

Structural Elucidation of Dextran Degradation Mechanism by *Streptococcus mutans* Dextranase Belonging to Glycoside Hydrolase Family 66*

Received for publication, January 13, 2012, and in revised form, February 8, 2012. Published, JBC Papers in Press, February 15, 2012, DOI 10.1074/jbc.M112.342444

Nobuhiro Suzuki^{#†}, Young-Min Kim^{#§1,2}, Zui Fujimoto^{#3}, Mitsuru Momma[‡], Masayuki Okuyama[§], Haruhiko Mori[§], Kazumi Funane[¶], and Atsuo Kimura[¶]

From the [#]Biomolecular Research Unit, National Institute of Agrobiological Sciences, 2-1-2 Kannondai, Tsukuba 305-8602, Japan, the [§]Research Faculty of Agriculture, Hokkaido University, Kita-9 Nishi-9, Kita-ku, Sapporo 060-8589, Japan, and the [¶]Applied Microbiology Division, National Food Research Institute, National Agriculture and Food Research Organization, 2-1-12 Kannondai, Tsukuba, Ibaraki 305-8642, Japan

Background: Dextranase hydrolyzes α -1,6-linkages of dextran, producing isomaltooligosaccharides.

Results: Crystal structure of *Streptococcus mutans* dextranase belonging to the glycoside hydrolase family 66 was determined.

Conclusion: The enzyme structures complexed with isomaltotriose and suicide substrate revealed the enzyme's catalytically important residues.

Significance: This is the first structural report for a GH-66 enzyme elucidating the enzyme's catalytic machinery.

Dextranase is an enzyme that hydrolyzes dextran α -1,6 linkages. *Streptococcus mutans* dextranase belongs to glycoside hydrolase family 66, producing isomaltooligosaccharides of various sizes and consisting of at least five amino acid sequence regions. The crystal structure of the conserved fragment from Gln¹⁰⁰ to Ile⁷³² of *S. mutans* dextranase, devoid of its N- and C-terminal variable regions, was determined at 1.6 Å resolution and found to contain three structural domains. Domain N possessed an immunoglobulin-like β -sandwich fold; domain A contained the enzyme's catalytic module, comprising a (β/α)₈-barrel; and domain C formed a β -sandwich structure containing two Greek key motifs. Two ligand complex structures were also determined, and, in the enzyme-isomaltotriose complex structure, the bound isomaltooligosaccharide with four glucose moieties was observed in the catalytic glycine cleft and considered to be the transglycosylation product of the enzyme, indicating the presence of four subsites, -4 to -1, in the catalytic cleft. The complexed structure with 4',5'-epoxypentyl- α -D-glucopyranoside, a suicide substrate of the enzyme, revealed that the epoxide ring reacted to form a covalent bond with the Asp³⁸⁵ side chain. These structures collectively indicated that Asp³⁸⁵ was the catalytic nucleophile and that Glu⁴⁵³ was the acid/base of the double displacement mechanism, in which the enzyme showed a

retaining catalytic character. This is the first structural report for the enzyme belonging to glycoside hydrolase family 66, elucidating the enzyme's catalytic machinery.

Streptococcus mutans is a Gram-positive bacterium that has been implicated as a major cariogenic bacteria (1, 2), which metabolizes sugars, including sucrose, glucose, fructose, and lactose, to lactic acid. With *S. mutans* accumulation on tooth surfaces as dental plaque, the lactic acid concentration increases and lowers the oral cavity pH, which leads to demineralization of tooth enamel, the origin of dental caries. Extracellular glucans, synthesized from sucrose by glucosyltransferase enzymatic activities, are known to enhance *S. mutans* biofilm formation. Glucosyltransferases synthesize both water-soluble glucans, such as dextrans with linear α -1,6-linked glucose units, and water-insoluble glucans, which are sticky polysaccharides composed of glucose units, predominantly in α -1,3-linkage and with various degrees of branching and are associated with bacterial cell adhesion to the tooth surface.

Endodextranase (EC 3.2.1.11; 6- α -D-glucan-6-glucanohydrolase) from *S. mutans* (SmDex)⁵ is an enzyme that hydrolyzes α -1,6-linkages of dextran (1, 3, 4) and produces isomaltooligosaccharides (IGs) of various sizes; it has been shown to modify glucan structure by controlling the amount and content of extracellular glucans (5, 6). SmDex belongs to glycoside hydrolase family 66 (GH-66) according to the CAZy database (available on the World Wide Web) (7). GH-66 enzymes consist of

* This work was supported, in part, by the Program for Promotion of Basic and Applied Researches for Innovations in Bio-oriented Industry (BRAIN).

The atomic coordinates and structure factors (codes 3VMN, 3VMO, and 3VMP) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (<http://www.rcsb.org/>).

¹ Both authors contributed equally to this work.

² Supported by Japan Society for the Promotion of Science Postdoctoral Fellowship for Foreign Researchers Grant-in-aid 17-05467. Present address: Korea Research Institute of Bioscience and Biotechnology, Jeonbul Branch Institute, Bioindustry Research Center, 1404 Sinjeong-dong, Jeongeup-si, Jeonbul 580-185, Korea.

³ To whom correspondence may be addressed. Tel./Fax: 81-29-838-7877; E-mail: zui@affrc.go.jp.

⁴ To whom correspondence may be addressed. Tel./Fax: 81-11-706-2808; E-mail: kimura@abs.agr.hokudai.ac.jp.

⁵ The abbreviations used are: SmDex, *S. mutans* endodextranase; SmDexTM, N- and C-terminal truncated mutant of *S. mutans* endodextranase; CBM, carbohydrate-binding module; CITase, cycloisomaltooligosaccharide glucanotransferase; E5G, 4',5'-epoxypentyl- α -D-glucopyranoside; GH, glycoside hydrolase family; IG, isomaltooligosaccharide; IG-3, isomaltotriose; PsCITase, *Paenibacillus* sp. 598K CITase; PsDex, *Paenibacillus* sp. dextranase; r.m.s., root mean square; SeMet, selenomethionine; TVAI, *T. vulgaris* R-47 α -amylase 1; TVAII, *T. vulgaris* R-47 α -amylase 2; aa, amino acid(s); PDB, Protein Data Bank.

Crystal Structure of *S. mutans* GH-66 dextranase

dextranases and cycloisomaltooligosaccharide glucanotransferases (CITases; EC 2.4.1.248), enzymes found mainly in bacteria. Both enzymes utilize dextran as a substrate, but CITases synthesize cycloisomaltooligosaccharide (cyclic saccharides linked by α -1,6-glucosyl bonds) from dextran by intramolecular transglycosylation (8–10). The *dex* gene containing an open reading frame of 2553 bp encoding SmDex enzyme, was first cloned from *S. mutans* Ingbratt, and its nucleotide and amino acid (aa) sequences were determined (11). The expressed protein is composed of 850 aa residues with a molecular mass of 94.5 kDa, but the formation of proteinase-associated multiple isoforms has been reported; similar observations have also been also reported regarding many other GH-66 dextranases of native and recombinant forms (12). According to amino acid sequence analysis of GH-66 enzymes, SmDex has been divided into five regions: a signal peptide sequence (N-terminal 24 aa), an N-terminal variable region (Ser²⁵–Asn⁹⁹), a conserved region (Gln¹⁰⁰–Ala⁶¹⁵), a glucan binding site (Leu⁶¹⁶–Ile⁷³²), and a C-terminal variable region (Asn⁷³³–Asp⁸⁵⁰) (13–15). CITase has an extra long insertion of ~90 aa inside the dextranase conserved region (16).

Biochemical studies using site-directed mutagenesis based on amino acid sequence comparison with other glucosyltransferases have revealed that Asp³⁸⁵ is essential for the catalytic reaction (14). On the other hand, Asp²⁷⁰ of CITase from *Bacillus circulans* T3040 (17) and Asp²⁴³ of endodextranase from *Thermotoga lettingae* TMO (18), both corresponding to Asp³⁸⁵ in SmDex, have been implicated as catalytic residues. Here, we recently conducted mutational analyses on two GH-66 enzymes, CITase and dextranase from *Paenibacillus* sp., and described three amino acid residues essential for catalysis (19, 20). However, the detailed catalytic mechanism of this enzyme family has not been elucidated due to the lack of the determination of the GH-66 enzyme's three-dimensional structure. We have recently produced a truncation mutant of SmDex, which is devoid of its N- and C-terminal variable regions (SmDexTM; bearing Gln¹⁰⁰–Ile⁷³²) (21). The recombinant full-size SmDex protein is proteolytically degraded to more than seven polypeptides (23–70 kDa) during long storage. SmDexTM, however, does not accept any further proteinase degradation and shows enhanced substrate hydrolysis, suitable for further biochemical analysis and industrial applications. In this article, the crystal structure of SmDexTM is introduced, representing the first described crystal structure of a GH-66 family enzyme, as well as insights into other family structures in relation to their enzymatic activities. The binary complex structures were also determined in the presence of IGs and of a covalently bonded structure from a suicide substrate, which allowed elucidation of the enzymatic reaction mechanism together with the substrate recognition site.

