B NDP sugar-dependent glycotransferases (AGT, OtsA, or WaaG), all of which have been crystallized bound to their intact NDP-sugar substrates (Fig. 3C). However, none of these structures contain an acceptor or acceptor mimic in their active sites. In our EcGS structures, the terminal hydroxyl group of HEPPSO makes a direct interaction to the β-phosphate, mimicking the interaction seen in structures of MalP that contain both phosphate and acceptor glycans (23, 24) (and Protein Data Bank 2asv). It is possible that this additional interaction sufficiently stabilizes developing negative charge on the phosphate, or is even the proton donor, making it a good enough leaving group to allow the hydrolysis reaction we see in the active site to occur. If true, this would be consistent with the concept that direct interaction between acceptor glycans and leaving group phosphate is an important aspect of the catalytic mechanism. Because none of the other retaining GT-B NDP sugar-dependent glycotransferase structures contain an acceptor mimic to facilitate the reaction, their NDP-sugars remain intact.

Therefore the combination of the active site of EcGS with HEPPSO acting as an unreactive glycans significantly destabilizes the phosphate sugar covalent bond, leading to the abortive products we see. However, given the quite long incubation times required to produce diffraction quality crystals (4–12 weeks) it is impossible to know the time course of this process. Crystal nucleation may very well occur when the ADP-Glc is still intact, with decomposition occurring over time as the crystals grow. The same can be said regarding the occupancies of each constituent in the active site. Although ADP is fully occupied in all the structures, the occupancy of the HEPPSO and glucose moieties vary between 0.8–1 and 0.5–0.8, respectively. Given that both ADP-Glc and HEPPSO are absolutely required for the production of this crystal form, it is reasonable to assume that both are bound during nucleation and occupancy is partially lost during or after crystal growth. In support of this is the fact that HEPPSO occupancy falls below 1 in only one of the crystals and the ADP occupancy is 1 in all forms.

The Glucose Moiety—We have reported five separate structures from co-crystallization of EcGS with ADP-Glc and the acceptor mimic HEPPSO, one from a crystal harvested after 4 weeks, three from crystals that were harvested after 12 weeks, and the last from an E377A mutant EcGS. We were surprised to see that in none of these structures do we see an intact ADP-Glc molecule. In wtGSs, the crystal harvested in the shortest period of time, we see density for most of the glucose moiety, but do not see density for the O4 or O1 glucose oxygen atoms, and it is clear that the glucose moiety is not covalently linked to the ADP moiety. However, our data does not allow us to make a firm identification of the glucose-like moiety (Fig. 1D). What is clear, however, is that the density is not consistent with the hydrolysis product glucose, which is clearly seen in all three of the wild-type enzyme crystals harvested after a considerably longer time. We therefore believe that we have captured some intermediate form that will subsequently be hydrolyzed to the glucose we see in the three later structures. Two possible intermediate candidates are β-glucopyranosylum (DGM) and the DGM C2 deprotonation product d-arabino-Hex-1-enitol that contains a C1–C2 double bond resulting from DGM deprotonation at the C2 position (Fig. 1D).

The dissociation of glucose and ADP in our structures is not seen in the other GT-B fold retaining NDP-sugar-dependent glycotransferases (AGT, OtsA, or WaaG), all of which have been crystallized bound to their intact NDP-sugar substrates (Fig. 3C). However, none of these structures contain an acceptor or acceptor mimic in their active sites. In our EcGS structures, the terminal hydroxyl group of HEPPSO makes a direct interaction to the β-phosphate, mimicking the interaction seen in structures of MalP that contain both phosphate and acceptor glycans (23, 24) (and Protein Data Bank 2asv). It is possible that this additional interaction sufficiently stabilizes developing negative charge on the phosphate, or is even the proton donor, making it a good enough leaving group to allow the hydrolysis reaction we see in the active site to occur. If true, this would be consistent with the concept that direct interaction between acceptor glycans and leaving group phosphate is an important aspect of the catalytic mechanism. Because none of the other retaining GT-B NDP sugar-dependent glycotransferase structures contain an acceptor mimic to facilitate the reaction, their NDP-sugars remain intact.

