Trehalulose Synthase Native and Carbohydrate Complexed Structures Provide Insights into Sucrose Isomerization

Stéphanie Ravaud, Xavier Robert, Hildegard Watzlavick, Richard Haser, Ralf Mattes, and Nushin Aghajari

From the Laboratoire de BioCrystallographie, Institut de Biologie et Chimie des Protéines, CNRS et Université de Lyon, UMR 5086, IFR 128 “BioSciences Gerland-Lyon Sud”, F-69367 Lyon Cedex 07, France and the Universität Stuttgart, Institut für Industrielle Genetik, Almandring 31, D-70569 Stuttgart, Germany

Various diseases related to the overconsumption of sugar make a growing need for sugar substitutes. Because sucrose is an inexpensive and readily available D-glucose donor, the industrial potential for enzymatic synthesis of the sucrose isomers trehalulose and/or isomaltulose from sucrose is large. The product specificity of sucrose isomerases that catalyze this reaction depends essentially on the possibility for tautomerization of sucrose, which is required for trehalulose formation. For optimal use of the enzyme, targeting controlled synthesis of these functional isomers, it is necessary to minimize the side reactions. This requires an extensive analysis of substrate binding modes and of the specificity-determining sites in the structure. The 1.6–2.2-Å resolution three-dimensional structures of native and mutant complexes of a trehalulose synthase from Pseudomonas mesoacidophila MX-45 mimic successive states of the enzyme reaction. Combined with mutagenesis studies they give for the first time thorough insights into substrate recognition and processing and reaction specificities of these enzymes.

Among the important outcomes of this study is the revelation of an aromatic clamp defined by Phe256 and Phe280 playing an essential role in substrate recognition and in controlling the reaction specificity, which is further supported by mutagenesis studies. Furthermore, this study highlights essential residues for binding the glucosyl and fructosyl moieties. The introduction of subtle changes informed by comparative three-dimensional structural data observed within our study can lead to fundamental modifications in the mode of action of sucrose isomerases and hence provide a template for industrial catalysts.

In 2003, a report commissioned by four United Nations agencies, the World Health Organization, and the Food and Agriculture Organization, stated that sugar should not account for more than 10% of a healthy diet aimed at reducing the growing burden of disorders related to cardiovascular pathologies, several forms of cancer, diabetes, obesity, osteoporosis, and dental infectious diseases. Thus, a number of synthetic sugar substitutes have been introduced on the market and still novel compounds emerge. They vary considerably in their degree of sweetness, volume, texture, and stability under various conditions, and no sweetener is a perfect replacement for sucrose, the main limitation being their taste profiles. Moreover, they are among the most controversial food additives because of allegations of adverse health effects including dermatological problems, headaches, mood variations, respiratory difficulties, seizures, and cancer.

Trehalulose (α-D-glucosylpyranosyl-1,1-D-fructofuranose) and isomaltulose (α-D-glucosylpyranosyl-1,6-D-fructofuranose) (Fig. 1, A and B) are structural isomers of sucrose (α-D-glucosylpyranosyl-1,2-β-D-fructofuranoside) that have a sweet taste and very similar physical and organoleptic properties to sucrose (Fig. 1C) (1, 2). Moreover, they are acaricidal (3), and their absorption reduces the rate with which monosaccharides and insulin are released into the bloodstream (4) and can thus be applied in diabetic and sports foods and drinks. Considering also that no side effects have been yet reported, these compounds could be ideal sucrose substitutes. Furthermore, the reducing properties of isomaltulose make it attractive as industrial precursor for the manufacturing of biosurfactants and biocompatible polymers (5).

Naturally present in honey in very small quantities, their chemical synthesis is very difficult, and their industrial production proceeds exclusively from sucrose using immobilized microorganisms. Several bacterial strains have been reported to convert sucrose into both isomaltulose and trehalulose, to act as a reserve material during periods of low carbon availability, because of the activity of a particular sucrose isomerase (sucrose mutase or α-glucosyltransferase, EC 5.4.99.11) (6–15). Sucrose isomerases are monomeric enzymes widely identified also in whiteflies (16), which catalyze the isomerization of sucrose into trehalulose and isomaltulose as main products and produce glucose and fructose in residual amounts because of sucrose hydrolysis. The ratio of the enzyme products varies from mainly isomaltulose (75–85%) to predominantly trehalulose (90%) depending on the bacterial strain (6–15).

Sucrose isomerases are classified into glycoside hydrolase family 13 of retaining enzymes (17) (www.cazy.org), the largest of the 110 families existing to date. It contains a wide range of polysaccharide-metabolizing enzymes, including glycogen and starch branching enzymes. All family 13 enzymes have structurally similar active sites with highly conserved catalytic residues and proceed via a double-displacement mechanism with...
for the formation of a covalent glycosyl-enzyme intermediate (18–23). What is still not understood is how these very divergent enzymes use the rest of their surface to confer both activity and specificity. For sucrose isomerases, the three-dimensional native structure of the isomaltulose synthase Pall from Klebsiella sp. LX3 (14, 24, 25) is the only structural information currently available, and a complex of this enzyme with a substrate, inhibitor, or substrate analogue is lacking. These authors suggested that a unique RLDRD motif located in a loop region is involved in the isomerization reaction and in the product specificity. To go beyond these first insights, more structural data especially on enzyme-substrate complexes are needed to characterize the enzyme specificity and the molecular mechanism controlling sucrose isomerization and consequently to deal with optimization of the catalytic properties of these enzymes to improve their industrial use.

We have recently cloned, purified, and crystallized the trehalulose synthase MutB from Pseudomonas mesoacidophila MX-45, a sucrose isomerase that mainly produces trehalulose (26), and described preliminary data on the structure of MutB in complex with Tris (27). Here, we investigate the catalytic mechanism and the specificities of MutB by combining mutagenesis and structural approaches and report the first three-dimensional structure of a trehalulose synthase and the first enzyme-substrate complex structures for a sucrose isomerase.