EXPERIMENTAL PROCEDURES

Expression and Crystallization of SmDex—SmDexTM originated from *S. mutans* ATCC 25175 (American Type Culture Collection, Manassas, VA) and was expressed, purified, and crystallized as previously reported (21, 22). The SmDex gene lacking the N-terminal 99 and C-terminal 118 residues was cloned into pET28 (Novagen Inc., Madison, WI) and overexpressed in *Escherichia coli* BL21 (DE3) cells (Agilent Technol-

ogies Inc., Santa Clara, CA). The expressed protein was purified by Ni²⁺-charged Hi-TrapTM chelating HP column chromatography (GE Healthcare) and anion exchange chromatography using a Hi-TrapTM Q-SepharoseTM HP column (GE Healthcare). The resulting protein was concentrated to 3.6 mg ml⁻¹ and crystallized by the sitting drop vapor diffusion method using an equal amount of precipitant solution composed of 30% polyethylene glycol monomethyl ether 2000 (w/v) (Hampton Research, Aliso Viejo, CA), 0.1 M phosphate-citrate buffer at pH 4.2 (Hampton Research) and 293 K. Selenomethionine (SeMet)-labeled SmDexTM was expressed in LeMaster medium (23) using the methionine-auxotrophic strain *E. coli* B834(DE3) and crystallized under the same conditions as for the native enzyme.

Data Collection and Structure Determination—Diffraction experiments for the native and SeMet-substituted crystals were conducted at the protein microcrystallography beamline, BL-17A, of the Photon Factory, High Energy Accelerator Research Organization (Tsukuba, Japan). Native diffraction data to 1.60 Å resolution (space group *P*2₁) were collected using an ADSC-Q270 CCD detector (Area Detector Systems Corp., Poway, CA). Crystals were cryocooled in a nitrogen gas stream to 95 K. While structural analyses of the enzyme complexed with isomaltotriose (IG-3; Seikagaku Corp., Tokyo, Japan) or 4',5'-epoxypentyl- α -D-glucopyranoside (E5G) (24, 25), SmDexTM crystals were soaked into a drop containing 5% (w/v, 99 mM) IG-3 or 1 mM E5G in the precipitant solution for 10 min or 1 h, respectively, before the diffraction experiment. Diffraction data of the IG-3 complex to 1.90 Å resolution were collected at 100 K, using R-axis VII imaging plate area detectors, and CuK α radiation was generated by a rotating anode generator, MicroMax007 (Rigaku Corp., Akishima, Japan). The E5G complex data were collected at beamline BL-5A, Photon Factory. All data were integrated and scaled using the DENZO and SCALEPACK programs in the HKL2000 program suite (26). Crystal structure was determined by the multiwavelength anomalous dispersion method using SeMet-labeled crystals (27), and 12 selenium atom positions were determined, and initial phases were calculated using the SOLVE/RESOLVE program (28, 29). The solution was subjected to the automodeling ARP/wARP program (30) contained in the CCP4 program suite (31). Manual model building and molecular refinement were performed using the Coot (32) and REFMAC5 programs (33, 34).

For the analyses of ligand-binding structures, structural determination was conducted using the ligand-free structure as the starting model and the bound ligand observed in the electron density difference map. Data collection and refinement statistics are shown in Table 1. Stereochemistry of the models was analyzed with the Rampage program (35), and structural drawings were prepared using the PyMOL program (Schrodinger, LLC, New York).

RESULTS

Overall Structure of SmDexTM—The crystal structure of SmDexTM was determined by the multiwavelength anomalous dispersion method using SeMet derivative data, and successively, the native and two ligand complex structures with IG-3 (SmDexTM-IG) or E5G (SmDexTM-E5G) were determined.

Crystal Structure of *S. mutans* GH-66 dextranase

TABLE 1

Data collection and refinement statistics

	SmDexTM		SeMet SmDexTM			SmDexTM-IG	SmDexTM-E5G
PDB code	3VMN					3VMO	3VMP
Data collection	PF-BL17A		PF-BL17A			MicroMax	PF-BL5A
Wavelength (Å)	0.97000 (Se-peak)	0.97888 (Se-edge)	0.97931 (Se-edge)	0.98400 (low remote)	0.96400 (high remote)	1.54178	1.00000
Space group	<i>P</i> 2 ₁			<i>P</i> 2 ₁		<i>P</i> 2 ₁	<i>P</i> 2 ₁
Unit cell (Å, °)	<i>a</i> = 53.2, <i>b</i> = 89.7, <i>c</i> = 63.3, β = 102.3			<i>a</i> = 56.6, <i>b</i> = 90.1, <i>c</i> = 63.3, β = 99.7		<i>a</i> = 57.0, <i>b</i> = 90.3, <i>c</i> = 63.0, β = 100.3	<i>a</i> = 53.4, <i>b</i> = 90.0, <i>c</i> = 62.8, β = 102.3
Resolution (Å)	50.0 – 1.60 (1.66 – 1.60)	50.0 – 1.80 (1.86 – 1.80)	50.0 – 1.80 (1.86 – 1.80)	50.0 – 1.80 (1.86 – 1.80)	50.0 – 1.80 (1.86 – 1.80)	50.0 – 1.90 (1.97 – 1.90)	50.0 – 1.90 (1.97 – 1.90)
No. reflections	571,898	423,407	421,508	431,201	393,582	476,322	310,988
Unique reflections	76,683 (7,641)	57,046 (5,676)	57,460 (5,362)	57,508 (5,685)	56,018 (4,483)	49,514 (4,954)	45,454 (4,112)
<i>R</i> _{merge}	0.063 (0.175)	0.078 (0.195)	0.078 (0.289)	0.060 (0.196)	0.073 (0.361)	0.057 (0.408)	0.085 (0.267)
Completeness (%)	100.0 (100.0)	100.0 (99.9)	99.3 (93.6)	99.9 (99.5)	96.9 (78.3)	99.7 (100.0)	98.3 (89.7)
Redundancy	7.5 (7.5)	7.4 (7.3)	7.3 (5.9)	7.5 (6.9)	7.4 (6.1)	9.6 (9.4)	6.9 (5.6)
<i>I</i> / <i>σ</i> (<i>I</i>)	39.2 (11.5)	38.7 (11.3)	33.9 (4.7)	37.9 (8.6)	30.3 (2.7)	46.1 (8.4)	28.9 (5.4)
Refinement							
Resolution (Å)	61.8 – 1.60					29.2 – 1.90	61.4 – 1.90
<i>R</i> / <i>R</i> _{free}	0.151 / 0.182					0.187 / 0.237	0.196 / 0.238
No. of reflections	72,799					46,979	43,068
No. of atoms	5,619					5,333	5,408
No. of protein atoms	5,029					4,985	5,012
No. of ions	1 PO ₄ ²⁻					1 PO ₄ ²⁻	1 PO ₄ ²⁻
No. of ligand atoms	-					68	42
No. of water molecules	587					6 sugar residues	(1 E5G)
Average B factor (Å ²)	12.7					275	373
protein atoms	11.6					35.1	26.5
ions	12.9					34.7	26.0
ligands	-					68.9	33.4
water molecules	22.4					51.2	41.1
r.m.s. deviations						37.4	33.1
bond length (Å)	0.019					0.018	0.017
bond angles (°)	1.823					1.570	1.538
Ramachandran plot	98.4 / 1.6 / 0 (favored/allowed/ disallowed %)					96.8 / 3.0 / 0.2	97.8 / 2.2 / 0

Structural refinement statistics are summarized in Table 1, and the quality and accuracy of the final structures were further demonstrated as more than 98% of their residues fell within the common regions of the Ramachandran stereochemistry plot. The recombinant SmDexTM molecule was composed of a sin-

gle polypeptide chain of 643 aa (positions 98–740), where the N-terminal Met⁹⁸ and Asp⁹⁹ and the C-terminal ⁷³³LEHHH-HHH⁷⁴⁰ were derived from the expression vector and purification tag. The N-terminal four residues Met⁹⁸–Lys¹⁰¹ and the C-terminal five residues His⁷³⁶–His⁷⁴⁰ were not identified due