Therefore the combination of the active site of EcGS with HEPPSO acting as an unreactive glycans significantly destabilizes the phosphate sugar covalent bond, leading to the abortive products we see. However, given the quite long incubation times required to produce diffraction quality crystals (4–12 weeks) it is impossible to know the time course of this process. Crystal nucleation may very well occur when the ADP-Glc is still intact, with decomposition occurring over time as the crystals grow. The same can be said regarding the occupancies of each constituent in the active site. Although ADP is fully occupied in all the structures, the occupancy of the HEPPSO and glucose moieties vary between 0.8–1 and 0.5–0.8, respectively. Given that both ADP-Glc and HEPPSO are absolutely required for the production of this crystal form, it is reasonable to assume that both are bound during nucleation and occupancy is partially lost during or after crystal growth. In support of this is the fact that HEPPSO occupancy falls below 1 in only one of the crystals and the ADP occupancy is 1 in all forms.

---

4 F. Sheng, X. Jia, A. Yep, J. Preiss, and J. H. Geiger, unpublished results.
Comparison of the donor glucose position in our structure to that of both the NDP-sugar dependent retaining GT-B glycosyltransferase (OtsA, AGT, and WaaG) and the phosphorylase structures (MalP and GP) is enlightening. The glucose in EcGS overlays very well with the equivalent glucose moieties in both MalP and GP, with very similar interactions between enzyme and the glucose oxygens (Fig. 3B). In the NDP-glucose-containing enzymes it is also true that most of the oxygens make similar interactions to enzyme. However, there is significant displacement of the glucose ring to accommodate the covalent bond between C1 and phosphate (Fig. 3C). This displacement is, not surprisingly, maximized at C1, where the difference in position between the C1 in glucose in our GS structures and the C1 in the OtsA ADP-glucose structure is 1.6 Å. However, the rest of the atoms in the ring are displaced far less, allowing them to make similar interactions in the active site in all the structures. For example, there is an intramolecular hydrogen bond between O4 of glucose and the α-phosphate in all of the NDP-sugar bound GT-B glycosyl transferase structures and this hydrogen bond is also seen between glucose O4 and the α-phosphate of ADP in our EcGS/ADP/glucose structures (Fig. 3C) (20, 21, 25).

Whether the active site contains ADP or an NDP-sugar, the positions of both phosphates are virtually identical in all of the NDP-sugar retaining GT-B glycosyltransferases, indicating that there is very little movement of the leaving group atoms during the reaction in these enzymes (Fig. 3C). The motion of the phosphate group has been beautifully characterized for the MalP enzyme (24). It would thus appear that in the case of the NDP-sugar enzymes the majority of the motion involves the C1 of glucose, with the diphosphate leaving group relatively static, whereas the phosphorylase enzymes involve more phosphate motion, consistent with the lessened stereochemical constraints on the motion of the phosphate in these enzymes (the substrate is free phosphate). This accentuates the fact that the presence or absence of a covalent linkage between the two phosphates results in significant changes in the chemistry of these enzymes, despite the close similarities of their active sites.

**Mechanistic Considerations**—The mechanism of the retaining GT-B enzymes is still an unsettled issue. The basic question is whether a double displacement mechanism occurs, where a covalent intermediate between enzyme and substrate is formed, or whether some form of oxocarbenium cation-like species is formed either as a transition state or short-lived intermediate (Fig. 4). Attempts to verify a double displacement mechanism for GT-B enzymes by labeling the catalytic nucleophile or trapping a covalent intermediate have so far been unsuccessful, but such a negative result certainly does not rule out double displacement (10). Our structures are in concert with the other known retaining GT-B structures in that the most well positioned candidate for the catalytic nucleophile in our structures is the main chain carbonyl of His-161. It is in fact in close proximity to the backside of C1 of the glucose moiety. It is significantly closer to the main chain carbonyl in our struc-