**EXPERIMENTAL PROCEDURES**

**Construction, Expression, and Purification of Enzymes**—MutB was overexpressed in Escherichia coli JM109 cells harboring the plasmid pHWG315 and purified by fast protein liquid chromatography (GE Healthcare) as described elsewhere (26). The active site MutB mutants D200A and E254Q were obtained by site-directed mutagenesis according to the QuikChange® site-directed mutagenesis protocol (Stratagene). The mutations were generated by using two synthetic oligonucleotide primers designed by the QuikChange® primer design program (Stratagene) (D200A, 5'-AGGCATGGGTCGCCGGACGGTGGCTACCT-3' and 5'-AGGTAGCCACCCTGCGGAAACCGCATGCCCT-3'; E254Q, 5'-GTCACGGGCCGACATCTTTCGCGGCTCC-3' and 5'-GGAGCGCCGAAAGTCTTGCGCCGGTGAC-3'; the mutations are indicated by bold letters, and silent mutations introducing a recognition site for the restriction enzyme BglII are indicated in italics), and confirmed by DNA sequencing. The cells of E. coli JM109 were transformed with mutant plasmid, i.e. pHWG678.6 (D200A) and pHWG679.9 (E254Q), and mutant enzymes were expressed and purified as described for the native protein (26).

Site-directed saturation mutagenesis followed the guidelines of the QuikChange® protocol using single degenerate codon oligonucleotides (NNS) to produce a mutant clone collection for residues 256, 280, and 414. All of the plasmids coding for mutB gene variants are derivatives of the l-rhamnose-inducible expression vector pJOE2702. Randomly selected clones were expressed, as described for the native enzyme, DNA-sequenced, and analyzed for enzyme properties.

**Enzyme Assay and Product Analysis**—Sucrose isomerase activity was determined as described previously (26). Enzyme mutants were analyzed by their ability to act on sucrose and the amount of released products. The product composition following 100 mM sucrose conversion was measured after ~90% substrate consumption, and concentrations of the individual sugars were determined by high performance anion exchange chromatography (HPLC)³ with pulse amperometric detection. The samples were separated on a CarboPac® PA1 column (Dionex) with 0.22 M NaOH, 0.02 M NaAc at a flow rate of 0.75 ml/min. Eluted sugars detected by pulsed amperometry with an analytical cell (5040, ESA Coulochem II, Bischoff) were identified by comparing retention times with those of standards. The relative percentage of each sugar was calculated from peak areas. Standard sugars, isomaltulose and trehalulose, were obtained from Südzucker. Protein concentration was determined by the method of Bradford (28) with bovine serum albumin as a standard.

**Inhibition Studies**—Inhibition studies were performed by measuring the enzyme activity at various concentrations of each inhibitor using sucrose as substrate. The amount of formed trehalulose was quantified by HPLC as described above and used for activity calculations. Lineweaver-Burk plots were performed for determination of the inhibition mode, and Dixon plots were used to calculate Kᵢ for the inhibitors.

**Crystallization**—Crystals were grown at 17 °C as described previously (Ref. 26 and Table 1).

MutB complexes with deoxyxojirimycin and castanospermine, respectively, were obtained by co-crystallization. 0.2 μl of a 100 mM inhibitor stock solution (Sigma) was mixed with 0.8 μl of protein stock and 1 μl of well solution, thus resulting in a final concentration of 10 mM inhibitor in the drops.

Co-crystallization of D200A in complex with glucose was performed by mixing 0.2 μl of a 200 mM sucrose solution (Sigma) with 0.8 μl of mutant enzyme (4.0 mg/ml) and 1 μl of well solution.

³ The abbreviation used is: HPLC, high performance liquid chromatography.
TABLE 1
Crystallization, data collection, and refinement statistics for MutB structures

| Protein conditions | Native MutB | MutB-Tris<sup>a</sup> | MutB-deoxyribozyme | MutB-castanospermine | MutB D200A-glucose | MutB E254Q-sucrose |
|--------------------|-------------|-----------------------|---------------------|----------------------|---------------------|---------------------|
| Protein concentration (mg/ml) | 8 | 7 | 7 | 8 | 4 | 4.6 |
| Ligand concentration | | 10 mm (co-crystallization) | | | | |
| Protein Data Bank entry code | 2PWH | IZJA | 2PWD | 2PWG | 2PWF | 2PWE |
| Synchrotron beamline (European Synchrotron Radiation Facility, Grenoble, France) | FIP BM30A | ID14-4 | FIP BM30A | ID14-2 | FIP BM30A | |
| Detector | MarCCD | ADSC Quantum 4 CCD | ADSC Quantum 4 CCD | MarCCD | ADSC Quantum 4 CCD | MarCCD |
| Wavelength (Å) | 0.9803 | 1.0720 | 0.9330 | 0.9803 | 0.9330 | 0.9797 |
| Data collection temperature (K) | 100 | 100 | 100 | 100 | 100 | 100 |
| Space group | P1 | P1 | P1 | P2<sup>2</sup> | P1 | P1 |
| Unit cell parameters (Å, °) | | | | | | |
| a = 63.6, b = 71.5, c = 81.7, α = 67.7, β = 73.0, γ = 70.4 | a = 63.4, b = 72.0, c = 82.2, α = 67.2, β = 73.1, γ = 70.8 | a = 63.4, b = 72.0, c = 82.0, α = 67.2, β = 73.4, γ = 70.9 | a = 64.2, b = 86.0, c = 123.5, β = 99.2 | a = 64.0, b = 85.8, c = 122.1, α = 81.8, β = 81.4, γ = 89.9 | a = 64.3, b = 73.7, c = 82.7, α = 66.2, β = 74.5, γ = 72.5 |
| Resolution range (Å) | 19.6–2.0 | 39.0–1.6 | 42.3–1.8 | 19.9–2.2 | 45.8–1.8 | 20.0–2.0 |
| Molecules in the AU | 2 | 2 | 2 | 2 | 4 | 2 |
| Outermost shell (Å) | 2.1–2.0 | 1.7–1.6 | 1.9–1.8 | 2.34–2.2 | 1.91–1.8 | 2.13–2.0 |
| Completeness (%) | 96.7 (95.9) | 95.0 (95.0) | 95.7 (93.7) | 99.8 (99.9) | 95.3 (94.4) | 96.4 (95.8) |
| Multiplicity | 2.0 | 1.9 | 2.1 | 3.8 | 1.9 | 2.0 |
| Total number of reflections | 160,045 | 294,913 | 233,837 | 255,546 | 439,132 | 170,239 |
| Number of unique reflections | 80,624 | 147,205 | 110,212 | 125,348 | 225,077 | 85,358 |
| R<sub>sym</sub> (%) | 7.8 (29.6) | 8.6 (27.7) | 8.6 (27.0) | 10.3 (35.3) | 6.3 (27.9) | 5.3 (19.3) |
| R<sub>merge</sub> (%) | 8.9 (2.8) | 4.7 (2.2) | 9.7 (3.9) | 10.2 (3.9) | 8.9 (3.1) | 12.4 (4.3) |
| R<sub>free</sub> (%) | 17.2 | 18.4 | 16.9 | 18.6 | 19.8 | 19.0 |