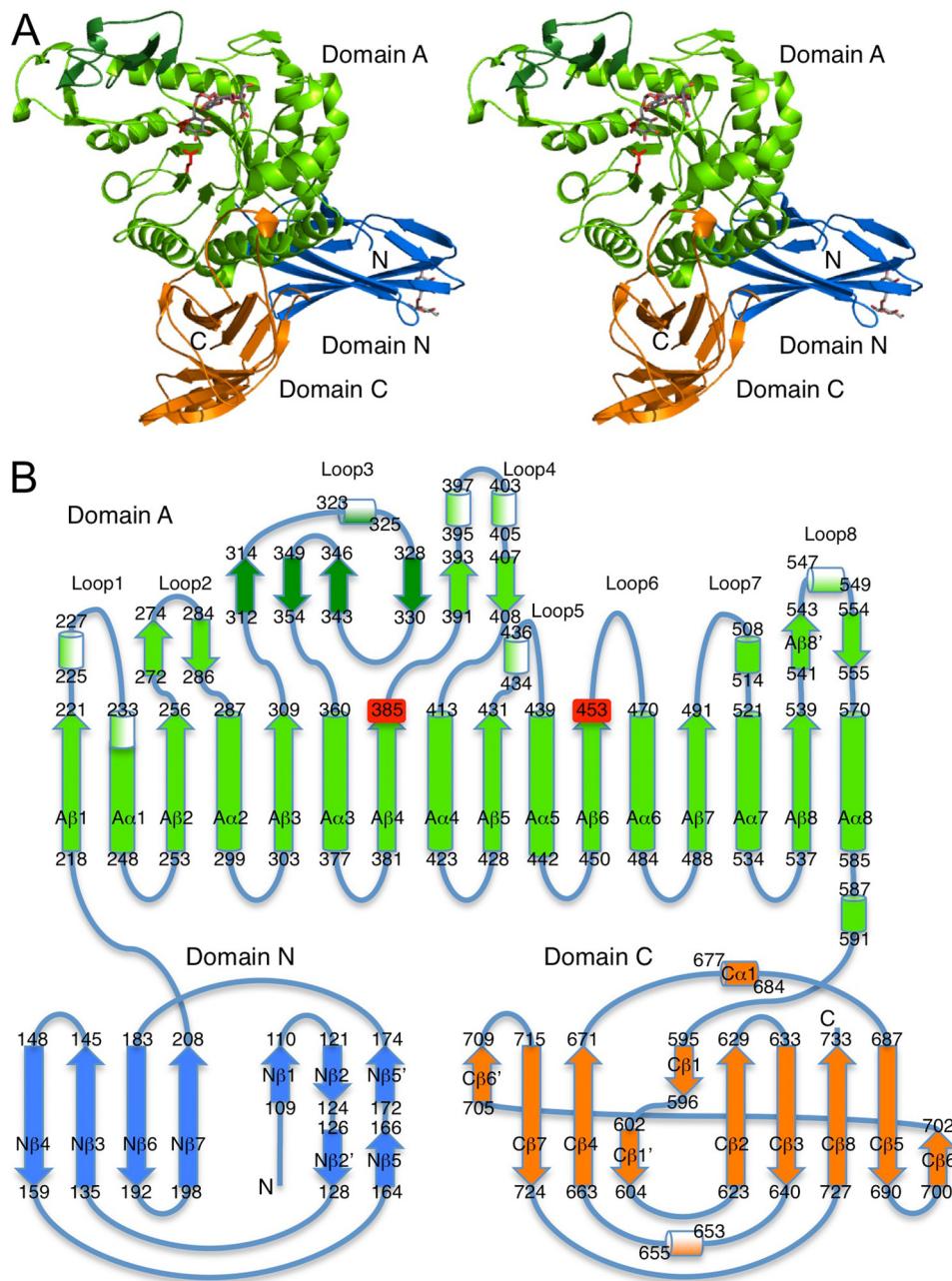


FIGURE 1. Structure of SmDexTM. *A*, stereoview of SmDexTM-IG complex ribbon model. Domains are shown in different colors: domains N, A, and C in blue, green, and orange, respectively. Dark green, loop 3 in domain A; red, two catalytic residues; gray, bound IG-3-derivative molecules. *B*, topological diagram of SmDexTM. α -Helices, 3_{10} -helices and β -strands are shown as filled cylinders, shaded cylinders, and filled arrows, respectively.

to lack of electron density. The final model consisted of one SmDexTM molecule accompanied with one phosphate ion.

The protein formed a multidomain structure composed of three domains, designated as domains N, A, and C (Fig. 1). Domain N (Asn¹⁰⁰–Asp²¹⁰) adopted an immunoglobulin-like β -sandwich fold, and its amino acid sequence showed no significant similarity with other domains in glycoside hydrolases, although a number of structures having similar immunoglobulin folds have been observed, including β -galactosidase (domains 2 and 4) (36), β -mannosidase (domains 2 and 4) (37), endoglucanase (PKD domain) (38), and bacterial sialidases (immunoglobulin-like linker domain) (39). The topology of domain N resembled a C2-type immunoglobulin fold, consist-

ing of seven antiparallel β -strands forming three-stranded and four-stranded β -sheets (Fig. 1B) (40, 41); a short loop split the 5th β -strand into two strands, N β 5 and N β 5', and, accordingly, the second strand was divided into N β 2 and N β 2'.

Domain A (Asp²¹¹–Lys⁵⁹³) was mainly composed of a $(\beta/\alpha)_8$ -barrel, which is a catalytic domain in many glycoside hydrolases. Two catalytic amino acid residues, which were detected by the SmDexTM-IG complex structure described below, were located on the concave surface formed by the central β -barrel C-terminal side. A structural homology search, using the Dali server (42), revealed that the domain A structure was similar to the catalytic domains of many GH-13 subfamilies, especially subfamily 20, such as in *Thermoactinomyces vul-*

Crystal Structure of *S. mutans* GH-66 dextranase

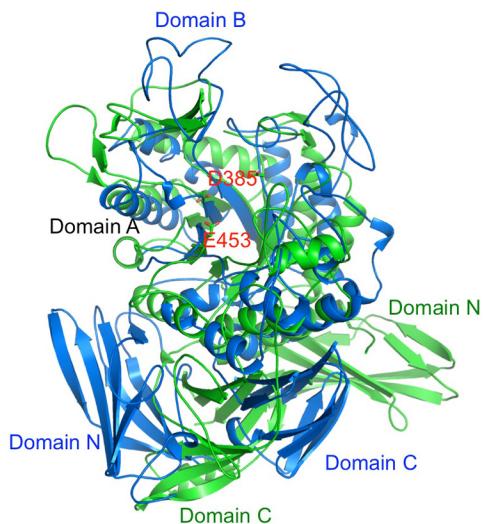


FIGURE 2. Superimposed models of SmDexTM (green) and TVAL (blue). Shown is the TVAL domain structure (43), similar to that of SmDex, from PDB entry 1J11; two catalytic residues of SmDexTM are indicated.

garis R-47 α -amylases 1 (TVAL, PDB code 2D0H, subfamily unknown) (43) and 2 (TVAII, PDB code 3A6O, subfamily 20) (44) and maltogenic amylase (PDB code 1GVI, subfamily 20) (45) with Z-scores of 18.2, 18.7, and 18.6, and root mean square (r.m.s.) differences of 3.6, 3.7, and 3.5 Å, respectively. There were five relatively large looped regions in this domain, ranging from 23 to 50 residues in length, and the 50-residue-long loop 3 associated with loop 2 to form a small subdomain structure (Fig. 1, dark green), which was similarly positioned in comparison with the domain B of α -amylases, forming one side wall of the catalytic cleft.

Domain C (Val⁵⁹⁴–Ile⁷³²) adopted an antiparallel β -sandwich structure, consisting of 10 β -strands, but basically belonged to two four-stranded Greek key motifs, which are found in many GHs. Sequence and deletion analyses of SmDex by Morisaki *et al.* (15) indicated that this domain as a dextran binding domain, but the deduced beginning of the domain, Leu⁶¹⁶, was different from Val⁵⁹⁴ in the actual domain structure observed here.

The domain arrangement of SmDex resembled some GH-13 proteins, including TVAL (43), which consists of domain N with an immunoglobulin fold; catalytic domain A with a $(\beta/\alpha)_8$ -barrel; domain B, which possesses a structure dissimilar from loop 3 of SmDex; and domain C with the Greek key motifs. The relative arrangement of domains N and C was, however, different between these enzymes, when the catalytic domains were superimposed (Fig. 2).

Crystal Structure of SmDexTM Complexed with Isomaltotriose–IG-3 was used for enzyme-product complex analysis as a means to elucidate the enzyme's catalytic mechanism. In the electron density map, bound sugars were observed at two positions, in the catalytic cleft and on domain N (Fig. 1). In the catalytic cleft, the modeled ligand was composed of four glucose moieties and appeared as a isomaltotetraose (Fig. 3, A and B), the product of a transglycosylation reaction by the enzyme. The other bound isomaltooligosaccharide showed two glucose moieties that appeared to be an isomaltose in the gap between domain N of an enzyme molecule and domain C of an adjacent

molecule in the crystal (Fig. 3C). The two domains contributed to the binding through a few hydrogen bonds and hydrophobic interactions, but no aromatic side chain was involved. Taking these facts into account, the latter ligand appeared to be a crystal packing artifact.

