Structural Enzymology of Cellvibrio japonicus Agd31B Protein Reveals \( \alpha \)-Transglucosylase Activity in Glycoside Hydrolase Family 31\(^*\)\(^{3,5}\)

Johan Larsbring\(^1\), Atsushi Izumi\(^5\), Glyn R. Hemsworth\(^4\), Gideon J. Davies\(^3,6\), and Harry Brumer\(^1\)*\(^{3,5}\)

From the \(^4\)Division of Glycoscience, School of Biotechnology, Royal Institute of Technology, AlbaNova University Centre, 106 91 Stockholm, Sweden, \(^5\)York Structural Biology Laboratory, Department of Chemistry, The University of York, York Yo10 Sdd, United Kingdom, and \(^6\)Michael Smith Laboratories and Department of Chemistry, University of British Columbia, 2185 East Mall, Vancouver, British Columbia V6T 1Z4, Canada

The metabolism of the storage polysaccharides glycogen and starch is of vital importance to organisms from all domains of life. In bacteria, utilization of these \( \alpha \)-glucans requires the concerted action of a variety of enzymes, including glycoside hydrolases, glycoside phosphorylases, and transglycosylases. In particular, transglycosylases from glycoside hydrolase family 13 (GH13) and GH77 play well established roles in \( \alpha \)-glucan side chain (de)branching, regulation of oligo- and polysaccharide chain length, and formation of cyclic dextrans. Here, we present the biochemical and tertiary structural characterization of a new type of bacterial \( 1,4-\alpha \)-glucan \( 4-\alpha \)-glucosyltransferase from \( \text{Agd31B} \).

†1 Both authors contributed equally to this work.

‡1 Work in Stockholm was supported by the Mizutani Foundation for Glycoscience and The Swedish Research Council Formas (via the European Union WoodWisdom-Net ERA-NET project “FiberSurf: New biotechnological tools for wood fiber surface modification and analysis” and via “CarboMat-the KTH Advanced Carbohydrate Materials Centre”), and work in York was supported by Biotechnology and Biological Sciences Research Council Grant BB/I014802/1.

§1 This article contains supplemental Figs. S1–S5.

The atomic coordinates and structure factors (codes 4b9y, 4b9z, and 4ba0) have been deposited in the Protein Data Bank (http://wwpdb.org/).

1 To whom correspondence may be addressed. Tel.: 44-1904-328260; Fax: 44-904-328266; E-mail: gideon.davies@york.ac.uk.

2 To whom correspondence may be addressed. Tel.: 44-1904-328260; Fax: 44-904-328266; E-mail: gideon.davies@york.ac.uk.

3 To whom correspondence may be addressed. Tel.: 604-822-3738; Fax: 604-822-2114; E-mail: brumer@msl.ubc.ca.

Glycogen is a highly branched, mixed linkage \( \alpha(1\rightarrow4)/\alpha(1\rightarrow6) \)-glucan polymer that serves as a readily accessible, osmotically neutral, cellular energy reserve in all domains of life (1–3). Glycogen is structurally related to amylpectin, which together with the linear polysaccharide amylose (\( \alpha(1\rightarrow4) \)-glucan) comprises the plant storage reserve starch (4). Prokaryotic and eukaryotic glycogen biosynthesis and degradation are a complex, highly conserved, and tightly controlled process involving a myriad of enzymes and regulatory factors (2, 3). In bacteria such as \( \text{Escherichia coli} \), glycogen is synthesized from ADP-glucose by the combined action of glycogen synthase, which builds linear \( \alpha(1\rightarrow4) \)-glucan chains, and glycogen branching enzyme, which catalyzes chain rearrangement via \( \alpha(1\rightarrow4) \)-to-\( \alpha(1\rightarrow6) \) transglucosylation, thereby yielding polydisperse molecules with molar masses of up to \( 10^8 \) Da (2, 3). In turn, catabolism under carbon-limited conditions occurs via the sequential action of glycogen phosphorylase and debranching enzyme to yield glucose 1-phosphate.

In bacteria, glycogen metabolism is closely linked to the metabolism of storage maltodextrins that involves the buildup and rearrangement of linear \( \alpha(1\rightarrow4) \)-gluco-oligosaccharides by transglycosylation (5). This process has been well described in \( \text{E. coli} \) in which the amylomaltsase MalQ, a member of glycoside hydrolase family 77 (GH77\(^3\), Ref. 6), catalyzes the transfer of a \( 4-\alpha \)-glucanosyl fragment from the non-reducing end of maltoligosaccharide donor substrates and possibly the disaccharide malto- \( \text{malto-oligosaccharide acceptors (4-\alpha \text{-glucano-}} \)

4 The abbreviations used are: GH, glycoside hydrolase family; HPAEC-PAD, high performance anion exchange chromatography with pulsed amperometric detection; \( \text{CJ, C. japonicus} \); Bistris propane, 1,3-bis(tris(hydroxymethyl)methylamino)propane; pNP, p-nitrophenol; \( \text{Ro, R. obeum} \).
bacterial 4-α-glucanotransferases of GH13 catalyze the disproportionate of maltotriose (10, 11), maltotetraose (12, 13), and longer 4-α-glucan chains.

Indeed, transglycosylation reactions leading to the rearrangement of α-glucans are widespread among bacteria. Glycosgen branching and debranching aside, diverse enzymes with 4-α-glucanotransferase activity (α(1→4)-glucan:α(1→4)-glucan transferase activity) can produce a range of linear and cyclic maltodextrin products via freely reversible disproportionation and cyclization reactions, respectively (9, 14). In addition to the aforementioned CtsY and CtsZ transglucosylase with strict transglucosylase activity (GH31 family), a number of retaining α-glucanotransferase (EC 2.4.1.19) of GH13 represents a previously undiscovered activity in GH31. Crystallography of the enzyme in free, acarbose-complexed, and trapped 5-fluoro-β-D-glucopyranosyl-enzyme intermediate forms has highlighted the structural basis for the strict transfer of a single glucosyl residue and preference for maltotriose and longer substrates. Taken together, the data suggest a biological role for CjAgd31B in glycogen or maltodextrin metabolism that may be complementary to that predicted for the GH77 homologue CjMal77Q.