<sup>a</sup> A preliminary report on the MutB-Tris complex has been published (27).
<sup>b</sup> R<sub>sym</sub> = (Σ<sub>i</sub> |<sup>Σ</sup><sub>j</sub> |F<sub>obs</sub>(h) − kF<sub>crys</sub>(h)|) / (Σ<sub>i</sub> |F<sub>crys</sub>(h)|).
<sup>c</sup> R<sub>merge</sub> = (Σ<sub>i</sub> |<sup>Σ</sup><sub>j</sub> |F<sub>obs</sub>(h) − kF<sub>crys</sub>(h)|) / (Σ<sub>i</sub> |F<sub>crys</sub>(h)|).
<sup>d</sup> R<sub>free</sub> = (Σ<sub>i</sub> |<sup>Σ</sup><sub>j</sub> |F<sub>obs</sub>(h) − kF<sub>crys</sub>(h)|) / (Σ<sub>i</sub> |F<sub>crys</sub>(h)|).
R<sub>free</sub> is calculated from a test set constituted by 5, 8, or 10% of the total number of reflections randomly selected (see "Experimental Procedures"), and k is a scaling factor.
The complex between the inactive E254Q mutant and sucrose was obtained by soaking mutant crystals for 20 min in the crystallization solution containing 1 mM sucrose.

Data Collection—Crystals used for the diffraction experiments typically measured \(0.8 \times 0.2 \times 0.03\) mm\(^3\) and were cryo-protected prior to data collection by rapid soaking (1 min) in mother liquor containing 20% (v/v) glycerol and for MutB-deoxynojirimycin, MutB-castanospermine, D200A-glucose, and E254Q-sucrose, the corresponding ligand molecule at the concentration used for co-crystallization or soaking experiments.

The diffraction data were collected at various beamlines (Table 1) at the European Synchrotron Radiation Facility (Grenoble, France). Diffracted intensities were integrated and scaled using programs from the XDS package (29), except for MutB-Tris, which was integrated with the program MOSFLM (30) and scaled with SCALA from the CCP4 software package (31). The data collection statistics are shown in Table 1.

Structure Determination and Refinement—The 1.6-Å resolution MutB-Tris structure (27) was used as starting model in a difference Fourier (water, Tris, and calcium ions were omitted) to solve the structures of native MutB and MutB-deoxynojirimycin. An initial rigid body refinement using data to 3 Å was performed, and in the subsequent refinement steps a simulated annealing protocol was used extending data to 2 and 1.8 Å, respectively.

The structure of MutB-castanospermine was solved by the molecular replacement method using the refined P2\(_1\) structure of MutB-Tris at 1.8 Å resolution (Protein Data Bank entry code 1ZJB) (27) as a search model, employing the program AMoRe (32). Diffraction data in the resolution range 15–3.5 Å were used in the molecular replacement search. A unique solution located the two molecules in the asymmetric unit with a correlation coefficient of 81.4% and an \(R\) factor of 31.3%. For D200A-glucose, the MutB-Tris structure (27) was used as a search model in AMoRe, and a unique solution was found, using 15–3.5-Å data. Correlation coefficients and \(R\) factors were, respectively, 75.0% and 33.1%.

The MutB-deoxynojirimycin structure was used for initial phasing of E254Q-sucrose data in a difference Fourier followed by a rigid body refinement, as described above.

All of the models were refined using the simulated annealing protocol as implemented in the software CNS (33), alternately with visual examination of electron density maps and manual building using TURBO-FRODO (34). \(R\) and \(R_{free}\) factors (35) were monitored, and the latter was calculated from 10% of the reflections randomly selected from the MutB-Tris and MutB-castanospermine data: 8% for native MutB data and 5% in the case of MutB-deoxynojirimycin, D200A-glucose, and E254Q-sucrose data.

Water molecules located at similar positions in the respective structures have the same numbering. Ligand models found in the HIC-Up data base (36) were manually inserted in the electron density maps. The quality of the three-dimensional structures was examined with PROCHECK (37) and WHATCHECK (38).

Active sites in the enzyme complex structures were inspected, and hydrogen bonds with a maximum length of 3.35 Å showing suitable hydrogen bond geometries were monitored with the program HBPLUS (39). The definition of the catalytic pocket and the calculation of its volume were done using the VOIDOO software (40).

Alignment and Figure Rendering—MutB and Pall sequences were aligned using the program CLUSTALW (41), and rendering of the alignment including superimposition of secondary structures and domains of the two enzymes was done with the program ESPript (42). The drawings were generated with the programs GRASP (43), DS Visualizer (Accelrys Inc., San Diego), ENDscript (42), and PyMOL (44).

RESULTS

Inhibition Studies—The inhibitory effects of the known \(\alpha\)-glucosidase inhibitors, castanospermine, deoxynojirimycin, and acarbose on MutB were tested. Among these castanospermine was a more potent inhibitor than deoxynojirimycin, with \(K_i\) values of 15 and 40 \(\mu\)M, respectively, whereas acarbose was completely noninhibitory. Tris, which inhibits some GH13 enzymes (45, 46), was also tested and found to be more weakly bound (\(K_i = 300 \mu\)M). Lineweaver-Burk plots of these inhibitors showed that all can be considered as competitive (data not shown).