The overall structure of the SmDexTM-IG complex was almost identical to that of ligand-free SmDexTM with an r.m.s. difference of 0.73 Å, implying that ligand binding had little effect upon the overall structure. The main differences in the backbone structures were confined to the solvent-exposed loops in domains A and C, and four residues (Ile²⁴⁹ to Lys²⁵²) in loop 2 of domain A were invisible in the electron density map of the SmDexTM-IG complex.

The catalytic cleft was located at the center surface of the $(\beta/\alpha)_8$ -barrel. The enzyme subsites were named according to Davies *et al.* (46), and the transglycosylated ligand occupied subsites –4 to –1 from the non-reducing end to the reducing end, with the glucose moieties in subsites –4 to –1 designated as Glc–4 to Glc–1. All glucose moieties were in the relaxed chair conformation, with the glucose moiety Glc–1 in the β -anomeric conformation and the O1 atom in the proximity of the Asp³⁸⁵ side chain, which was the candidate of a catalytic nucleophile. However, dextran has α -1,6-linkages, and hydrogen bonds cannot be formed with the scissile bond of the natural substrate. The Asp³⁸⁵ Oδ2 atom was also close to the Glc–1 C1 atom with a distance of 3.1 Å, implying its role as a nucleophile. The other acidic residue, Glu⁴⁵³, formed a hydrogen bond with the Glc–1 O2 atom via the Oε2 atom. When Glc–1 adopts an α -anomeric conformation, as in natural substrates, the Glu⁴⁵³ Oε1 atom would be situated within hydrogen bonding distance of the Glc–1 O1 atom and could provide hydrogen in the catalytic bond, suggesting that it would act as an acid/base in the catalytic reaction. The O3 and O4 atoms of Glc–1 hydrogen-bonded to the Ala⁵⁵⁹ main chain oxygen atom and the Tyr²⁵⁷ Oη atom, respectively, and Glc–1 was recognized by five direct hydrogen bonds.

Glc–2 was also found to participate in five direct hydrogen bonds with three tyrosine residues, Tyr²⁵⁷, Tyr²⁶⁰, and Tyr³⁰⁷, and aspartic acid Asp²⁵⁸. Asp²⁵⁸ recognized two oxygen atoms, O3 and O4 of Glc–2, playing an important role in substrate recognition at subsite –2. Besides the hydrogen bonds, subsite –2 was built by the hydrophobic residues of Trp²⁷⁷, Trp²⁸⁰, and Met³⁰⁹, and Glc–2 was strictly recognized by the protein. In contrast, Glc–3 showed no direct hydrogen bonding to the protein, and Glc–4 exhibited one hydrogen bond between its O2 atom and the side chain of Thr⁵⁶³. These two glucose moieties were embedded into the two aromatic residues of Trp²⁸⁰ and Tyr⁵⁶⁰, appearing to be loosely recognized, with their electron density, especially for Glc–4, being relatively weak. The average *B*-factors of six cyclic atoms of the glucose rings were 51.6, 45.7, 56.7, and 64.8 Å² for Glc–1 to Glc–4, respectively, which also implied that subsites –1 and –2 showed strong glucose binding.

Crystal Structure of SmDexTM Complexed with ESG—The SmDexTM-E5G complex structure was determined to label and visually identify the enzyme's nucleophile residue. The resulting overall structure superimposed well with the ligand-free SmDexTM and SmDexTM-IG complex structures, with

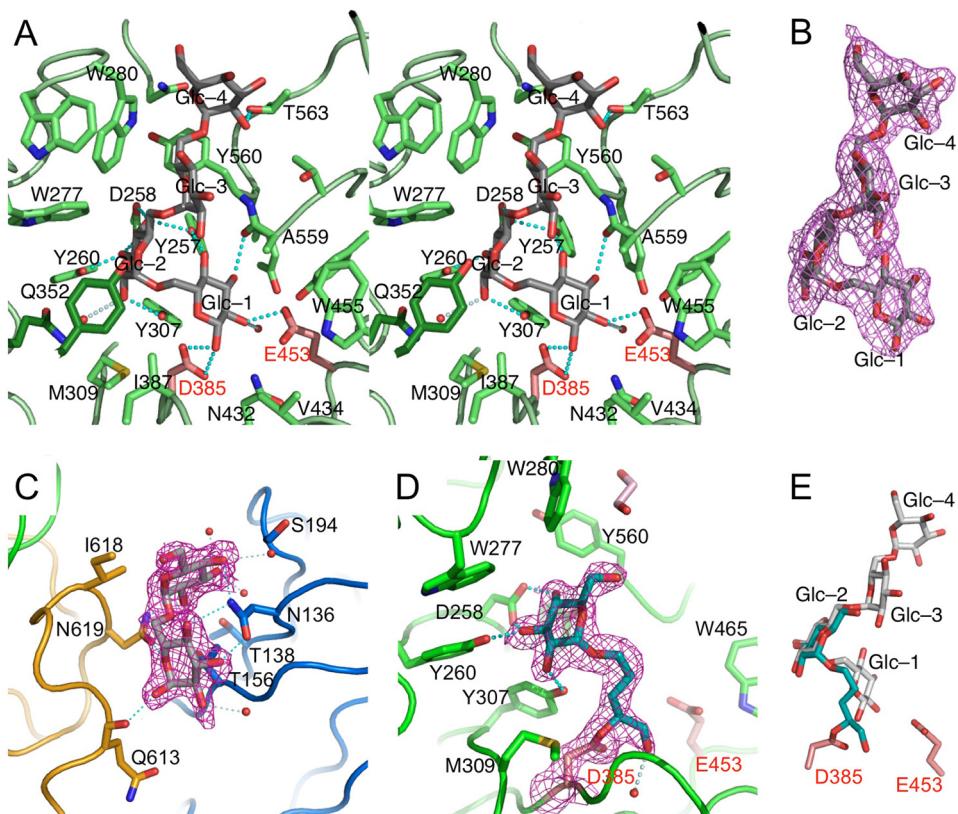


FIGURE 3. Bound ligand structures in SmDexTM. *A*, stereoview of SmDexTM-IG catalytic cleft. Gray stick model, bound IG; pale red, two catalytic residues; cyan break lines, estimated hydrogen bonds. *B*, $2F_o - F_c$ electron density map of the bound IG-3 derivative in the catalytic cleft (contour level, 1σ). *C*, a second IG-3 derivative found between domains N and C of the different molecules in the SmDexTM-IG complex structure; electron density map around the ligand (0.5σ contour level). *D*, E5G molecule covalently bonded to Asp³⁸⁵ in the SmDexTM-E5G catalytic cleft. Teal stick model, E5G; electron density map around the ligand (1σ contour level). *E*, superimposition of IG-3 derivative and E5G molecules bound in the catalytic cleft of the SmDexTM-IG and SmDexTM-E5G complex structures.

r.m.s. differences of 0.20 and 0.65 Å, respectively. The electron density at the active site accommodated a single E5G moiety, and the E5G density was connected to the Asp³⁸⁵ Oδ2 atom, revealing that E5G was covalently bound to the enzyme. The E5G glucose moiety was located at subsite -2, identical to Glc-2 of the SmDexTM-IG complex structure, and recognized in the same manner by Tyr²⁵⁷, Asp²⁵⁸, Tyr²⁶⁰, and Tyr³⁰⁷ (Fig. 3*D*). The E5G alkyl chain, consisting of the 1'-4' carbon atoms, occupied the similar positions of C6, C5, O5, and C1 atoms of the Glc-1 of the IG-3 complex, respectively, but was slightly shifted toward Asp³⁸⁵ (Fig. 3*E*). This was probably due to the covalent linkage between the E5G 4' carbon atom and the Asp³⁸⁵ Oδ2 atom after opening of the epoxide ring formed by the 4' and 5' carbons and epoxy oxygen.

DISCUSSION

In this study, the crystal structure of SmDexTM was determined, yielding the first example of a fully described GH-66 protein. The length of GH-66 proteins (including precursor forms) varies from 536 to 1686 aa residues. The 633-aa sequence used in this study covered the entire conserved region of GH-66 proteins, and thus, the SmDexTM crystal structure obtained could be considered a structural representative of this family. Although CITases have a carbohydrate-binding module 35 (CBM-35) family that splits the catalytic domain A into two regions in their primary structures (Fig. 4) (10, 16), the GH-66

protein minimal component comprises three domains, N, A, and C. In addition, the larger GH-66 proteins have C-terminal variable regions and are expected to contain different domain structures, as can be deduced from the Conserved Domain Database (47). These proteins include CBM-6 or CBM-35 in *Paenibacillus* sp. dextranase (48); CITase (10), CBM-4/9, or CBM-61 in *Paenibacillus* sp. dextranase; and CBM-2 or a fibronectin type III fold in *Catenulispora acidiphila* dextranase (49). The roles of these C-terminal domains remain unclear, but CBMs generally are involved in binding complicated substrates and in assisting catalysis in the catalytic domain, and they are assumed to be dextran-binding modules. In addition, some GH-66 enzymes possess regions in their C terminus for cell wall association and include sorting signals with LPXTG cell wall anchor motifs in SmDex (13) and S-layer homology domains, which associate noncovalently with cell walls (50).