EXPERIMENTAL PROCEDURES

Curve fitting and processing of kinetics data were performed using Origin 8 software (OriginLab). p-Nitrophenyl (pNP) α-glycosides, sucrose, D-maltose, and starch from corn were purchased from Sigma. Malto-oligosaccharides (maltotriose to maltohexaose), isomaltose, melibiose, and acarbose were purchased from CarboSynth. α-Glucosyl fluoride and 5-fluoro-α-D-glucopyranosyl fluoride were kind gifts from Professor Steven Withers (Department of Chemistry, University of British Columbia, Canada). Ultrapure water was used in all experiments and refers to water purified on a Milli-Q system (Millipore) with a resistivity (ρ) >18.2 megaohms cm.

Cloning of CjAgd31B

The open reading frame encoding CjAgd31B (GenBank accession number ACE84782.1) was amplified by PCR from genomic DNA of C. japonicus Ueda107 using Phusion polymerase (Finnzymes) and the following primers (Thermo Fischer Scientific): 5'-CACCATGAATCCGGTCAAACG-3' and 5'-ATGCAACCTG-NP-AAGTTAAGCGCTTC-3' with the forward primer incorporating the CACC overhang (underlined) needed for TOPO cloning and excluding the predicted signal peptide (cleavage site between amino acid residues 24 and 25). The PCR product was cloned into the pENTR/SDF-TOPO entry vector (Invitrogen) and recombined into the pET-DEST42 destination vector (Invitrogen) as described previously (26).

Gene Expression and Protein Purification

Plasmids harboring the CjAgd31B gene were transformed into E. coli BL21(DE3) by electroporation, the gene was expressed and the resulting protein was purified by immobilized metal affinity chromatography following an established protocol (26). Analysis by SDS-PAGE showed the protein to be electrophoretically pure. LC electrospray ionization MS was used for protein molar mass determination as described previously (28). For crystallization studies, the protein was further purified by size exclusion chromatography and ion exchange chromatography. The eluted protein solution was concentrated to 5 ml by a Vivaspin 20 concentrator (Sartorius Stedim Biotech) and loaded onto a HiLoad 16/60 Superdex 200 prep grade column (GE Healthcare) equilibrated with 20 mM Tris (pH 8.0), 300 mM sodium chloride. The eluted protein solution was dialyzed into 20 mM Tris (pH 8.0) at 4°C for 16 h. The dialyzed protein solution was loaded onto a Resource Q column (GE Healthcare) equilibrated with 20 mM Tris (pH 8.0) and eluted with a linear gradient of 20 mM Tris (pH 8.0), 400 mM sodium chloride. Two major peaks of CjAgd31B were obtained, and the peak eluted in lower salt concentration was collected and used for protein crystallization. Protein concentrations were deter-
A Bacterial \(\alpha\)-Transglucosylase from GH31

mined from \(A_{280}\) values of suitably diluted samples using an extinction coefficient of 139,245 M\(^{-1}\) cm\(^{-1}\) as calculated by the ProtParam tool on the ExPASy server (29).

**Thin Layer Chromatography (TLC)**

TLC was performed using normal phase silica on aluminum plates eluted with acetonitrile–water (2:1). Analytes were visualized by immersion in 8% \(\text{H}_2\text{SO}_4\) in ethanol followed by charring.

**High Performance Anion Exchange Chromatography with Pulsed Amperometric Detection (HPAEC-PAD)**

Oligo- and monosaccharides were analyzed on a Dionex ICS-3000 HPLC system operated by Chromelion software version 6.80 (Dionex) using a Dionex Carbopac PA200 column. Solvent A was water, solvent B was 1 M sodium hydroxide, and solvent C was 1 M sodium acetate.

**Gradient A**—Conditions used were 0–5 min, 10% B, 2% C; 5–12 min, 10% B and a linear gradient from 2 to 30% C; 12–12.1 min, 50% B, 50% C; 12.1–13 min, an exponential gradient of B and C back to the initial conditions; 13–17 min, initial conditions.

**Gradient B**—Conditions used were 0–4 min, 10% B, 5% C; 4–8 min, 10% B and a linear gradient from 5 to 25% C; 8–8.1 min, 50% B, 50% C; 8.1–9 min, an exponential gradient of B and C back to the initial conditions; 9–13 min, initial conditions.

**Gradient C**—Conditions used were 0 to 4 min, 10% B, 6% C; 4 to 17 min, 10% B and a linear gradient from 6–25% C; 17 to 17.1 min, 50% B, 50% C; 17.1 to 18 min, an exponential gradient of B and C back to initial conditions; 18 to 22 min, initial conditions.

**Gradient D**—Conditions used were 0–4 min, 10% B, 6% C; 4–10 min, 10% B and a linear gradient from 5 to 25% C; 10–10.1 min, 50% B, 50% C; 10.1–11 min, an exponential gradient of B and C back to the initial conditions; 11–15 min, initial conditions.

**Gradient E**—Conditions used were 0–4 min, 10% B, 6% C; 4–15 min, 10% B and a linear gradient from 5 to 25% C; 15–15.1 min, 50% B, 50% C; 15.1–16 min, an exponential gradient of B and C back to the initial conditions; 16–20 min, initial conditions.

**pH-Rate Profile**

Measurements of the pH dependence of \(Cj\text{Agd31B}\) catalysis were carried out using maltose as substrate in an HPAEC-PAD-based transglycosylation assay described below. The buffers used (50 mM) were sodium citrate (pH 3–6.5), sodium phosphate (pH 6.5–8), glycylglycine (pH 8–9), and glycine (pH 9–10) (supplemental Fig. S1).

**Enzyme Assays**

Activity on pNP-glycosides was analyzed by a stopped assay as described previously (26) using an enzyme concentration of 6.5 \(\mu\)M. The transglycosylation activity of \(Cj\text{Agd31B}\) on various oligosaccharides was performed in 100-\(\mu\)l reactions at 25 °C in 50 mM citrate buffer (pH 6). For initial rate saturation kinetics experiments, \(Cj\text{Agd31B}\) was added to a final concentration of 1.6 \(\mu\)M for maltose and 270 pm for maltotriose, maltotetraose, and maltopentaose. Reactions typically proceeded for 10 min and were stopped by addition of 4 \(\mu\)l of 5 M sodium hydroxide.

HPAEC-PAD was used for product analysis using Gradients A, B, D, and E for reactions on maltose, maltotriose, maltotetraose, and maltopentaose, respectively. Commercial malto-oligosaccharides (maltose to maltohexaose) were used as standards.