Overall Structure—We determined the crystal structure of MutB in its native state and in complex with the competitive inhibitors Tris, deoxynojirimycin, and castanospermine (Fig. 2 and Table 1). Also, the structures of two MutB inactive mutants, D200A and E254Q, were solved in complex with \(\alpha\)-glucose and with its natural substrate sucrose, respectively (Fig. 2 and Table 1). All of the three-dimensional structures MutB-Tris, MutB-deoxynojirimycin, MutB-castanospermine, D200A-glucose, and E254Q-sucrose are highly similar to that of the free enzyme, as indicated by root mean square deviation values between C\(_{\alpha}\) in these five structures and the native form (0.24, 0.16, 0.32, 0.18, and 0.27 Å, respectively). Each MutB monomer (557 amino acids) forms an ellipsoid with dimensions of \(78 \times 51 \times 44\) Å\(^3\). The topology of MutB is similar to that observed in the vast majority of known GH13 enzyme structures and displays three domains (Figs. 3 and 4A), including a major central domain with a \((\beta/\alpha)_8\) barrel super secondary structure (N-terminal domain, also known as domain A in other GH13 proteins). As in Pall (24) (Fig. 4B), the N-terminal domain differs from the classical TIM barrel (47), by having five additional \(\alpha\)-helices, N\(_{42}\), N\(_{71}\), N\(_{88}\), N\(_{92}\), and N\(_{108}\), extending from \(\beta\)-strands N\(_{74}\), N\(_{87}\), and N\(_{88}\), respectively (Fig. 3). A loop-rich structure of 69 residues (subdomain, also referred to as domain B in other GH13 enzyme structures), comprising an \(\alpha\)-helix, a 3\(_{10}\)-helix, and three anti-parallel \(\beta\)-strands, protrudes from the \((\beta/\alpha)_8\) barrel between N\(_{73}\) and N\(_{93}\). This subdomain strongly interacts with the N-terminal domain by forming 22 hydrogen bonds and one salt bridge. Domain C, a 79-residue-long C-terminal domain, is made up of two antiparallel \(\beta\)-sheets. It is tightly connected to the N-terminal domain through a network of hydrogen bonds, clustered in three regions: between N\(_{74}\) and the segment including C\(_{B1}\) and C\(_{B2}\), between N\(_{87}\) and C\(_{B2}/C\(_{B7}\), and between N\(_{88}\) and C\(_{B2}\). An interdomain salt bridge connects Arg\(^{303}\) N\(_{74}\) to Asp\(^{485}\) O\(_{81}\). A calcium ion, Ca\(_{7001}\), identified in all MutB structures, is bound to the catalytic N-terminal domain in a...
loop easily accessible to the solvent. This calcium ion, which is hepta-coordinated, is located ~22 Å from the active site (Fig. 4A). An interaction between this ion and the peptide oxygen of Ile28 suggests that this calcium may be structural.

Active Site Architecture—The crystal structures of MutB show an active site defined by 14 residues, among which Asp200, Glu254, Asp327, Tyr64, His104, and His326 are conserved in the active sites of GH13 enzymes. Five additional residues (Asp61, Phe145, Phe164, Gln168, and Arg414) are conserved in AS (Neisseria polysaccharea amylosucrase), OGL (Bacillus cereus oligo-1,6-glucosidase), and Pall, all members of the so-called oligo-1,6-glucosidase subfamily (48). Finally, Phe144, Asn328, and Glu386 are specific for the sucrose isomerases Pall and MutB. The active site forms a deep pocket having a volume of 1000 Å³, which is perfectly adapted for binding a sucrose molecule (volume of ~450 Å³ (Fig. 4D)). This topology, also characteristic of Pall and OGL, was considered to be a crucial determinant of the substrate specificity of AS for sucrose (46).

Within the pocket Arg198, Asp384, and Arg284 determine the orientations of the catalytic residues and provide stability to the active site. At the bottom of the pocket Tyr64 and the salt bridge Asp200–Arg198 close the pocket, thus delimiting the nonreducing end to a single subsite, −1.

Flanking the entrance of the pocket is a pair of phenylalanines, Phe256 and Phe280, which are ~6 Å apart. They form a gateway ~11 Å above the bottom of the pocket and limit the reducing end of the substrate-binding site to subsite +1 (Figs. 4D and 5). Thr201 is located at the same level as these phenylalanines, as well as Arg284 and Arg291 situated at the entrance of the pocket. Thr201 is involved in a water-mediated hydrogen bond to the O-1 atom of the fructose moiety in sucrose and is only found in Pall and MutB structures. Arg284 is also specific to these enzymes, whereas Arg291 is present in AS as well. Subsite +1 is directly accessible to the solvent through the active site entrance channel, suggesting that the entry also serves as exit.

A short water channel of three adjacent water molecules, Wat4019, Wat4296, and Wat4073 connects the surface of the enzyme to the catalytic pocket in all structures. Both Wat4019 and Wat4296 are stabilized by direct hydrogen bonds to residues Asp61 and Arg328, whereas Wat4073 interacts with Thr60 and Ile385.

MutB-Tris Complex—Tris, known to inhibit a number of α-amylase family enzymes (45, 46), binds also to the MutB active site with direct hydrogen bonds to the nucleophile Asp200, to the general acid-base Glu254, and to Asp61, His104, and Arg314 (Fig. 2A and supplemental material) and mimics a sugar molecule bound in the −1 subsite. In the MutB-Tris com-
plex, double conformations (referred to as 1 and 2) of Phe$^{256}$ and Phe$^{280}$ are seen, indicating a possible concerted movement of these residues during the enzymatic reaction.