SmDex and the other GH-66 dextranases and CITases are retaining enzymes, targeting substrates with α -anomeric conformations and releasing an α -anomeric product, and thus their catalytic mechanism is supposed to be a double displacement mechanism (51). There have been reports suggesting that Asp³⁸⁵ of SmDex (14) and the corresponding residues in other GH-66 enzymes are the nucleophiles (17, 18). Here, the SmDexTM-IG complex structure revealed that the Asp³⁸⁵ Oδ2 atom was located close to the Glc-1 C1 atom, with a distance of

Crystal Structure of *S. mutans* GH-66 dextranase

SmDex

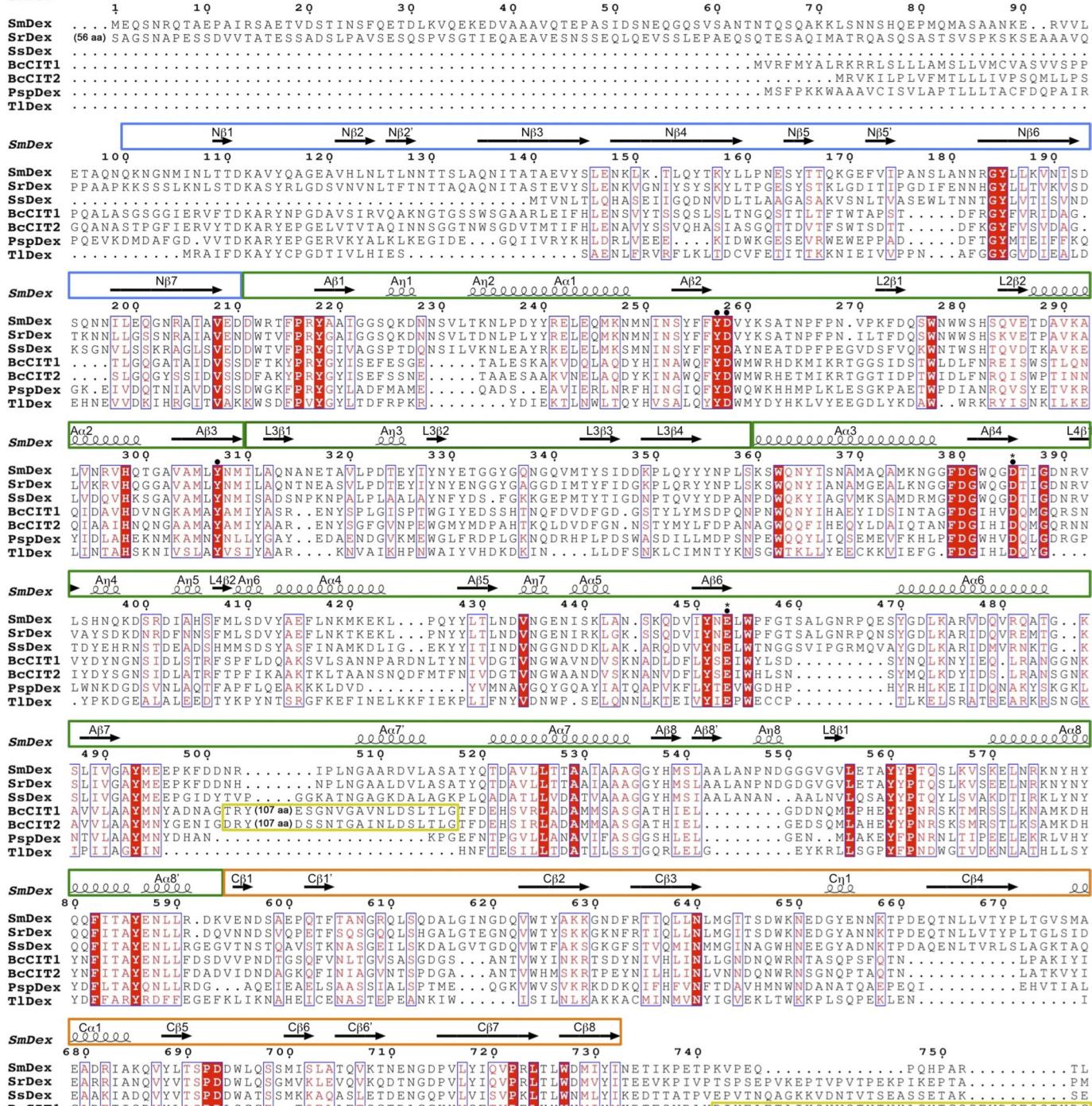


FIGURE 4. Amino acid sequence alignment of GH-66 proteins. Shown are the amino acid number, the secondary structural elements, and domain color for *SmDex*. Black closed circles, residues involved in hydrogen bonding with sugar molecules at subsites -1 and -2 ; asterisks, catalytic residues; boxes, conserved residues; yellow boxes, putative CBM-35 domains in CTases; black box, LPXTG motif. In order, proteins are dextranase from *S. mutans* ATCC 25175 dextranase (*SmDex*; GenBank™ accession number AEB70967), dextranase from *Streptococcus ratti* ATCC196645 (*SrDex*; AEJ54248), dextranase from *Streptococcus salivarius* M-33 (*SsDex*; BAA06127), CTase from *B. circulans* T-3040 (*BcCIT1*; BAA09604), CTase from *B. circulans* U-155 (*BcCIT2*; BAA13595), dextranase from *Paenibacillus* sp. Dex70-1B (*PspDex*; AAQ91301), and dextranase from *T. lettingae* TMO (*TlDex*; ABV33789). Alignment was prepared using the ESPript program (60).

3.1 Å. In addition, the SmDexTM structure when complexed with E5G, a widely used suicide substrate employed to label the nucleophile of the retaining glycosidases (52), showed that the epoxy ring opened up, and the 4' carbon atom of the alkyl chain moiety formed a covalent bond with the Asp³⁸⁵ Oδ2 atom. Structural evidence obtained here visually proved that Asp³⁸⁵ is the nucleophile of SmDex. On the other hand, there has been no report identifying the catalytic acid/base entity in GH-66 proteins. The SmDexTM-IG complex structure here showed that the C-1 hydroxyl group of Glc-1 took a β-anomeric conformation and that the distance between the Glu⁴⁵³ Oε1 atom and the Glc-1 C1 atom was 3.9 Å. However, the distance between the Glc-1 O1 and Glu⁴⁵³ Oε1 atoms would be within hydrogen-bonding distance, assuming the Glc-1 O1 atom was in the α-anomeric conformation. Furthermore, in the ligand-free SmDexTM structure, the Glu⁴⁵³ Oε1 and Asp³⁸⁵ Oδ2 atoms were apart by 6.0 Å, a reasonable distance between the two catalytic residues of a retaining type glycosidase (51). Taking these observations into account, Glu⁴⁵³ was identified as this enzyme's catalytic acid/base.

In addition to structural analyses, a separate mutational approach has been utilized to clarify the catalytic residues of GH-66 enzymes, using *Paenibacillus* sp. 598K CITase (PsCITase) (20) and *Paenibacillus* sp. dextranase (PsDex) (19). Three acidic residues, Asp¹⁴⁴, Asp²⁶⁹, and Glu³⁴¹, are predicted to be the catalytically important residues for PsCITase; similarly, Asp¹⁸⁹, Asp³⁴⁰, and Glu⁴¹² are important for PsDex. When a chemical rescue reaction is applied to D340G or E412Q mutants of PsDex by using α-isomaltotetraosyl fluoride with sodium azide NaN₃, the D340G and E412Q mutants formed β- and α-isomaltotetraosyl azides, implying that Asp³⁴⁰ and Glu⁴¹² are a nucleophile and an acid/base catalyst, respectively. PsCITase Asp²⁶⁹ and PsDex Asp³⁴⁰ correspond to Asp³⁸⁵ of SmDex, and PsCITase Glu³⁴¹ and PsDex Glu⁴¹² correspond to Glu⁴⁵³ of SmDex. Biochemical analyses and structural observations agree that aspartic acid is the catalytic nucleophile and that glutamic acid is the catalytic acid/base of the GH-66 enzymes.

Also from mutational analyses, another aspartic acid, Asp¹⁴⁴ of PsCITase and Asp¹⁸⁹ of PsDex, is essential for catalysis. They correspond to Asp²⁵⁸ of SmDex, which formed two hydrogen bonds to the two oxygen atoms of Glc-2. This indicated that accurate substrate binding at subsite -2 was essential for the enzymatic activity in addition to the catalytic residues directly involved in the hydrolysis.