To test different acceptors, starch was used as a glucosyl donor. Starch from corn was dissolved to 1% (w/v) in water followed by dialysis in deionized water using a 5-kDa-cutoff membrane to remove monosaccharides and small oligosaccharides. 50-\(\mu\)l reactions containing 0.4% starch as donor, 1 mM acceptor (glucose or isomaltose), and 2 \(\mu\)M enzyme were incubated at 25 °C for 10 min and terminated by addition of 3 \(\mu\)l of 5 M sodium hydroxide. Products were analyzed by HPAEC-PAD using Gradient C.

**IC\(_{50}\) Measurements**

The inhibition of \(Cj\text{Agd31B}\) by acarbose (0–1050 \(\mu\)M) was determined by using maltotriose (100 \(\mu\)M) as a substrate in reactions as described above. Product formation was analyzed by HPAEC-PAD (Gradient B).

**Crystallization and Data Collection**

\(Cj\text{Agd31B}\) was stored in 5 mM Bistris propane (pH 8.5) and concentrated to 7 mg/ml by using a Vivaspin 20 concentrator. In initial crystal screens using Crystal Screen HT, Index HT, SaltRx HT (Hampton Research), and modified Newcastle Screen prepared at the York Structural Biology Laboratory, small single crystals were obtained in several conditions. Well diffracting crystals were obtained after 3–4 days in 1.8 M ammonium sulfate, 0.1 M HEPES (pH 7.0), 2% PEG 400 at 20 °C by the sitting drop vapor diffusion method. The structure was solved using experimental phasing with an iodine derivative. This was prepared by placing \(\sim 1 \mu\)l of a 0.25 g/ml potassium iodide solution into a 2-\(\mu\)l crystallization droplet to allow slow diffusion of iodine into the crystal. Crystallization droplets with iodine solution were left at 20 °C for 16 h prior to freezing and data collection. For the complex structures, crystals were soaked in 1.8 M ammonium sulfate, 0.1 M HEPES (pH 7.0), 2% PEG 400 with either 5 mM 5-fluoro-\(\alpha\)-d-glucopyranosyl fluoride or 5 mM acarbose for 1 h at 20 °C.

All crystals were cryoprotected by 2.0 mM lithium sulfate, 0.1 M HEPES (7.0), 2% PEG 400. The x-ray data for the free enzyme and iodine-soaked crystals were collected at 100 K on an ADSC Q315 charge-coupled device detector on beamline I02 of the Diamond Light Source. X-ray data for the complex with acarbose and the 5F\(\beta\)Glc-enzyme were collected at 100 K on an ADSC Q315 charge-coupled device detector at BL-ID14-4 and BL-ID29 at the European Synchrotron Radiation Facility, respectively. The details of the data collections are listed in Table 1. All data were processed using iMOSFLM (30) and programs from the CCP4 suite (31) unless otherwise stated. The statistics of the data processing and structure refinement are listed in Table 1.

Experimental phasing was performed by single wavelength anomalous diffraction methods at a wavelength of 1.8 Å. Heavy atom substructure solution and initial phasing was performed on the 2.9-Å resolution iodine-derivatized crystal data with autoSHARP (32) followed by phase extension with the 1.9-Å free enzyme data set using DM (33). The 1.9-Å data were used.
encodes two other GH31 members, the putative C. japonicus receptors, transporters, and extracellularly. In addition to thus likely to be localized in the periplasm or to be secreted genomic neighbors include a predicted genome of DECEMBER 21, 2012 • VOLUME 287 • NUMBER 52

| TABLE 1 |
| X-ray data collection and refinement statistics |
| OXL, oxalate; EDO, 1,2-ethanediol; PEG, 2-(2-hydroxyethoxy)ethanol; PG4, 2-[2-(2-hydroxyethoxy)ethoxy] ethanol; PG4, 2-[2-(2-hydroxyethoxy)ethoxy] ethanol; r.m.s.d., root mean square; 5FβGlc, 5-fluoro-β-glucosyl; ESRF, European Synchrotron Radiation Facility; SO4, sulfate. |
| Name (Protein Data Bank code) |
| Free enzyme (4b9y) | 5FβGlc (4ba0) | Acarbose (4b9r) |
| Data collection |
| Beamline | Diamond I02 | ESRF ID29 | ESRF14-4 |
| Space group | P622 | P622 | P622 |
| Cell dimensions |
| a, b, c (Å) | 197.3, 197.3, 103.0 | 197.2, 197.3, 103.1 | 196.9, 196.9, 102.8 |
| α, β, γ (°) | 90, 90, 120 | 90, 90, 120 | 90, 90, 120 |
| Resolution (Å) (outer shell) | 49.85–1.90 (2.00–1.90) | 44.48–1.85 (1.95–1.85) | 51.38–2.0 (2.11–2.00) |
| Rmerge | 0.10 (0.52) | 0.10 (0.52) | 0.11 (0.46) |
| l/σL | 21.5 (6.0) | 19.8 (5.9) | 22.2 (7.4) |
| Completeness (%) | 100.0 (100.0) | 100.0 (100.0) | 100.0 (100.0) |
| Redundancy | 22.0 | 20.8 | 21.5 |
| Refinement |
| Resolution (Å) | 49.3–1.90 | 44.5–1.85 | 49.7–2.00 |
| No. reflections | 92,667 | 100,216 | 79,078 |
| Rmerge | 0.16/0.19 | 0.17/0.19 | 0.16/0.19 |
| No. atoms |
| Protein | 6,273 | 6,270 | 6,251 |
| Ligand | 93 | 80 | 142 |
| Water | 647 | 557 | 444 |
| B-factors (Å2) |
| Protein (TLS refinement) | 27.9 | 31.2 | 29.3 |
| Water | 33.8 | 36.8 | 32.4 |
| Ligands | SO4, 80.8; PG4, 44.8; OXL, 42.7; EDO, 48.2 | 5FβGlc, 21.2; SO4, 71.3; PGE, 47.7; Arg, 48.2; PEG, 34.4; EDO, 53.8 | Acarbose, 40.2; SO4, 55.3; OXL, 55.1; PEG, 37.5; EDO, 53.1 |
| r.m.s.d. |
| Bond lengths (Å) | 0.007 | 0.006 | 0.007 |
| Bond angles (°) | 1.1 | 1.1 | 1.1 |

α-glucosidase CjAgd31A and the biocatalytically and structurally characterized α-xylanosidase CjXyl31A (26). Although all are members of GH31, CjAgd31B has a low sequence similarity to both CjAgd31A and CjXyl31A with amino acid identities of 28 and 27% and similarities of 43 and 45%, respectively. From our recent phylogenetic analysis (26), the biocatalytically characterized member of GH31 most similar to CjAgd31B is YihQ from E. coli (sequence identity, 28%; similarity, 44%). E. coli YihQ has been annotated as an α-glucosidase based on a weak ability to hydrolyze the artificial substrate α-glucosyl fluoride, although the enzyme was impotent toward a range of other α-glucosides (38).