**MutB-Sucrose Substrate Complex**—The structure of the general acid/base knock-out mutant, E254Q, in complex with the natural substrate allowed us to define precisely substrate recognition and binding interactions. Sucrose (as other ligands complexed to MutB) is bound at the bottom of the deep pocket, which encloses the terminal glucose. It occupies two binding subsites, with the glucosyl ring at the $1_{\text{Sucrose}}$ position and the fructosyl moiety at the $1_{\text{Sucrose}}$ position. The conformation of sucrose may be described by the torsion angles $\varphi$ and $\psi$, where $\varphi$ is defined by O-5, C-1, O-1, and C-2 and $\psi$ by C-1, O-1, C-2, and O-2' (see Fig. 1C for a schematic representation of sucrose). In the three-dimensional structure of MutB-sucrose, $\varphi$ and $\psi$ angles are 20° and −66°, respectively. The glucosyl ring is in an undistorted $4_{\mathrm{I}}$ conformation, and the fructosyl ring is in a $4_{\mathrm{E}}$ conformation. As seen on Fig. 5B, the electron density is very well defined, and sucrose obviously adopts only one conformation. Binding of the glucose moiety is mediated by 14 direct hydrogen bonds and by stacking to Tyr$^{64}$ and Phe$^{164}$ (Fig. 2E and supplemental material). Tyr$^{64}$, in turn stacks with Pro$^{53}$, which is structurally conserved in the related TAKA-amylase (Aspergillus niger $\alpha$-amylase), OGL, and in Pall, and is further stabilized by hydrogen bonds to Asp$^{99}$ and Arg$^{198}$. Subsite +1 is defined only by Glu$^{254}$, Asp$^{327}$, and Arg$^{414}$, which hydrogen bond to the fructosyl part of sucrose and by Phe$^{164}$ and Phe$^{256}$ involved in hydrophobic interactions. Like the two catalytic residues Glu$^{254}$ and Asp$^{327}$, Arg$^{414}$ is a part of both binding subsites, suggesting an important role of this residue in substrate binding and positioning. Despite differences of 0.5–1 Å observed in the backbone conformation of the regions defining

---

**FIGURE 3.** Sequence alignment of Pall and MutB. Secondary structure elements of MutB are indicated above the alignment, and flexible regions described in the text and shown in Fig. 4C are indicated with similar color coding.
The catalytic pocket and involving residues 286–295 and 376–414 (Fig. 4C), comparison of E254Q-sucrose and MutB native structures show that they are highly similar but that upon binding of sucrose in the active site, Arg414 adopts clearly a different conformation, displacing its side chain by 1.4 Å to accommodate both the glucose and the fructose rings (Fig. 5, A and B). The importance of Arg414 as indicated by the three-dimensional structures was confirmed by site-directed saturation mutagenesis, with most of the obtained mutants exhibiting very low activity on sucrose. Within the obtained mutants, only R414T showed weak isomerase activity, whereas mutants R414K/S/A/G were enzymes with low hydrolytic activity. Finally, R414L/W destroyed the enzyme activity.

Whereas five water-mediated hydrogen bonds are seen between the fructosyl moiety and MutB, only one water molecule associates the enzyme to the pyranosyl ring of sucrose by hydrogen bonding to O-3 and O-4. This probably reflects that fructose has to be more mobile than the glucosyl moiety to undergo conformational changes and to easily exit the pocket during the catalytic cycle.

A comparable substrate-binding site was observed in the structure of the inactive mutant E328Q (E254Q in MutB) of AS in complex with sucrose (49). Locations of the glucosyl moiety and interacting residues at subsite −1 are conserved, whereas the position of the fructosyl ring is quite different because it rotates 20° around the C-1–O-1 bond, resulting in distinct dihedral angles being (φ,ψ) = (44°,−93°) in AS mutant E328Q-sucrose compared with (φ,ψ) = (20°,−66°) in MutB mutant E254Q-sucrose. Moreover, oxygen atoms O-1’ and O-6’ (fructosyl ring) both point in opposite directions when compared with their counterparts in the AS E328Q-sucrose structure, leading to distinct interaction patterns in subsite +1 between the two structures. In the MutB complex, only a water-mediated interaction to the O-1’ atom is observed, compared with two direct hydrogen bonds between O-1’ and the catalytic acid-base in the AS structure. O-6’ interacts with Arg414, with Asp327, and with a water molecule in the MutB complex but performs in AS-sucrose three direct hydrogen bonds to Arg446 (corresponding to Arg291 in MutB), and to Asp394 that is not conserved in MutB. In AS, Arg509 (Arg414 in MutB) is too distant (3.6 Å) to perform hydrogen bonding to O-6’. The tighter interactions in the catalytic pocket at subsite −1 of the MutB complex and points now directly toward the O-3 and O-4 hydroxyls.
groups, suggesting again, along with Arg414 mutants (described earlier), that it plays a key role in recognition and stabilization of the ligand. Binding of castanospermine also results in a 0.9-Å shift of Phe144 and in 0.7- and 1-Å displacements of Arg284 and Glu386, respectively.

The hydroxyl group on castanospermine C-6 lays axially, points toward the entrance of the pocket, and makes only one contact to the catalytic Asp327. Also, a salt bridge (2.6 Å) is formed between the nucleophile Asp200 O and the imino-NH of castanospermine. When comparing MutB_castanospermine and E254Q_sucrose complexes, considerable movements of the backbone are observed in three regions surrounding the active site, in that residues ranging from 57 to 62, 212 to 226, and 286 to 294 make a concerted movement of 0.5–1 Å (Fig. 4C).

**MutB-Deoxynojirimycin Transition State Mimic and D200A-Glucose Complexes**—The electron density maps (Fig. 5, D and E) show that the conformation and the interaction patterns (Fig. 2, B and D) of deoxynojirimycin and glucose within subsite −1 are strongly similar to those described for the glucose ring in E254Q_sucrose (Fig. 2E). When the positions of glucosyl rings are compared between glucose and sucrose complexes, the major difference is a 10° rotation around the C-3–C-4 bond (the glucosyl moiety of sucrose being rotated away from co-planarity with Tyr64) and a small translation. The major differences in atomic positions are therefore observed for atoms C-1 and O-5, which are ~0.5 Å apart from one glucose ring to another in the glucose and sucrose complexes, respectively.