The catalytic nucleophile of SmDex Asp³⁸⁵ was located at the end of strand Aβ4 and the acid/base catalyst Glu⁴⁵³ at the end of strand Aβ6. Many GHs have the (β/α)₈-barrel as a catalytic domain, but SmDex was similar to the clan GH-D enzymes in that the catalytic nucleophile was positioned at the end of the fourth β-strand and the catalytic acid/base on the loop adjacent to the end of the sixth β-strand of the (β/α)₈-barrel (Fig. 5A). The difference here was that GH-D enzymes possess two catalytic aspartate residues, whereas SmDex had glutamate Glu⁴⁵³ for the acid/base catalyst. The CAZy database lists three families, GH-27, GH-31, and GH-36, in clan GH-D. SmDex and GH-D proteins show similar domain arrangements, such that the three-dimensional structures of GH-27 enzymes are com-

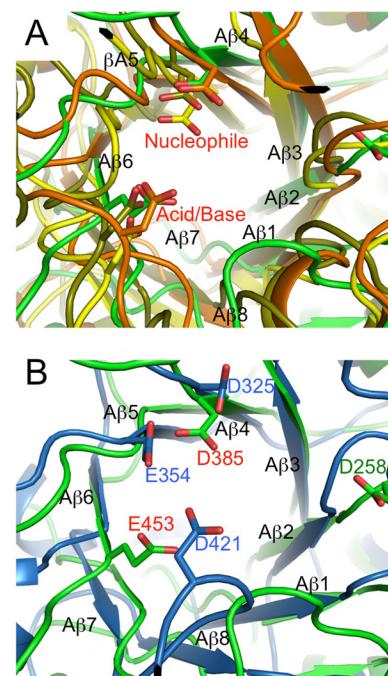


FIGURE 5. Superimposed model of the catalytic sites of SmDexTM on the related structures. A, comparison with GH-D proteins. Green, SmDexTM; orange, α-galactosidase from *Trichoderma reesei* (PDB code 1SZN) belonging to GH-27 (61); dark olive, α-xylosidase from *E. coli* (PDB code 1XSJ) belonging to GH-31 (62); yellow, α-galactosidase from *Lactobacillus acidophilus* NCFM (PDB code 2XN0) belonging to GH-36. Nucleophiles Asp³⁸⁵, Asp⁴¹⁶, and Asp⁴⁷⁹; general acids Asp²²⁶, Asp⁴⁸², and Asp⁵⁴⁹ of these proteins; and the corresponding residues of SmDex Asp³⁸⁵ and Glu⁴⁵³ are shown as stick models. B, comparison with GH-36 enzymes (dark blue; PDB code 1XSJ) (44), whose catalytic domain showed highest structural homology by DALI search; catalytically important residues of enzymes are shown as stick models.

posed of two domains corresponding to domains A and C in SmDex, with the exception of β-arabinopyranosidase from *Streptomyces avermitilis*, which contains extra two domains, domains III and IV (CBM-13) (53). Similarly, domains A and C are also conserved in GH-31 and GH-36 proteins, with GH-31 possessing two more domains (e.g. domain N and a distal C-terminal domain in α-xylosidases (PDB codes 1WE5 and 2XVG) (54, 55)), whereas α-galactosidases in GH-36 (PDB codes 2XN0 and 2YFO) (56, 57) have one extra domain similar to domain N in GH-31. Enzymes in GH-13 also have a (β/α)₈-barrel domain A and antiparallel β-domain C, but GH-13 is grouped in clan GH-H, and the positions of the catalytic residues are different from SmDex. The catalytic nucleophile is the aspartic acid located at the end of the (β/α)₈-barrel fourth β-strand, and the proton donor is the glutamic acid located at the fifth β-strand end (Fig. 5B) (44, 58).

Most enzymes in the clan GH-D are exo-type enzymes, working from oligosaccharide chain ends, and their active sites form a pocket-type structure, but, in contrast, an isomaltodextranase belonging to GH-27 is considered to be an endo-type exception, acting within oligosaccharide chains (59). SmDex is an endo-type enzyme, which possesses a catalytic cleft positioned across the surface of the (β/α)₈-fold. The cleft was surrounded by the loops of the (β/α)₈-barrel, such that loops 2, 3, and 4 formed one side wall, and loops 6 and 7 formed the other (Fig. 6). In the SmDexTM-IG structure, four glucose moieties were observed in the glycone side of the cleft, and the bound

Crystal Structure of *S. mutans* GH-66 dextranase

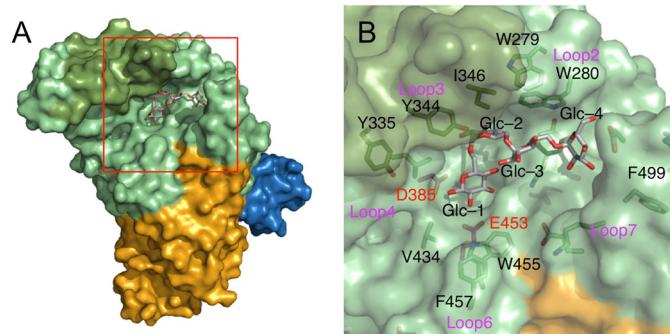


FIGURE 6. Surface model of the SmDexTM-IG complex structure. *A*, overall structure; domains N, A, and C are shown in blue, green, and orange, respectively. *B*, close-up view of catalytic cleft; bound IG-3 derivative, two catalytic residues, and hydrophobic residues forming the cleft are shown as stick models.

ligand appeared similar to isomaltotetraose, which was assumed to be the product of the reverse or transglycosylation enzyme reaction, because the protein used was a wild-type protein, and a relatively high IG-3 concentration was employed in the soaking experiment. Furthermore, continuous electron density was observed around the Glc-4 O6 atom, implying that the bound ligand might be isomaltooligosaccharide with a degree of polymerization of 5 or greater. Because Glc-4 was located at the catalytic cleft end, subsite -5 did not exist, and the remaining glucoses might have been disordered. Among the four subsites, Glc-2 showed the strongest electron density, compared with the other glucose moieties, and had the lowest *B*-factor. In addition, the SmDexTM-E5G complex structure demonstrated that the E5G glucose moiety occupied subsite -2. These observations suggested that it was necessary for the substrate to be captured by the enzyme at subsite -2 for a catalytic reaction to be possible. Amino acid sequence alignment with GH-66 proteins showed that five amino acids, which hydrogen-bond with Glc-1 and Glc-2 of the bound IG molecule (Tyr²⁵⁷, Asp²⁵⁸, Tyr³⁰⁷, Asp³⁸⁵, and Glu⁴⁵³), are strictly conserved in GH-66 enzymes, whereas Tyr²⁶⁰ is nevertheless conserved in streptococcal dextranases (Fig. 4). This implied that other GH-66 enzymes could be expected to recognize dextran in a manner almost identical to SmDex. In contrast, amino acid residues contributing hydrophobic interactions with Glc-3 and Glc-4 are less conserved and represent regions that might influence an enzyme's product specificity.

On the other hand, no ligand was observed in the aglycone side of the catalytic cleft. The glycone side was covered by protruding loops 2 and 7, whereas the aglycone side of the cleft was open and wide (Fig. 6). Dextran shows high solubility and appears to have a flexible structure, such that the wide cleft might preferably take in such an unstructured substrate. The aglycone side of the cleft mainly consisted of loops 3, 4, and 6 with some aromatic residues that were arranged at the loops and also appeared to help with substrate uptake. Subsites of the aglycone side remained unclear, but judging from the position of Glc-1, subsite +1 appeared to be surrounded by the hydrophobic residues Ile³⁸⁷, Val⁴³⁴, and Trp⁴⁵⁵, but Glc+1 did not seem to be specifically recognized.

Original SmDex contains a C-terminal variable region after domain C and a N-terminal variable region before domain N. In

a previous paper, recombinant full-size SmDex protein (95.4 kDa) was expressed but was also proteolytically degraded to form a shorter, truncated isoform of 89.8 kDa (21). When series of truncation mutants, with deleted C-terminal and/or N-terminal variable regions, were constructed and examined, SmDexTM was found to be devoid of its N- and C-terminal variable regions and to be proteinase-resistant. SmDexTM also displayed the enhanced substrate hydrolysis compared with full-size SmDex protein. Here, SmDexTM structure revealed that Asn¹⁰², the N-terminal residue, was located in the proximal region of the Aα8 head, rather close to the catalytic cleft's glycone side, and the C terminus His⁷³⁵ was positioned in the opposite side. Although the structure of these regions remains unclear, it cannot be denied that, as a proenzyme, the N-terminal variable region might approach the catalytic cleft's aglycone side, and the C-terminal variable region might approach the glycone side, and thus both might hinder pro-SmDex from accessing the substrate. Another possibility is that the unstructured N- and C-terminal variable regions might cause aggregation of the enzyme, leading to decreased hydrolytic activity.