**Gene Expression**—A gene construct encoding CjAgd31B with a C-terminal hexahistidine tag and lacking the predicted native signal peptide was expressed in E. coli BL21(DE3) cells. The protein product was purified by immobilized metal affinity chromatography for kinetics analyses and additionally by size exclusion chromatography and ion exchange for crystallization studies; purity in both cases was confirmed by SDS-PAGE (data not shown). The molar mass of CjAgd31B, corresponding to the C-terminal His-tagged enzyme starting at Asn-25 (the natural site of signal peptide cleavage), was verified by LC electrospray ionization MS (expected, 49,478.8 Da; observed, 49,478.1 Da; supplemental Fig. S3). The overall yield was typically around 100 mg/liter of culture broth.

**Transglycosylation Activity on Malto-oligosaccharides**—Based on membership in GH31, the substrate specificity of CjAgd31B was initially tested using pNP-α-glucoside and pNP-α-xylanoside; the enzyme showed no apparent liberation of the aglycone from either of these substrates after extended incubation (1 mm substrate and 6.5 μM enzyme and up to 4 h incubation). The enzyme also displayed no detectable activity on sucrose, meli-
biose, isomaltose, or α-glucosyl fluoride (a substrate for E. coli YihQ (38)). On maltose (Glc\textsubscript{pca}(\textsubscript{1→4})Glc), however, formation of both glucose and longer oligosaccharides could be observed using TLC (data not shown). The transglycosylation potential of CjAgd31B was confirmed by HPAEC-PAD following incubation of 6.5 μM enzyme with 10 mM maltose for 30 min, which led to a buildup of malto-oligosaccharides. Products with a degree of polymerization of up to 14 glucose residues could be detected (Fig. 1A).

To further analyze the catalytic properties of CjAgd31B, HPAEC-PAD was used to measure product formation from reactions on malto-oligosaccharides. Under conditions of low substrate conversion (<10% of substrate consumed), the enzyme was shown to transfer a single glucose moiety from a donor to an acceptor molecule. Incubation of the enzyme with linear malto-oligosaccharides (maltotriose to maltopentaose) exclusively yielded Glc\textsubscript{u} – 1 and Glc\textsubscript{u} + 1 products via transglycosylation (Fig. 1B). On maltotriose, maltotetraose, and maltopentaose, the production of glucose, which would indicate competing substrate hydrolysis, was not observed under initial rate conditions. Thus, apparent Michaelis-Menten kinetics parameters for these substrates could be directly determined from plots of $v_0/\left[E\right]$ versus $[Glc]_u$ (Fig. 2A) where the rate of transglycosylation product formation is given by Equation 1 (in this case, the rate of Glc\textsubscript{u} – 1 formation can also be used). The best substrate for the enzyme was maltotriose with a $k_{cat}/K_m$ app value of 196 s\textsuperscript{-1}·mm\textsuperscript{-1}, whereas maltotetraose and maltopentaose displayed comparable $k_{cat}/K_m$ app values of 72 and 58 s\textsuperscript{-1}·mm\textsuperscript{-1}, respectively (Table 3).

$$v_{\text{transglycosylation}} = \frac{d[Glc_u + 1]}{dt} \quad (\text{Eq. 1})$$

Hydrolysis could, however, be detected when maltose was used as a substrate. Here, the production of Glc was measurably higher than the 1:1 stoichiometric ratio of Glc to Glc\textsubscript{3} expected for disproportionation (Glc\textsubscript{2} → Glc\textsubscript{3} + Glc). In this case, the velocity of the transglycosylation reaction was directly measured according to Equation 1 by quantifying the maltotriose produced. Determination of the hydrolytic rate required subtraction of the amount of glucose co-produced by dispropor-
Transglycosylating activity of CjAgd31B on various substrates

\[ \Delta G^\circ = R T \ln \left( \frac{k_{\text{cat}}}{K_m} \frac{[\text{maltotriose}]}{[\text{maltooligosaccharides}]} \right) \]

No activity was detected on pNP-\( \alpha \)-Glc, pNP-\( \alpha \)-Xyl, \( \alpha \)-glucosyl fluoride, isomaltose, sucrose, or melibiose.

| Substrate          | \( k_{\text{cat}}/K_m \) (s\(^{-1}\)) | \( K_m \) (mM) | \( k_{\text{cat}}/K_m \) (s\(^{-1}\)mM\(^{-1}\)) | \( \Delta G^\circ \) (kJ mol\(^{-1}\)) |
|--------------------|--------------------------------------|---------------|-----------------------------------------------|-------------------------------------|
| Maltose            | 341 ± 12                             | 8.8 ± 0.87    | 38.8                                          | 4.17                                |
| Maltotriose        | 239 ± 8.3                            | 1.2 ± 0.13    | 195.5                                         | 2.57                                |
| Maltotetraose      | 123 ± 9.0                            | 1.7 ± 0.29    | 72.3                                          | 2.57                                |
| Maltopentaose      | 181 ± 17                             | 3.1 ± 0.55    | 57.8                                          | 3.14                                |

Inhibition by Acarbose—The inhibitory effect of the pseudotetrasaccharide acarbose, a common \( \alpha \)-glucanase and \( \alpha \)-glucosidase inhibitor, was assayed using maltotriose as the substrate at a fixed concentration of 100 \( \mu \)M. The IC\(_{50}\) value was determined to be 75.1 ± 3.4 \( \mu \)M by plotting the relative activity versus the concentration of acarbose and fitting Equation 3 by non-linear regression (Fig. 4). With reactions performed at a substrate concentration of 100 \( \mu \)M, which is much lower than the apparent \( K_m \) value (1.2 mM), the IC\(_{50}\) value is approximately equal to the \( K_m \) value. The \( K_i \) may be more accurately calculated using Equation 4 (40), which yielded 81.4 \( \mu \)M.