The catalytic Asp327 makes a very tight hydrogen bond (2.3 and 2.5 Å in MutB_castanospermine and D200A-glucose structures, respectively) to O-2 of the ligand. In MutB_deoxynojirimycin, Glu254 interacts with the inhibitor via a water molecule, whereas it is directly linked to glucose O-1 in the D200A-glucose structure. In MutB_deoxynojirimycin, Asp200 (mutated to alanine in the complex with glucose) performs a strong interaction with N-5 of deoxynojirimycin. An overlay of the two structures shows that the distance between O-1 at the anomeric carbon of the bound glucose (here in β-configuration) and the nucleophile Asp200 Oβ2, residing at the same position in MutB_deoxynojirimycin would be 1.4 Å, indicating the covalent character of this interaction. Because glucose is a 40%/60% mixture of α/β anomers at the temperature of the co-crystallization experiment, an obvious question is how this ligand binds to the active site pocket. When the α-anomer is superimposed on the bound glucose, the only difference in binding is that O-1, which hydrogen bonds to Glu254 Oe1 instead would bond to Glu254 Oe2. The latter is,
however, geometrically less favorable, and the distance between donor and acceptor atoms would be ~3 Å whereas the distance in the β-anomer complex is 2.5 Å. The interaction with Asp\(^{200}\) would also be clearly impossible in the case of the α-anomer (not shown). These interaction energy differences seem to be responsible for the β-anomer specificity. In covalent glucosyl-enzyme intermediates as seen in experimentally determined three-dimensional structures of cyclodextrin glycosyltransferase and AS (22, 23), the sugar units are also in β-conformations.

**Product Specificity**—The pair of phenylalanines, Phe\(^{256}\) and Phe\(^{280}\), forming an aromatic clamp and flanking the active site entrance were present in double conformations in the MutB-Tris complex, indicating a possible concerted movement of these residues during the enzymatic reaction. In the other MutB structures, Phe\(^{256}\) and Phe\(^{280}\) adopt only one conformation corresponding to one of the two conformations seen in the MutB-Tris complex (Fig. 5). Mutagenesis studies at these positions resulted in mutant enzymes with strongly altered product specificity. When mutating either of these phenylalanines to small amino acids (G, A, S, or T) or to K, an enzyme exhibiting major hydrolytic activity was obtained, whereas aromatic transformants maintain the isomerase activity. Notably, the double mutant F256A/F280A has nearly exclusively hydrolytic activity.

In MutB, only a few water molecules are directly available for the hydrolytic reaction, although two or three are enclosed in a small cavity located inside the catalytic pocket close to Phe\(^{256}\). Within this cavity Wat\(^{4248}\) is presumably the catalytic water molecule, because its location and its distance from the putative glucosyl-enzyme intermediate covalent bond are consistent with a hydrolytic mechanism previously described (13, 50). A direct interaction of Wat\(^{4248}\) with the acid-base Glu\(^{254}\) strengthens the hypothesis of this water being catalytic, and Wat\(^{4248}\) is hydrogen-bonded in all structures to Wat\(^{4996}\) and occasionally to Wat\(^{4754}\). Wat\(^{4248}\) also hydrogen bonds to Asp\(^{327}\), Arg\(^{286}\), and Asn\(^{328}\), and Wat\(^{4754}\) is linked to Asn\(^{325}\) and Asn\(^{328}\). Surprisingly, in Pall despite an identical environment, only one water molecule (Wat\(^{401}\), counterpart to Wat\(^{4248}\) in MutB) is observed in this site.

**DISCUSSION**

All GH13 enzymes operate using a two-step double-displacement mechanism involving formation and breakdown of a covalent glucosyl-enzyme intermediate (18–23). Both steps are suggested to proceed via oxocarbenium ion transition states.

Structural analysis and comparative studies of MutB structures allowed, for the first time, formation of a detailed image of the first step in this catalytic reaction. The E254Q-sucrose structure illustrates not only how the substrate binds prior to catalysis but also the role of Glu\(^{254}\) as a proton donor to the interglycosidic oxygen of sucrose. Indeed, the nearly linear interaction observed between sucrose O-1 and Glu\(^{254}\) in the wild type enzyme would minimize the proton transfer energy barrier during catalysis and favor the transition state analogue binding preceding the covalent intermediate. The distance between Asp\(^{200}\) O61 and sucrose C-1 in the E254Q-sucrose structure is 2.7 Å, which favors a possible nucleophilic attack. Asp\(^{200}\), the catalytic nucleophile attacks C-1, leading to the formation of an oxocarbenium ion transition state, as mimicked by the deoxynojirimycin inhibitor in the MutB-deoxynojirimycin structure (Fig. 5D). As concerns the slightly twisted boat conformation adopted by the castanospermine in the MutB-castanospermine structure (Fig. 5C), it may be the precursor of the transition state, assumed to be in a half-chair conformation (51, 52). In the latter the glycosidic linkage at the anomeric carbon would be in an axial position, which is needed for the terminal glycosyl residue to be completely enclosed in the pocket with the nucleophilic Asp\(^{200}\) adjacent to C1 and to allow the aglycon to depart following bond cleavage.