---

**FIGURE 4.** Proposed mechanisms of glucosyl transfer catalyzed by GT retaining enzymes. Two essential components of double-displacement, a catalytic nucleophile and a glucosyl-enzyme intermediate, are circled and framed, respectively. In $S_{n1}$- and $S_{n1}'$-like mechanisms, the positively charged DGM (framed) is stabilized by the leaving group AMP-phosphate and the incoming nucleophile 4-hydroxyl group of the sugar.
Crystal Structure of Closed Conformation of E. coli GS

structures than in the other NDP-sugar containing structures because C1 is still covalently bound to the β-phosphate in ADP located further above this main chain carbonyl in these structures. Whether it acts by forming an intermediate covalent bond or by transition state/intermediate stabilization via interaction at C1, this main chain carbonyl is likely to play an important role in the reaction for all of these enzymes.

The intermediate species we observe between the hydrolysis product glucose and substrate ADP-glucose is intriguing, but our inability to confidently identify this species does not allow a conclusion to be made regarding this species, save that it is in an almost identical location to that of glucose in our structures. A glucose-derived intermediate also lacking the 1-hydroxyl group was found at the superimposable position in the MalP active site in a 1.95-Å PLP-PO₄₋–maltopentose complex structure (PDB 2asv) (Fig. 3B), produced by rapid soaking and freezing. This intermediate was assigned as 1,5-anhydroscorbital, which would be the result of hydride reduction of the 2-glucopyranosylpyranosyl intermediate. Interestingly, the Geremia group also has twice the reported crystallographic isolation of an oxocarbenium cation intermediate using a carefully developed quick soak/freeze protocol, although the structure has not been formally published (48, 49). The isolation of such an intermediate would provide strong evidence for an S₁,1-like mechanism (Fig. 4).

Our structures are in concert with all other GT-B retaining enzyme structures in that an Arg and Lys (Arg-300 and Lys-305 in EcGS) directly interact with the leaving group phosphate (the β-phosphate of an NDP-sugar) and that a carboxylate (Glu-377 in EcGS) be in position to directly interact with the Lys. Even though these residues are not always conserved in sequence, their structural analogues are present in all GT-B retaining enzyme structures so far determined (21, 22, 25, 27, 47). In addition, HEPPSO, which binds in the acceptor glycan binding site in the EcGS structures, also makes a direct interaction with this β-phosphate, perhaps providing the critical interaction necessary for the slow ADP-Glc hydrolysis that occurs in the active site of the crystalline enzyme. Our conclusion is that although HEPPSO is quite different from a glycan, it is able to mimic the interaction with phosphate, activating it as a leaving group but not the nucleophilic attack, leading to the hydrolysis product we see. Consistent with the idea that interaction between leaving group phosphate and glycan substrate is important, both MalP structures that show both glycan and phosphate show tight interaction between the glycan terminal O4 and phosphate (2.4 and 2.5 Å in the two structures) (PDB 2asv and Ref. 24). Similarly, a close distance between the incoming nucleophile hydroxyl group of the acceptor sugar and the phosphate oxygen of the leaving group is also seen in the GT-B retaining enzyme OtsA (21).

In conclusion, our structures of EcGS have shown that the active site, once fully formed for GS, is remarkably similar to that of the other GT-B retaining enzymes whose structures are known. However, formation of this active site appears to require binding of both sugar nucleotide and glycan or glycan mimic to bring the two domains of the enzyme together and form a catalytically competent active site. This relatively extreme example of induced fit is in contrast to the other GT-B retaining enzymes where more local loop motions are responsible for closing the active site. Our structures also illustrate the dual role of Glu-377 in both maintaining the positive charge on Lys-305 and holding and orienting the donor glucose moiety in the active site. The signature of the mechanism is the precise cooperation between leaving group phosphate, donor sugar, and the incoming acceptor that requires an elaborate geometric arrangement of these elements and the active site of the enzymes residues. To date all GT-B retaining enzymes have displayed strikingly similar active sites (45), especially the catalytically essential residues, implying that a similar mechanism is employed by all.