\[
k_{\text{obs}} = \frac{k_{\text{max}}}{1 + \frac{[S]}{IC_{50}}} \quad \text{(Eq. 3)}
\]

\[
K_i = \frac{IC_{50}}{1 + \frac{[S]}{K_m}} \quad \text{(Eq. 4)}
\]

The calculated \( K_i \) value of CjAgd31B is similar to \( K_i \) values for acarbose with other GH31 enzymes such as the human maltase–glucoamylase N-terminal subunit (62 \( \mu \)M; Ref. 41) and sucrase N-terminal subunit (14 \( \mu \)M; Ref. 42), whereas it is sig-
significantly higher than the $K_i$ values for the human maltase-glucosaminase C-terminal subunit (1.72 \text{M}; Ref. 43), the *Gracilariopsis lemaneiformis* -glucan lyase (0.02 \text{M}; Ref. 44), and the *Thermoplasma acidophilum* AglA -glucosidase (2.99 \text{M}; Ref. 23).

**Tertiary Structures of CjAgd31B**—Three tertiary structures of CjAgd31B were obtained at 1.9-, 2.0-, and 1.85-Å resolution, respectively: the free enzyme, a non-covalent complex with the inhibitor acarbose, and a trapped 5-fluoro-D-glucopyranosyl-enzyme intermediate. All crystals contained one molecule in the asymmetric unit. The CjAgd31B structures consisted of a typical GH31 fold comprising four domains with two insertions (Fig. 5A): the N-terminal domain (N-terminal; residues 35–240), the catalytic (\(\beta/\alpha\)_8) domain (residues 241–586) with insertion domain 1 (Insert 1; residues 345–384) and insertion domain 2 (Insert 2; residues 415–435), the C-terminal proximal domain (C-proximal; residues 587–667), and the C-terminal distal domain (C-distal; residues 668–817). In all structures, the electron density map of the 10 N-terminal residues from 25 to 34 and C-terminal residues from 818 to the end (859), including the V5 epitope and His tag provided by the expression vector, were disordered. The free and acarbose structures had a disordered region from 137 to 140, and the 5-fluoro-D-glucosyl-enzyme structure had a disordered region from 139 to 140.

The visible secondary structures of the four domains, including the insertion domains, in CjAgd31B were well conserved with the human sucrase-isomaltase (Protein Data Bank code 3lpp) of GH31 with a root mean square deviation of 1.9–2.4 Å.

The free enzyme structure reveals a water-lined pocket where the conserved catalytic aspartic acid residues (Asp-412 and Asp-480) are located (45, 46). The pocket was 318 Å³ as calculated by the Pocket-Finder server (47) with a depth of 11 Å. The active site pocket was composed of residues Tyr-179, Phe-271, Asp-299, Leu-300, Ile-307, Met-311, Ile-341, Glu-343, Tyr-376, Phe-377, Trp-410, Asp-412, Leu-413, Glu-417, Arg-463, Trp-477, Asp-480, Asp-509, Phe-513, Arg-538, His-540, and Gln-542 (supplemental Fig. S4). To define enzyme-substrate interactions, two complexes, with the inhibitory tetrasaccharide acarbose (see above) and the covalent 5-fluoro-D-glucopyranosyl-enzyme intermediate, were obtained.

**Structure of the Acarbose Complex**—The CjAgd31B complex with acarbose revealed clear, unambiguous density for the inhibitor in the subsites (Fig. 5B; subsite nomenclature according to Ref. 39). In the −1 subsite, the enzyme-derived nucleophile, Asp-412, is indeed poised for nucleophilic attack, lying 3.2 Å “above” the pseudo-anomeric carbon of acarbose and with a nucleophile-C1-NH angle of 164.1°. The catalytic acid, Asp-480, lies 2.5 Å from the “interglycosidic” nitrogen of acarbose as expected. The hydrophobic residues Leu-300, Ile-341, Trp-410, Trp-477, Phe-271, and Phe-513 all lie within a 4-Å distance from the −1 subsite pseudosugar of acarbose (supplemental Fig. S4).
The +1 subsite contains the 6-deoxyglucosyl moiety of acarbose. Hydrogen bonds are made to Arg-463, Glu-417, and a water molecule, and enzyme-substrate distances suggest van der Waals contacts to Phe-377. In most other solved structures of GH31, the hydrogen bonds provided here by Glu-417 are made instead by an aspartate residue from a loop in the N-terminal domain (42, 48–50); the other exception to this is found in CjXyl31A in which a PA14 domain insert in the N-terminal domain extends the active site (26, 42, 48–50).

The substrate-interacting residues in the +1 and +2 subsites of the CjAgd31B structure are essentially homologous to those in the human maltase-glucoamylase and the Ro-αG1β glucosidase from Ruminococcus obeum with the exception of Phe-271, Leu-300, Phe-377, and Leu-413, which instead are Trp, Ile, Trp, and Met, respectively, in the maltase-glucoamylase and Ro-αG1 structures (supplemental Fig. S5 and Refs. 48 and 49). These side chains all make van der Waals contacts to acarbose as well as the 5-fluoro-β-glucosyl residue (see below). It is possible that Phe-271 of CjAgd31B, corresponding to Trp-169 and Tyr-299 in Ro-αG1 and maltase-glucoamylase, respectively, contributes to substrate specificity as wild type Ro-αG1 prefers isomaltose to maltose as a substrate, whereas the W169Y mutant inverts this preference (49).