Morisaki *et al.* (15) have reported that SmDex domain C is necessary for dextran binding in deletion mutant studies, but here, no apparent dextran-binding site was found in domain C of the SmDexTM-IG structure. This domain widely contacts the catalytic domain and appears to contribute to its stabilization. In particular, an elongated loop between the β-strands Cβ3 and Cβ4 was observed here to be in contact with the catalytic domain's loops 6 and 7, which were involved in the formation of the catalytic cleft, and appeared to hold their relative positions. Therefore, amino acid deletion in domain C might result in conformational changes in the catalytic cleft, causing reduction of dextran binding.

Acknowledgments—We thank the beamline researchers and staff at the Photon Factory.

REFERENCES

1. Hamada, S., and Slade, H. D. (1980) Biology, immunology, and cariogenicity of *Streptococcus mutans*. *Microbiol. Rev.* **44**, 331–384
2. Loesche, W. J. (1986) Role of *Streptococcus mutans* in human dental decay. *Microbiol. Rev.* **50**, 353–380
3. Igarashi, T., Yamamoto, A., and Goto, N. (1992) Characterization of the dextranase purified from *Streptococcus mutans* Ingbratt. *Microbiol. Immunol.* **36**, 969–976
4. Wanda, S. Y., and Curtiss, R., 3rd. (1994) Purification and characterization of *Streptococcus sobrinus* dextranase produced in recombinant *Escherichia coli* and sequence analysis of the dextranase gene. *J. Bacteriol.* **176**, 3839–3850
5. Walker, G. J., Pulkownik, A., and Morrey-Jones, J. G. (1981) Metabolism of the polysaccharides of human dental plaque. Release of dextranase in batch cultures of *Streptococcus mutans*. *J. Gen. Microbiol.* **127**, 201–208
6. Colby, S. M., Whiting, G. C., Tao, L., and Russell, R. R. (1995) Insertional inactivation of the *Streptococcus mutans* dexA (dextranase) gene results in altered adherence and dextran catabolism. *Microbiology* **141**, 2929–2936
7. 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 glycogenomics. *Nucleic Acids Res.* **37**, D233–D238
8. Oguma, T., Matsuyama, A., Kikuchi, M., and Nakano, E. (1993) Cloning and sequence analysis of the cyclomaltodextrinase gene from *Bacillus sphaericus* and expression in *Escherichia coli* cells. *Appl. Microbiol. Bio-*

Crystal Structure of *S. mutans* GH-66 dextranase

- technol.* **39**, 197–203
9. Oguma, T., Tobe, K., and Kobayashi, M. (1994) Purification and properties of a novel enzyme from *Bacillus* spp. T-3040, which catalyzes the conversion of dextran to cyclic isomaltooligosaccharides. *FEBS Lett.* **345**, 135–138
 10. Funane, K., Kawabata, Y., Suzuki, R., Kim, Y. M., Kang, H. K., Suzuki, N., Fujimoto, Z., Kimura, A., and Kobayashi, M. (2011) Deletion analysis of regions at the C-terminal part of cycloisomaltooligosaccharide glucanotransferase from *Bacillus circulans* T-3040. *Biochim. Biophys. Acta* **1814**, 428–434
 11. Igarashi, T., Yamamoto, A., and Goto, N. (1995) Characterization of the dextranase gene (*dex*) of *Streptococcus mutans* and its recombinant product in an *Escherichia coli* host. *Microbiol. Immunol.* **39**, 387–391
 12. Khalikova, E., Susi, P., and Korpela, T. (2005) Microbial dextran-hydrolyzing enzymes. Fundamentals and applications. *Microbiol. Mol. Biol. Rev.* **69**, 306–325
 13. Igarashi, T., Asaga, E., and Goto, N. (2004) Roles of *Streptococcus mutans* dextranase anchored to the cell wall by sortase. *Oral Microbiol. Immunol.* **19**, 102–105
 14. Igarashi, T., Morisaki, H., Yamamoto, A., and Goto, N. (2002) An essential amino acid residue for catalytic activity of the dextranase of *Streptococcus mutans*. *Oral Microbiol. Immunol.* **17**, 193–196
 15. Morisaki, H., Igarashi, T., Yamamoto, A., and Goto, N. (2002) Analysis of a dextran-binding domain of the dextranase of *Streptococcus mutans*. *Lett. Appl. Microbiol.* **35**, 223–237
 16. Aoki, H., and Sakano, Y. (1997) A classification of dextran-hydrolysing enzymes based on amino acid sequence similarities. *Biochem. J.* **323**, 859–861
 17. Yamamoto, T., Terasawa, K., Kim, Y. M., Kimura, A., Kitamura, Y., Kobayashi, M., and Funane, K. (2006) Identification of catalytic amino acids of cycloextran glucanotransferase from *Bacillus circulans* T-3040. *Biosci. Biotechnol. Biochem.* **70**, 1947–1953
 18. Kim, Y. M., and Kim, D. (2010) Characterization of novel thermostable dextranase from *Thermotoga lettingae* TMO. *Appl. Microbiol. Biotechnol.* **85**, 581–587
 19. Kim, Y. M., Kiso, Y., Muraki, T., Kan, M. S., Nakai, H., Saburi, W., Lang, W., Kang, H. K., Okuyama, M., Mori, H., Suzuki, R., Funane, K., Suzuki, N., Momma, M., Fujimoto, Z., Oguma, T., Kobayashi, M., Kim, D., and Kimura, A. (2012) Novel dextranase catalyzing cycloisomaltooligosaccharide formation and identification of catalytic amino acids and their functions using chemical rescue approach. *J. Biol. Chem.* **287**, 19927–19935
 20. Suzuki, R., Terasawa, K., Kimura, K., Fujimoto, Z., Momma, M., Kobayashi, M., Kimura, A., and Funane, K. (2012) Biochemical characterization of a novel cycloisomaltooligosaccharide glucanotransferase from *Paenibacillus* sp. 598K. *Biochim. Biophys. Acta*, in press
 21. Kim, Y. M., Shimizu, R., Nakai, H., Mori, H., Okuyama, M., Kang, M. S., Fujimoto, Z., Funane, K., Kim, D., and Kimura, A. (2011) Truncation of N- and C-terminal regions of *Streptococcus mutans* dextranases enhances catalytic activity. *Appl. Microbiol. Biotechnol.* **91**, 329–339
 22. Suzuki, N., Kim, Y. M., Fujimoto, Z., Momma, M., Kang, H. K., Funane, K., Okuyama, M., Mori, H., and Kimura, A. (2011) Crystallization and preliminary crystallographic analysis of dextranase from *Streptococcus mutans*. *Acta Crystallogr. F Struct. Biol. Cryst. Commun.* **67**, 1542–1544
 23. LeMaster, D. M., and Richards, F. M. (1985) ^1H – ^{15}N heteronuclear NMR studies of *Escherichia coli* thioredoxin in samples isotopically labeled by residue type. *Biochemistry* **24**, 7263–7268
 24. Kang, H. K., Kim, Y. M., Nakai, H., Kang, M. S., Hakamada, W., Okuyama, M., Mori, H., Nishio, T., and Kimura, A. (2010) Suicide substrate-based inactivation of endodextranase by ω -epoxyalkyl α -glucopyranosides. *J. Appl. Glycosci.* **57**, 269–272
 25. Kimura, A., Nishio, T., Hakamada, W., Oku, T., Mar, S. S., Okada, G., and Chiba, S. (2000) Affinity labelling of glycosidase by ω -epoxyalkyl α -glucoside. *J. Appl. Glycosci.* **47**, 235–241
 26. Otwinowski, Z., and Minor, W. (1997) Processing of x-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326
 27. Hendrickson, W. A., and Ogata, C. M. (1997) Phase determination from multiwavelength anomalous diffraction measurements. *Methods Enzymol.* **276**, 494–523
 28. Terwilliger, T. C. (2003) SOLVE and RESOLVE. Automated structure solution and density modification. *Methods Enzymol.* **374**, 22–37
 29. Terwilliger, T. C. (2003) Automated main-chain model building by template matching and iterative fragment extension. *Acta Crystallogr. D Biol. Crystallogr.* **59**, 38–44
 30. Perrakis, A., Morris, R., and Lamzin, V. (1999) Automated protein model building combined with iterative structure refinement. *Nat. Struct. Biol.* **6**, 458–463
 31. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., and Wilson, K. S. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 235–242
 32. Emsley, P., and Cowtan, K. (2004) Coot. Model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132
 33. Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F., and Vagin, A. A. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 355–367
 34. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* **53**, 240–255
 35. Lovell, S. C., Davis, I. W., Arendall, W. B., 3rd, de Bakker, P. I., Word, J. M., Prisant, M. G., Richardson, J. S., and Richardson, D. C. (2003) Structure validation by $\text{C}\alpha$ geometry. ϕ , ψ and $\text{C}\beta$ deviation. *Proteins* **50**, 437–450
 36. Jacobson, R. H., Zhang, X. J., DuBose, R. F., and Matthews, B. W. (1994) Three-dimensional structure of β -galactosidase from *E. coli*. *Nature* **369**, 761–766
 37. Tailford, L. E., Money, V. A., Smith, N. L., Dumon, C., Davies, G. J., and Gilbert, H. J. (2007) Mannose foraging by *Bacteroides thetaiotaomicron*. Structure and specificity of the β -mannosidase, BtMan2A. *J. Biol. Chem.* **282**, 11291–11299
 38. Najmudin, S., Guerreiro, C. I., Carvalho, A. L., Prates, J. A., Correia, M. A., Alves, V. D., Ferreira, L. M., Romão, M. J., Gilbert, H. J., Bolam, D. N., and Fontes, C. M. (2006) Xyloglucan is recognized by carbohydrate-binding modules that interact with β -glucan chains. *J. Biol. Chem.* **281**, 8815–8828
 39. Newstead, S., Watson, J. N., Knoll, T. L., Bennet, A. J., and Taylor, G. (2005) Structure and mechanism of action of an inverting mutant sialidase. *Biochemistry* **44**, 9117–9122
 40. Williams, A. F., and Barclay, A. N. (1988) The immunoglobulin superfamily. Domains for cell surface recognition. *Annu. Rev. Immunol.* **6**, 381–405
 41. Williams, A. F., Davis, S. J., He, Q., and Barclay, A. N. (1989) Structural diversity in domains of the immunoglobulin superfamily. *Cold Spring Harb. Symp. Quant. Biol.* **54**, 637–647
 42. Holm, L., Kääriäinen, S., Rosenström, P., and Schenkel, A. (2008) Searching protein structure databases with DaliLite version 3. *Bioinformatics* **24**, 2780–2781
 43. Kamitori, S., Abe, A., Ohtaki, A., Kaji, A., Tonoza, T., and Sakano, Y. (2002) Crystal structures and structural comparison of *Thermoactinomyces vulgaris* R-47 α -amylase 1 (TVAI) at 1.6 Å resolution and α -amylase 2 (TVAI) at 2.3 Å resolution. *J. Mol. Biol.* **318**, 443–453
 44. Ohtaki, A., Mizuno, M., Tonoza, T., Sakano, Y., and Kamitori, S. (2004) Complex structures of *Thermoactinomyces vulgaris* R-47 α -amylase 2 with acarbose and cyclodextrins demonstrate the multiple substrate recognition mechanism. *J. Biol. Chem.* **279**, 31033–31040
 45. Lee, H. S., Kim, M. S., Cho, H. S., Kim, J. I., Kim, T. J., Choi, J. H., Park, C., Lee, H. S., Oh, B. H., and Park, K. H. (2002) Cyclomaltodextrinase, neopullulanase, and maltogenic amylase are nearly indistinguishable from each other. *J. Biol. Chem.* **277**, 21891–21897
 46. Davies, G. J., Wilson, K. S., and Henrissat, B. (1997) Nomenclature for sugar-binding subsites in glycosyl hydrolases. *Biochem. J.* **321**, 557–559
 47. Marchler-Bauer, A., Anderson, J. B., Chitsaz, F., Derbyshire, M. K., DeWeese-Scott, C., Fong, J. H., Geer, L. Y., Geer, R. C., Gonzales, N. R., Gwadz, M., He, S., Hurwitz, D. I., Jackson, J. D., Ke, Z., Lanczycki, C. J., Liebert, C. A., Liu, C., Lu, F., Lu, S., Marchler, G. H., Mullokandov, M., Song, J. S., Tasneem, A., Thanki, N., Yamashita, R. A., Zhang, D., Zhang, N., and Bryant, S. H. (2009) CDD: Specific functional annotation with the Conserved Domain Database. *Nucleic Acids Res.* **37**, D205–D210