Information on what the covalent glycosyl-enzyme intermediate would look like is obtained from the D200A-glucose structure. Reaction steps succeeding the formation of the covalent glycosyl-enzyme intermediate are more speculative. The low number of interactions between the fructosyl ring and the enzyme active site may allow this moiety to be displaced, to rotate, and/or to tautomerize to form the two isomers, isomaltohexose and trehalulose. The product specificity of sucrose isomerases depends substantially on the opportunity for tautomerization of sucrose during the reaction that is required for trehalulose formation. We speculate that in subsite −1 the glucose moiety of trehalulose most probably will bind in a chair conformation as seen in all structures within this study with the exception of MutB-castanospermine and that differences in interactions between the enzyme and sucrose isomers are to be found in the acceptor subsite. Our study highlights structural features involved in the distinct product specificities of these enzymes, especially the nature of the aromatic clamp and the associated volume of the catalytic pocket. The mutant F256A/F280A, which destroys the clamp, would result in a pocket being continuously open, thus allowing the hydrolyzed glucose to diffuse rapidly away as opposed to the clamp, which locks the glucose moiety in the enzyme intermediate and prevents release of the fructosyl ring. Although aromatic transformants preserve isomerase activity, the nature and position of the aromatic residues seem to modulate reaction specificity by altering the volume of the pocket and hence the space for tautomerization. We therefore propose that the aromatic clamp phenylalanines not only participate in substrate recognition but also partially control the reaction specificity. An aromatic clamp with similar functions has been described in GH5 family enzymes such as the fungal exoglucanase subfamily (53) and the endocellulase CelCCA (54). The clamp is also suggested to be involved in both entry and release to/from the active site and therefore in isomerase versus hydrolytic activity of the enzyme. Besides the narrow size of the pocket, this may explain the weak hydrolytic activity of these enzymes because there is little room for water to enter when a substrate molecule occupies the pocket. In other glycoside hydrolases, optimized systems have been described in which the active site can be continuously and efficiently supplied with water (55–57). As concerns the water channel connecting the surface to the catalytic pocket, also identified in AS (46), its function is still uncertain. We hypothesize that it serves to refill the catalytic pocket with water, a process that may be advantageous in terms of efficiency.
because water refill and product release can be performed simultaneously.

Our structures indicate mobile regions (Fig. 4C) that most probably play important roles during the enzymatic reaction: the subdomain, the loops between Nβ4 and Nα4, between Nβ6 and Nα6 (which contains the so-called “isomerization motif” described for Pall (24, 25)), and between Nβ8 and Nα8', which includes η8. A significant deviation in the backbone is observed in this isomerization motif between MutB and Pall (Fig. 4C), the maximum deviation between Ca of MutB and Pall being 4.3 Å around Asp290 (MutB). The deletion of one residue four positions upstream from Asp290 in MutB as compared with Pall, as well as the charge differences three to five positions downstream from Asp290 between these isomerases may explain this deviation. Further, a shift of 2 Å between the two structures is observed (residues 376–384 in MutB). Indeed, examination of MutB and Pall surfaces colored as a function of charge (Fig. 6) revealed that MutB is more negatively charged on the face displaying the entrance of the catalytic pocket. Analysis of the surface distribution of invariant residues between MutB and Pall reveals that most of the nonconserved residues are located at the surface of the enzyme on this side, which may contribute to explain the distinct charge distributions between the two enzymes.

A complete analysis of important structural features requires a comparison with an enzyme-substrate complex of an isomaltulose synthase, which so far is not available. However, analyses of the structures described herein give a detailed image of conformational changes of important enzyme residues upon ligand binding and provide, together with the mutagenesis results, a solid basis for the understanding of substrate and product specificities and for rational protein engineering.

However, one should also consider remote sequence differences that alter the shape and/or stability of the substrate pocket. Distant prolines or residues involved in salt bridges or even hydrogen bonds that alter the rigidity of loops enclosing the substrate seem to be important, as recently found for the structurally related AS (58). Future work should target residues distant from the catalytic pocket to minimize the impact of such mutations on its structure, and changes resulting in midrange or even long range effects have to be envisioned.

Acknowledgments—We are grateful to the staff members of ID14-2, ID14-4, and FIP BM30A beamlines at the European Synchrotron Radiation Facility (Grenoble, France) for technical advice and precious help during data collection.

REFERENCES

1. Takazoe, I. (1989) in Progress in Sweeteners (Grenby, T. H., ed.) pp. 143–167, Elsevier, Barking, UK
2. Hamada, S. (2002) Pure Appl. Chem. 74, 1293–1300
3. Ooshima, T., Izumitani, A., Minami, T., Fujitani, T., Nakajima, Y., and Hamada, S. (1991) Caries Res. 25, 277–282
4. Yamada, K., Shinhara, H., and Hosoya, N. (1985) Nutr. Rep. Int. 32, 1211–1222
5. Lichtenhaller, F. W., and Peters, S. (2004) C. R. Chimie 7, 65–90
6. Cheetham, P. S. (1984) Biochem. J. 220, 213–220
7. Miyata, Y., Sugitani, T., Tsuyuki, K., Ebashi, T., and Nakajima, Y. (1992) Biosci. Biotechnol. Biochem. 56, 1680–1681
8. Park, Y. K., Uekane, R. T., and Pupin, A. M. (1992) Biotech. Lett. 14, 547–551
9. Nagai-Miyata, J., Tsuyuki, K., Sugitani, T., Ebashi, T., and Nakajima, Y. (1993) Biosci. Biotechnol. Biochem. 57, 2049–2053
10. Nagai, Y., Sugitani, T., and Tsuyuki, K. (1994) Biosci. Biotechnol. Biochem. 58, 1789–1793
11. Huang, J. H., Hsu, L. H., and Suc, J. C. (1998) J. Ind. Microbiol. Biotechnol. 21, 22–27
12. Mattes, R., Klein, K., Schiwech, H., Kunz, M., and Munir, M. (Month July 28, 1998) U. S. Patent 5786140
13. Veronese, T., and Perlot, P. (1998) FEBS Lett. 441, 348–352
14. Zhang, D., Li, X., and Zhang, L. H. (2002) Appl. Environ. Microbiol. 68,
Trehalulose Synthase Structures and Sucrose Isomerization