Notable features of the +2 and +3 subsites are the tyrosine clamp of Tyr-179 (part of a long loop extending from the N-terminal domain) and Tyr-376 of Insert 1, which together form van der Waals contacts to the internal glucose moiety (third ring) of acarbose (cf. Fig. 5B and supplemental Fig. S4). A similar hydrophobic clamp has not been found in other GH31 structures apart from the PA14-mediated protein-sugar interaction in CjXyl31A (26).
A bacterial α-transglucosylase from GH31

Structure of the Trapped Covalent 5-Fluoro-β-glucosyl-enzyme Intermediate—To assess factors that may lead to the strict transglycosylation activity of CjAgd31B, a near mimic of the covalent glycosyl-enzyme intermediate was accessed using a classic “Withers” reagent, 5-fluoro-α-D-glucopyranosyl fluoride. The electron density map clearly reveals the trapped 5-fluoro-β-D-glucopyranosyl-enzyme observed in 1S3, skew boat conformation (conformational aspects of catalysis are reviewed in Ref. 51) formed via covalent linkage to the O82 atom of Asp-412 with 1.34-Å distance (Fig. 6A). The majority of interactions of this sugar are the same as previously observed for the −1 subsite sugar of the acarbose pseudotetrasaccharide (supplemental Fig. S4), but in addition, two water molecules (Fig. 6A) bind to O82 atom of Asp-480, the acid-base residue. These water molecules form a hydrogen bond network with Asp-480, Arg-463, Glu-417, and a third water molecule binding to Gln-542. Although transglycosylation is always kinetically favored over hydrolysis (52), what is unusual about CjAgd31B and indeed other transglycosylases is how they overcome the thermodynamically favored hydrolysis reaction in 55 m water as discussed below.

DISCUSSION

Through a combination of enzymological and structural analysis, we have revealed that CjAgd31B from the soil saprophyte C. japonicus possesses the ability to exclusively transfer single glucosyl units from α(1→4)-glucans to the non-reducing terminal 4-OH of glucose and α(1→4)- and α(1→6)-linked glucosyl residues; weak hydrolysis activity is only observed on the disaccharide maltose. As outlined in the Introduction and discussed below, this type of transglycosylase has not previously been described in GH31 nor any other CAZyme family to our knowledge.

GH31 enzymes utilize a double displacement mechanism involving a covalent glycosyl-enzyme intermediate, which, as was the case here for CjAgd31B, can be trapped and directly observed using kinetic probes derived from fluorosugars (46). In the natural reactions catalyzed by GH31 members, the glycosyl-enzyme is most commonly decomposed by water, yielding substrate hydrolysis. However, this intermediate can also be intercepted by saccharide acceptor substrates to generate transglycosylation products with varying efficiencies in a substrate- and enzyme-dependent manner (53, 54). Indeed, several members of GH31 have been shown to possess transglycosylation ability, although yields are typically low due to dominating hydrolytic reactions (21–23, 55, 56). In this context, the strict transglycosylating activity of CjAgd31B on malto-oligosaccharide substrates with degree of polymerization ≥3 is particularly noteworthy.

CjAgd31B is also distinct from homologues of bacterial ctsY and ctsZ gene products, which are the only other predominant transglycosylases to have been identified in GH31 thus far. Working in concert, CtsY and CtsZ generate cycloalternan tetrasaccharides from α(1→4)-glucans via a three-step reaction (19, 20). In the first step, CtsZ acts as an α(1→4)-to-α(1→6) transglucosylase to generate isomaltosyl moieties at the end of α(1→4)-glucan chains. Our HPAEC-PAD data indicate that in contrast CjAgd31B effects α(1→4)-to-α(1→4) transglycosylation and moreover cannot address (1→6) linkages; isomaltose is not a donor substrate. In the second and third steps of cycloalternan tetrasaccharide synthesis, CtsY catalyzes an intramolecular isomaltosyl transfer to yield a α-D-GlcP-(1→6)-α-D-GlcP-(1→3)-α-D-GlcP-(1→6)-α-D-GlcP-(1→4)-α-glucan structure followed by intramolecular cyclization to yield cyclo-(1→6)-α-D-GlcP-(1→3)-α-D-GlcP-(1→6)-α-D-GlcP-(1→3)-α-D-GlcP-(1→4). These two reactions are clearly distinct from that catalyzed by CjAgd31B. Unfortunately, the three-dimensional structures of both CtsZ and CtsY are currently unknown, which precludes comparison with CjAgd31B to understand the structural basis for these divergent transglycosylation activities in GH31.

The reactions catalyzed by CjAgd31B bear some similarity to, but are again distinct from, those catalyzed by 4-α-glucanotransferases (EC 4.2.1.25) of GH13 and GH77. GH13 encompasses a huge diversity of α-glucan-hydrolyzing and -transglycosylating enzymes (57) of which the Thermotoga spp. 4-α-glucanotransferases are perhaps the most relevant to the present study (12, 13, 58). Thermotoga maritima and Thermotoga neapolitana 4-α-glucanotransferases catalyze disproportionation reactions of malto-oligosaccharides, utilizing maltotetrose as the smallest donor substrate (12, 13). The structure of the T. maritima enzyme has revealed that the active site is an open cleft comprising at least 5 subsites (−2 to +3), which provides clear rationalization for the ability of the enzyme to randomly transfer longer α-glucan chains. In contrast, the T. maritima maltosyltransferase, also of GH13, strictly transfers maltosyl (Glc2) units from maltotriose and longer malto-oligosaccharides to the 4-position of α-glucan acceptor substrates (58) due to the presence of a unique protein motif that blocks the active site cleft (59).

GH77 enzymes are structurally related to GH13 enzymes in clan GH-H and thus generally possess open cleft-shaped active sites (60–62) that confer specificity for longer glucan donor substrates (9). For example, the eukaryotic Solanum tuberosum (potato) starch disproportionating (“D”) enzyme has such an extended active site (Protein Data Bank code 1x1n), transfers long α-glucan chains, and does not use maltose as a donor substrate (14, 63, 64). The disaccharide is also not a substrate for the Thermus aquaticus amylo maltase of GH77 that is distinguished by its propensity to form large cyclic α-glucans from long α(1→4)-glucan donors (14, 60). In the context of bacterial malto-oligosaccharide metabolism, the E. coli GH77 amylo maltase MalQ appears to favor the transfer of longer α-glucan chains, although there appears to be some debate whether this enzyme can utilize maltose as a donor, thereby transferring a single glucosyl residue to longer congeners.

CjAgd31B thus occupies a unique catalytic place as a 4-α-glucosyltransferase among the broader spectrum of 4-α-glucanotransferases. As a member of GH31, CjAgd31B belongs to clan GH-D, which is also composed of GH27 and GH36. Clan GH-D and GH-H members are built on a common triose isomerase (B/α)s barrel scaffold and may share a distant evolutionary relationship (27). However, in contrast to the clefited clan GH-H members, clan GH-D members are typified by shallow, pocket-shaped active sites comprising only one negative subsite accommodating a monosaccharide residue, the glyco-
sidic bond of which undergoes catalysis. CjAgd31B likewise presents an active site pocket as revealed by acarbose and 5-fluoro-glycosyl-enzyme complex structures, allowing speculation regarding its strict glucosyl transfer capacity.