Acknowledgments—We especially thank Dr. Zdzislaw Wawrzak at beamline DND-CAT of the Advanced Photon Source for help with data collection and processing. We also thank Dr. Xiangshu Jin for helpful discussions during model refinement.

REFERENCES

1. Preiss, J., and Romeo, T. (1989) Adv. Microbiol. Physiol. 30, 183–238
2. Preiss, J., and Sivak, M. (1999) in Comprehensive Natural Products Chemistry (Pinto, B. M., ed) pp. 441–495, Elsevier Science B.V., Amsterdam, The Netherlands
3. Ballicora, M. A., Iglesias, A. A., and Preiss, J. (2004) Photosynth. Res. 79, 1–24
4. Ballicora, M. A., Iglesias, A. A., and Preiss, J. (2003) Microbiol. Mol. Biol. Rev. 67, 213–225
5. Iglesias, A. A., and Preiss, J. (1992) Biochem. Educ. 20, 196–203
6. Boyer, C., and Preiss, J. (1977) Biochemistry 16, 3693–3699
7. Abad, M. C., Binderup, K., Rios-Stener, J., Armi, R. K., Preiss, J., and Geiger, J. H. (2002) J. Biol. Chem. 277, 42164–42170
8. Coutinho, P. M., Deleury, E., Davies, G. J., and Henrissat, B. (2003) J. Mol. Biol. 328, 307–317
9. Bowles, D., Hayenkova, J., Lim, E.-K., and Poppenberger, B. (2005) Curr. Opin. Plant Biol. 8, 254–263
10. Lairson, L. L., and Withers, S. G. (2004) Chem. Commun. (Camb.) 20, 2243–2248
11. Ünlügil, U. M., and Rini, J. M. (2000) Curr. Opin. Struct. Biol. 10, 510–517
12. Lairson, L. L., Chiu, C. P. C., Ly, H. D., He, S., Wakarchuk, W. W., Strynadka, N. C. J., and Withers, S. G. (2004) J. Biol. Chem. 279, 28339–28344
13. Sinnott, M. L., and Jencks, W. P. (1980) J. Am. Chem. Soc. 102, 2026–2032
14. Banait, N. S., and Jencks, W. P. (1991) J. Am. Chem. Soc. 113, 7951–7958
15. Banait, N. S., and Jencks, W. P. (1991) J. Am. Chem. Soc. 113, 7978–7986
16. Chiappe, C., Moro, G. L., and Munforde, P. (1997) Tetrahedron 53, 10471–10478
17. Klein, H. W., Im, M. J., and Palm, D. (1986) Eur. J. Biochem. FEBS 157, 107–114
18. Sinnott, M. L. (1990) Chem. Rev. 90, 1171–1202
19. Patenaude, S. I., Seto, N. O. L., Borisova, S. N., Szpacenko, A., Marcus, S. L., Palec, M. M., and Evans, S. V. (2002) Nat. Struct. Biol. 9, 685–690
20. Gibson, R. P., Tarling, C. A., Roberts, S., Withers, S. G., and Davies, G. J. (2004) J. Biol. Chem. 279, 1950–1955
21. Gibson, R. P., Turkenburg, J. P., Charnock, S. J., Lloyd, R., and Davies, G. J. (2002) Chem. Biol. 9, 1337–1346
22. Watson, K. A., Schinzel, R., Palm, D., and Johnson, L. N. (1997) EMBO J. 16, 1–14
23. Watson, K. A., McCleverty, C., Geremia, S., Cottaz, S., Driguez, H., and Johnson, L. N. (1999) EMBO J. 18, 4619–4632
24. Geremia, S., Campagnolo, M., Schinzel, R., and Johnson, L. N. (2002) J. Mol. Biol. 322, 413–423
25. Lariviere, L., Sommer, N., and Morera, S. (2005) J. Mol. Biol. 352, 139–150
26. Goldsmith, E. J., Sprang, S. R., Hamlin, R., Xuong, N. H., and Fletterick, R. J. (1989) Science 245, 528–532
Crystal Structure of Closed Conformation of E. coli GS