Crystal Structure of *S. mutans* GH-66 dextranase

48. Finnegan, P. M., Brumbley, S. M., O'Shea, M. G., Nevalainen, H., and Bergquist, P. L. (2005) Diverse dextranase genes from *Paenibacillus* species. *Arch. Microbiol.* **183**, 140–147
49. Copeland, A., Lapidus, A., Glavina Del Rio, T., Nolan, M., Lucas, S., Chen, F., Tice, H., Cheng, J. F., Bruce, D., Goodwin, L., Pitluck, S., Mikhailova, N., Pati, A., Ivanova, N., Mavromatis, K., Chen, A., Palaniappan, K., Chain, P., Land, M., Hauser, L., Chang, Y. J., Jeffries, C. D., Chertkov, O., Brettin, T., Detter, J. C., Han, C., Ali, Z., Tindall, B. J., Göker, M., Bristow, J., Eisen, J. A., Markowitz, V., Hugenholtz, P., Kyriades, N. C., and Klenk, H. P. (2009) Complete genome sequence of *Catenulispora acidiphila* type strain (ID 139908). *Stand. Genomic Sci.* **1**, 119–125
50. Mesnage, S., Fontaine, T., Mignot, T., Delepierre, M., Mock, M., and Fouet, A. (2000) Bacterial SLH domain proteins are non-covalently anchored to the cell surface via a conserved mechanism involving wall polysaccharide pyruvylation. *EMBO J.* **19**, 4473–4484
51. Davies, G., and Henrissat, B. (1995) Structures and mechanisms of glycosyl hydrolases. *Structure* **3**, 853–859
52. Withers, S. G., and Aebersold, R. (1995) Approaches to labeling and identification of active site residues in glycosidases. *Protein Sci.* **4**, 361–372
53. Ichinose, H., Fujimoto, Z., Honda, M., Harazono, K., Nishimoto, Y., Uzura, A., and Kaneko, S. (2009) A β -L-Arabinopyranosidase from *Streptomyces avermitilis* is a novel member of glycoside hydrolase family 27. *J. Biol. Chem.* **284**, 25097–25106
54. Larsbrink, J., Izumi, A., Ibatullin, F. M., Nakhai, A., Gilbert, H. J., Davies, G. J., and Brumer, H. (2011) Structural and enzymatic characterization of a glycoside hydrolase family 31 α -xylosidase from *Cellvibrio japonicus* involved in xyloglucan saccharification. *Biochem. J.* **436**, 567–580
55. Kitamura, M., Ose, T., Okuyama, M., Watanabe, H., Yao, M., Mori, H., Kimura, A., and Tanaka, I. (2005) Crystallization and preliminary x-ray analysis of α -xylosidase from *Escherichia coli*. *Acta Crystallogr. F Struct. Biol. Cryst. Commun.* **61**, 178–179
56. Fredslund, F., Hachem, M. A., Larsen, R. J., Sørensen, P. G., Coutinho, P. M., Lo Leggio, L., and Svensson, B. (2011) Crystal structure of α -galactosidase from *Lactobacillus acidophilus* NCFM. Insight into tetramer formation and substrate binding. *J. Mol. Biol.* **412**, 466–480
57. Bruel, L., Sulzenbacher, G., Cervera Tison, M., Pujo, A., Nicoletti, C., Perrier, J., Galinier, A., Ropartz, D., Fons, M., Pompeo, F., and Giardina, T. (2011) α -Galactosidase/sucrose kinase (AgaSK), a novel bifunctional enzyme from the human microbiome coupling galactosidase and kinase activities. *J. Biol. Chem.* **286**, 40814–40823
58. MacGregor, E. A., Janecek, S., and Svensson, B. (2001) Relationship of sequence and structure to specificity in the α -amylase family of enzymes. *Biochim. Biophys. Acta* **1546**, 1–20
59. Iwai, A., Ito, H., Mizuno, T., Mori, H., Matsui, H., Honma, M., Okada, G., and Chiba, S. (1994) Molecular cloning and expression of an isomaltodextranase gene from *Arthrobacter globiformis* T6. *J. Bacteriol.* **176**, 7730–7734
60. Gouet, P., Courcelle, E., Stuart, D. I., and Métoz, F. (1999) ESPript. Analysis of multiple sequence alignments in PostScript. *Bioinformatics* **15**, 305–308
61. Golubev, A. M., Nagem, R. A., Brandão Neto, J. R., Neustroev, K. N., Eneyskaya, E. V., Kulminskaya, A. A., Shabalina, K. A., Savel'ev, A. N., and Polikarpov, I. (2004) Crystal structure of α -galactosidase from *Trichoderma reesei* and its complex with galactose. Implications for catalytic mechanism. *J. Mol. Biol.* **339**, 413–422
62. 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 -glycosidase and its glycosyl-enzyme intermediate. *J. Biol. Chem.* **280**, 2105–2115