The inherent challenge for transglycosylases is overcoming the thermodynamic preference for water as a nucleophile versus saccharide acceptor substrates. It has long been established by Withers et al. (52) through analysis of the reactivation of trapped intermediates that transglycosylation is kinetically favored over hydrolysis. Crystallographic analysis of the trapped covalent glycosyl-enzyme intermediate of CjAgd31B suggests that the hydrogen-bonding scheme does not place a water molecule with appropriate geometry or interaction with the catalytic acid-base residue Asp-480 to facilitate hydrolysis of the intermediate. Instead, two water molecules (one of them disordered) appear to lie on either side of the position expected of a catalytically competent nucleophile. However, the interglycosidic nitrogen of acarbose does interact with the acid-base (and with O3 and C5 groups, binding in positions corresponding to the observed waters of the trapped intermediate).

This solvent hydrogen-bonding arrangement suggests that the CjAgd31B active site has evolved to avoid deprotonation and activation of water, whereas optimization of hydrogen bonds to O6 and O3 of acceptor glucosides in the +1 subsite allows a favorable placement of the O4 atom for deprotonation and concerted electrophilic migration of C1 of the β-glucosyl enzyme. In this process, Glu-417 and Arg-463 play particularly important roles in binding the O3 of the +1 sugar and legislating against a water molecule positioned to enable hydrolysis. A caveat is that the trapped intermediate observed is that of a 5-fluorglycoside, so it is possible that the observed solvent network of the free enzyme structure is perturbed by the unnatural 5-fluoro substituent. However, the solvent network of the free enzyme structure is similar to that of the trapped intermediate complex especially in context of the O6- and O3-mimicking water molecules, where there is no suitably poised “nucleophilic” water molecule bound to Asp-480.

It is currently unclear what role(s) CjAgd31B might play in the biology of C. japonicus. Interestingly, the gene encoding CjAgd31B is located among a cluster of genes predicted to encode α-glucan-active enzymes (α-amylase, cyclomaltodextrin glucanotransferase, 6-phospho-β-glucosidase, and glucoamylase) and transporter proteins (TonB-dependent receptors and ATP-binding cassette transporters; supplemental Fig. S2). Whether these genes are co-regulated or comprise an operon is currently not known. This genomic association together with the observation that CjAgd31B is encoded with a native secretion signal peptide hints toward a role in glucogen or starch metabolism in the periplasm. Indeed, CjAgd31B could possibly have a function similar to the GH77 amylomalase MalQ from E. coli, which creates longer α-glucan chains from shorter malto-oligosaccharides as substrates for maltodextrin phosphorolyses; these phosphorolyses require maltopentaose as a minimal substrate to generate glucose 1-phosphate for further metabolism (5, 65, 66). C. japonicus, however, does possess a predicted GH77 homologue (CjMal77Q, CJA_1882; Ref. 25), which is located elsewhere in the genome and in proximity to other predicted glucogen/starch-active enzymes. This would suggest that CjAgd31B and CjMal77Q most likely have independent or perhaps complementary functions.

Another possible clue to the physiological function of CjAgd31B can be gleaned from analysis of potential GH31 orthologs. The biochemically characterized GH31 member closest to CjAgd31B is the α-glucosidase YihQ from E. coli K12 MG1655 (38). Biological data on E. coli YihQ are currently lacking; however, a reverse genetics analysis of a YihQ orthologue in Salmonella enterica serovar Enteriditis indicates that ΔyihQ mutants are deficient in capsular polysaccharide formation (67). Notably, the LPS of this organism consists of a repeating core glycan comprising tyvelose, l-rhamnose, galactose, and mannose that is appended with extended α(1→4)-glucan chains. It is therefore tempting to speculate that YihQ and by extension CjAgd31B may act as a transglycosylase to extend or restructure these chains. In this context, it is interesting to note that E. coli YihQ has previously been designated as an α-glucosidase based on a weak activity on α-glucosyl fluoride but no other α-glucosides (38). A reassessment of YihQ activity both in vitro and in vivo in light of the transglycosylation capacity of GH31 enzymes demonstrated in the present study may well be warranted. In conclusion, the detailed enzyme structure-function analysis of Agd31B from the model soil bacterium C. japonicus presented here that has defined a previously unknown α-transglucosylase activity in GH31 will inform future functional genomics studies in bacteria and other microorganisms.

Acknowledgments—We are grateful to Prof. Harry J. Gilbert (Newcastle University) for insightful discussions on C. japonicus and early access to genomic sequence data. We thank Prof. Stephen G. Withers (University of British Columbia) for gifts of α-glucosyl fluoride and 5-fluoro-α-D-glucopyranosyl fluoride and Dr. Gustav Sundqvist (KTH Glycoscience) for assistance with protein mass spectrometry. The staffs of the Diamond Light Source and European Synchrotron Radiation Facility are thanked for provision of data collection facilities.

REFERENCES
1. Elbein, A. D. (2009) in Microbial Glycobiology: Structures, Relevance and Applications (Moran, A., Holst, O., Brennan, P. J., and von Itzstein, M., eds) pp. 185–201, Academic Press, London
2. Wilson, W. A., Roach, P. J., Montero, M., Baroja-Fernández, E., Muñoz, F. J., Eydallin, G., Viale, A. M., and Pozueta-Romero, J. (2010) Regulation of glycogen metabolism in yeast and bacteria. FEMS Microbiol. Rev. 34, 952–985
3. Roach, P. J., Depaoli-Roach, A. A., Hurley, T. D., and Tagliabucci, V. S. (2012) Glycogen and its metabolism: some new developments and old themes. Biochem. J. 441, 763–787
4. Ball, S., Colleoni, C., Conci, U., Raj, J. N., and Tiriaux, C. (2011) The evolution of glycogen and starch metabolism in eukaryotes gives molecular clues to understand the establishment of plastid endosymbiosis. J. Exp. Bot. 62, 1775–1801
5. Boos, W., and Shuman, H. (1998) Maltose/maltodextrin system of Escherichia coli: transport, metabolism, and regulation. Microbiol. Mol. Biol. Rev. 62, 204–229
6. Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B. (2009) The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycomics. Nucleic Acids Res. 37, D233–D238
7. Palmer, T. N., Ryman, B. E., and Whelan, W. J. (1976) Action pattern of
A Bacterial α-Transglucosylase from GH31

towards improved functional annotations of α-amylase-related proteins. 