27. Sprang, S. R., Withers, S. G., Goldsmith, E. J., Fletterick, R. J., and Madsen, N. B. (1991) *Science* **254**, 1367–1371
28. Buschiazzo, A., Ugalde, J. E., Guerin, E. M., Shepard, W., Ugalde, R. A., and Alzari, P. M. (2004) *EMBO J.* **23**, 3196–3205
29. Cristina, H., Guinovart, J. I., Fita, I., and Ferrer, J. C. (2006) *J. Biol. Chem.* **281**, 2923–2931
30. Yep, A., Ballicora, A. M., Sivak, N. M., and Preiss, J. (2004) *J. Biol. Chem.* **279**, 8359 – 8367
31. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* **276**, 307–326
32. Howard, A. J. (2000) in *Crystallographic Computing 7: Proceedings from the Macromolecular Crystallographic Computing School* (Bourne, P. E., and Watenpaugh, K. D., eds) pp. 1–12, Oxford University Press, Oxford
33. Brünger, A. T. (1992) *Nature* **355**, 472–475
34. Delano, W. L. (2002) *The PyMOL Molecular Graphics System*, DeLano Scientific, San Carlos, CA
35. Fox, J., Kawaguchi, K., Greenberg, E., and Preiss, J. (1976) *Biochemistry* **15**, 849 – 857
36. Qi, G., Lee, R., and Hayward, S. (2005) *Bioinformatics* **21**, 2832–2838
37. Yep, A., Ballicora, M., and Preiss, J. (2006) *Arch. Biochem. Biophys.* **453**, 188–196
38. Furukawa, K., Tagaya, M., Inouye, M., Preiss, J., and Fukui, T. (1990) *J. Biol. Chem.* **265**, 2086–2090
39. Furukawa, K., Tagaya, M., Tanizawa, K., and Fukui, T. (1993) *J. Biol. Chem.* **268**, 23837–23842
40. Cleland, W. W. (1963) *Biochim. Biophys. Acta* **67**, 173–187
41. Cleland, W. W. (1967) *Annu. Rev. Biochem.* **36**, 77–124
42. Schinzel, R., and Palm, D. (1990) *Biochemistry* **29**, 9956–9962
43. Cid, E., Gomis, R. R., Geremia, R. A., Guinovart, J. I., and Ferrer, J. C. (2000) *J. Biol. Chem.* **275**, 33614–33621
44. Boix, E., Yingnan, Z., Swaminathan, G. J., Brew, K., and Acharya, K. R. (2002) *J. Biol. Chem.* **277**, 28310–28318
45. Yep, A. A., Ballicora, M., and Preiss, J. (2004) *Biochem. Biophys. Res. Commun.* **316**, 960–966
46. Sprang, S., Goldsmith, E., and Fletterick, R. (1987) *Science* **237**, 1012–1019
47. Martinez-Fleites, C., Proctor, M., Roberts, S., Bolam, D. N., Gilbert, H. J., and Davies, G. J. (2006) *Chem. Biol.* **13**, 1143–1152
48. Campagnolo, M., Campa, C., Zorzi, R. D., Wuerges, J., and Geremia, S. (2008) *Arch. Biochem. Biophys.* **471**, 11–19
49. Geremia, S., Campagnolo, M., Demitri, N., and Johnson, L. N. (2006) *Structure* **14**, 393–400
50. Jones, T. A. (1985) *Methods Enzymol.* **115**, 157–171