2676–2682
15. Wu, L., and Birch, R. G. (2005) *Appl. Environ. Microbiol.* 71, 1581–1590
16. Salvucci, M. E. (2003) *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* 135, 385–395
17. Henrisat, B., and Davies, G. (1997) *Curr. Opin. Struct. Biol.* 7, 637–644
18. Koshland, D. E. (1953) *J. Biol. Chem.* 203, 201–212
19. Davies, G., and Henrisat, B. (1995) *Structure* 3, 853–859
20. Yoshioka, Y., Hasegawa, K., Matsuura, Y., Katsube, Y., and Kubota, M. (1997) *J. Mol. Biol.* 271, 619–628
21. Mosi, R., He, S., Uitdehaag, J., Dijkstra, B. W., and Withers, S. G. (1997) *Biochemistry* 36, 9927–9934
22. Uitdehaag, J. C., Mosi, R., Kalk, K. H., van der Veen, B. A., Dijkhuizen, L., Withers, S. G., and Dijkstra, B. W. (1999) *Nat. Struct. Biol.* 6, 432–436
23. Jensen, M. H., Mirza, O., Albenne, C., Remaude-Simeon, M., Monsan, P., Gajhede, M., and Skov, L. K. (2004) *Biochemistry* 43, 3104–3110
24. Zhang, D., Li, N., Lok, S. M., Zhang, L. H., and Swaminathan, K. (2003) *J. Biol. Chem.* 278, 35428–35434
25. Zhang, D., Li, N., Swaminathan, K., and Zhang, L. H. (2003) *FEBS Lett.* 534, 151–155
26. Ravaud, S., Watzlawick, H., Haser, R., Mattes, R., and Aghajari, N. (2005) *Acta Crystallogr. F Struct. Biol. Crystalliz. Comm.* 61, 100–103
27. Ravaud, S., Watzlawick, H., Mattes, R., Haser, R., and Aghajari, N. (2005) *Biologia Cell. Mol. Biol. Sect.* 60, 99–105
28. Bradford, M. M. (1976) *Anal. Biochem.* 72, 248–254
29. Kabsch, W. (1993) *J. Appl. Crystallogr.* 26, 795–800
30. Leslie, A. G. W. (1991) in *Crystallography Computing* (Moras, D., Podjarny, A. D., and Thierry, J. C., eds) pp. 50–61, Oxford University Press, Oxford, UK
31. CCP4 Collaborative Computational Project Number 4 (1994) *Acta Crystallogr. Sect. D Biol. Crystallogr.* 50, 760–763
32. Navaza, J. (2001) *Acta Crystallogr. Sect. D Biol. Crystallogr.* 57, 1367–1372
33. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. Sect. D Biol. Crystallogr.* 54, 905–921
34. Roussel, A., and Camblalou, C. (1992) *TURBO-FRODO*, Biographics, AFMB, Marseille, France
35. Brünger, A. T. (1992) *Nature* 355, 472–475
36. Kleywegt, G. J., and Jones, T. A. (1998) *Acta Crystallogr. Sect. D Biol. Crystallogr.* 54, 1119–1131
37. Laskowski, R. A., MacArthur, M. W., and Moss, D. S. (1993) *J. Appl. Crystallogr.* 26, 283–291
38. Hooft, R. W., Vriend, G., Sander, C., and Abola, E. E. (1996) *Nature* 381, 272
39. McDonald, I. K., and Thornton, J. M. (1994) *J. Mol. Biol.* 238, 777–793
40. Kleywegt, G. J., and Jones, T. A. (1994) *Acta Crystallogr. Sect. D Biol. Crystallogr.* 50, 178–185
41. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucleic Acids Res.* 22, 4673–4680
42. Gouet, P., Robert, X., and Courcelle, E. (2003) *Nucleic Acids Res.* 31, 3320–3323
43. Nicholls, A., Sharp, K. A., and Honig, B. (1991) *Proteins* 11, 281–296
44. DeLano, W. L. (2002) *PyMOL*, DeLano Scientific, San Carlos, CA
45. Aghajari, N., Feller, G., Gerday, C., and Haser, R. (1998) *Structure* 6, 1503–1516
46. Skov, L. K., Mirza, O., Henriksen, A., De Montalk, G. P., Remaude-Simeon, M., Sarcabal, P., Willemot, R. M., Monsan, P., and Gajhede, M. (2001) *J. Biol. Chem.* 276, 25273–25278
47. Banner, D. W., Bloomer, A. C., Petsko, G. A., Phillips, D. C., Pogson, C. I., Wilson, I. A., Corran, P. H., Furth, A. J., Milman, J. D., Offord, R. E., Priddle, J. D., and Waley, S. G. (1975) *Nature* 255, 609–614
48. Oslanovea, A., and Janecek, A. (2002) *Cell Mol. Life Sci.* 59, 1945–1959
49. Mirza, O., Skov, L. K., Remaude-Simeon, M., Potocki de Montalk, G., Albenne, C., Monsan, P., and Gajhede, M. (2001) *Biochemistry* 40, 9032–9039
50. Veronese, T., and Perlot, P. (1999) *Enzyme Microbial Technol.* 24, 263–269
51. Strynadka, N. C., and James, M. N. (1991) *J. Mol. Biol.* 220, 401–424
52. Kuroki, R., Weaver, L. H., and Matthews, B. W. (1993) *Science* 262, 2030–2033
53. Cutfield, S. M., Davies, G. J., Murshudov, G., Anderson, B. F., Moody, P. C., Sullivan, P. A., and Cutfield, J. F. (1999) *J. Mol. Biol.* 294, 771–783
54. Ducros, V., Czjzek, M., Belaich, A., Gaudin, C., Fierobe, H. P., Belaich, J. P., Davies, G. J., and Haser, R. (1995) *Structure* 3, 939–949
55. Boel, E., Brady, L., Brzozowski, A. M., Derewenda, Z., Dodson, G. G., Jensen, V. J., Petersen, S. B., Swift, H., Thim, L., and Woldike, H. F. (1990) *Biochemistry* 29, 6244–6249
56. Aghajari, N., Feller, G., Gerday, C., and Haser, R. (1998) *Protein Sci.* 7, 564–572
57. Robert, X., Haser, R., Gottschalk, T. E., Ratajczak, F., Driguez, H., Svensson, B., and Aghajari, N. (2003) *Structure* 11, 973–984
58. van der Veen, B. A., Potocki-Veronese, G., Albenne, C., Joula, G., Monsan, P., and Remaude-Simeon, M. (2004) *FEBS Lett.* 560, 91–97