46. Lovering, A. L., Lee, S. S., Kim, Y. W., Withers, S. G., and Strynadka, N. C. (2005) Mechanistic and structural analysis of a family 31 α-glucosidase and its glycosyl enzyme intermediate. J. Biol. Chem. 280, 2105–2115

47. Hendlich, M., Rippmann, F., and Barnickel, G. (1997) LIGSITE: automatic and efficient detection of potential small molecule-binding sites in proteins. J. Mol. Graph. Model. 15, 359–363, 389

48. Sim, L., Quezada-Calvillo, R., Sterchi, E. E., Nichols, B. L., and Rose, D. R. (2008) Human intestinal maltase-glucoamylase: crystal structure of the N-terminal catalytic subunit and basis of inhibition and substrate specificity. J. Mol. Biol. 375, 782–792

49. Tan, K., Tesar, C., Wilton, R., Keigher, L., Babnigg, G., and Joachimiak, A. (2009) Novel α-glucosidase from human gut microbiome: substrate specificities and their switch. FASEB J. 24, 3939–3949

50. Ernst, H. A., Lo Leggio, L., Willemoes, M., Leonard, G., Blum, P., and Larsen, S. (2006) Structure of the Sulfolobus solfataricus α-glucosidase: implications for domain conservation and substrate recognition in GH31. J. Mol. Biol. 358, 1106–1124

51. Davies, G. J., Planas, A., and Rovira, C. (2012) Conformational analyses of the reaction coordinate of glycosidases. Acc. Chem. Res. 45, 308–316

52. Davies, G. J., Planas, A., and Rovira, C. (2012) Conformational analyses of the reaction coordinate of glycosidases. Acc. Chem. Res. 45, 308–316

53. Sinnott, M. L. (1990) Novel α-glucosidase from human gut microbiome: substrate specificities and their switch. FASEB J. 24, 3939–3949

54. Ernst, H. A., Lo Leggio, L., Willemoes, M., Leonard, G., Blum, P., and Larsen, S. (2006) Structure of the Sulfolobus solfataricus α-glucosidase: implications for domain conservation and substrate recognition in GH31. J. Mol. Biol. 358, 1106–1124

55. Kato, N., Suyama, S., Shirokane, M., Kato, M., Kobayashi, T., and Tsugakoshi, N. (2002) Novel α-glucosidase from Aspergillus nidulans with strong transglycosylation activity. Appl. Environ. Microbiol. 68, 1250–1256

56. Yamamoto, T., Unno, T., Watanabe, Y., Yamamoto, M., Okuyama, M., Mori, H., Chiba, S., and Kimura, A. (2004) Purification and characterization of Acremonium implicatum α-glucosidase having regioselectivity for α-1,3-glucosidic linkage. Biochim. Biophys. Acta 1700, 189–198

57. Stom, M. R., Danchin, E. G., Rancurel, C., Coutinho, P. M., and Henrissat, B. (2006) Dividing the large glycose hydrolase family 13 into subfamilies: towards improved functional annotations of α-amylase-related proteins. Protein, Eng. Des. Sel. 19, 555–562

58. Humphry, D. R., Black, G. W., and Cummings, S. P. (2003) Reclassification of ‘Pseudomonas fluorescens’ subsp. ‘cellulosum’ NCIMB 10462 (Usda et al. (1952) as Cellvibrio iaponicus sp. nov., and revival of Cellvibrio vulgaris sp. nov., nom. rev., and Cellvibrio fulvus sp. nov., nom. rev. Int. J. Syst. Evol. Microbiol. 53, 393–400

59. Meissner, H., and Liebl, W. (1998) Thermotoga maritima maltosyltransferase, a novel type of maltodextrin glycosyltransferase acting on starch and malto-oligosaccharides. Eur. J. Biochem. 258, 1050–1058

60. Przybas, I., Tomoo, K., Terada, Y., Takaha, T., Fujii, K., Saenger, W., and Strätler, N. (2000) Crystal structure of amylomaltase from Thermus aquaticus, a glycosyltransferase catalysing the production of large cyclic glucans. J. Mol. Biol. 296, 873–886

61. Jung, J. H., Jung, T. Y., Seo, D. H., Yoon, S. M., Choi, H. C., Park, B. C., Park, C. S., and Woo, E. J. (2011) Structural and functional analysis of substrate recognition by the 250s loop in amylomaltase from Thermus brockianus. Proteins 79, 633–644

62. Barends, T. R., Bultema, J. B., Kaper, T., van der Maarel, M. J., Dijkhuizen, L., and Dijkstra, B. W. (2007) Three-way stabilization of the covalent intermediate in amylomaltase, an α-amylase-like transglycosylase. J. Biol. Chem. 282, 17242–17249

63. Jones, G., and Whelan, W. J. (1969) Action pattern of D-enzyme, a transmaltodextrinylase from potato. Carbohydr. Res. 9, 483

64. Takaha, T., Yanase, M., Okada, S., and Smith, S. M. (1993) Disproportionating enzyme (4-α-glucanotransferase; EC: 2.4.1.25) of potato. Purification, molecular-cloning, and potential role in starch metabolism. J. Biol. Chem. 268, 1391–1396

65. Schwartz, M., and Hofnung, M. (1967) Maltodextrin phosphorylase of Escherichia coli. Eur. J. Biochem. 2, 132–145

66. Becker, S., Palm, D., and Schirz, R. (1994) Dissecting differential binding in the forward and reverse reaction of Escherichia coli maltodextrin phosphorylase using 2-deoxyglucosyl substrates. J. Biol. Chem. 269, 2485–2490

67. Gibson, D. L., White, A. P., Snyder, S. D., Martin, S., Heiss, C., Azadi, P., Surette, M., and Kay, W. W. (2006) Salmonella produces an O-antigen capsule regulated by AgfD and important for environmental persistence. J. Bacteriol. 188, 7722–7730