THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 285, NO. 23, pp. 17763–17770, June 4, 2010

Structural Basis for Substrate Selectivity in Human Maltase-Glucoamylase and Sucrase-Isomaltase N-terminal Domains*[S]

Received for publication, November 10, 2009, and in revised form, February 19, 2010 Published, JBC Papers in Press, March 31, 2010, DOI 10.1074/jbc.M109.078980

Lyann Sim[‡1], **Carly Willemsma**[‡], **Sankar Mohan**[§], **Hassan Y. Naim**[¶], **B. Mario Pinto**[§], **and David R. Rose**[‡‖2]

From the [‡]*Ontario Cancer Institute and Department of Medical Biophysics, University of Toronto, Toronto, Ontario M56 2M9, Canada, the* [§]*Department of Chemistry, Simon Fraser University, Burnaby, British Columbia V5A 1S6, Canada, the* [¶]*Department of Physiological Chemistry, University of Veterinary Medicine, D-3059 Hannover, Germany, and the* [‖]*Department of Biology, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada*

Human maltase-glucoamylase (MGAM) and sucrase-isomaltase (SI) are small intestinal enzymes that work concurrently to hydrolyze the mixture of linear α-1,4- and branched α-1,6-oligosaccharide substrates that typically make up terminal starch digestion products. MGAM and SI are each composed of duplicated catalytic domains, N- and C-terminal, which display overlapping substrate specificities. The N-terminal catalytic domain of human MGAM (ntMGAM) has a preference for short linear α-1,4-oligosaccharides, whereas N-terminal SI (ntSI) has a broader specificity for both α-1,4- and α-1,6-oligosaccharides. Here we present the crystal structure of the human ntSI, in apo form to 3.2 Å and in complex with the inhibitor kotalanol to 2.15 Å resolution. Structural comparison with the previously solved structure of ntMGAM reveals key active site differences in ntSI, including a narrow hydrophobic +1 subsite, which may account for its additional substrate specificity for α-1,6 substrates.

In humans, six enzyme activities (two α-amylase and four α-glucosidase activities) are involved in the breakdown of dietary starches and sugars into glucose. The α-glucosidase activities are associated with two small intestinal membrane-bound enzymes: maltase-glucoamylase (MGAM)[3] and sucrase-isomaltase (SI) (for a review, see Refs. 1 and 2). MGAM and SI are composed of duplicated catalytic domains: an N-terminal membrane-proximal domain (ntMGAM and ntSI) and a C-terminal luminal domain (ctMGAM and ctSI). The domains are anchored to the small intestinal brush-border membrane via an *O*-glycosylated stalk stemming from the N-terminal domain. Given that MGAM and SI genes arose from duplication and divergence of an ancestral gene, which itself has undergone tandem duplication (3), the N-terminal domains of MGAM and SI are more similar to one another in sequence, as are the C-terminal domains (~60% sequence identity), than are the N- and C-terminal domains associated with the same enzyme (~40% sequence identity).

Within the carbohydrate-active enzymes (CAZY) classification system (36), which groups enzymes based on sequence similarity and reflects the functional and structural similarities of family members, N- and C-terminal MGAM and SI domains are members of the glycoside hydrolase 31 family (GH31). The four domains exhibit *exo*-glucosidase activities against α-1,4-linked maltose substrates (Fig. 1*A*) but display different specificities for malto-oligosaccharides of various lengths (4–6). ntSI and ctSI subunits have additional activity for the α-1,6 linkages of starch branch points (and isomaltose substrates; Fig. 1*A*) and the α-1,2 linkage of sucrose, respectively (7), and are historically referred to as isomaltase and sucrase.

As they are involved in the breakdown of dietary sugars and starches, MGAM and SI are attractive targets for inhibition by α-glucosidase inhibitors as a means of controlling blood glucose levels in individuals with type 2 diabetes (8). Acarbose (Fig. 1*B*) is the most widely used α-glucosidase inhibitor currently on the market and has been shown to be an efficient inhibitor of α-amylase (9) and of the C-terminal domains of MGAM and SI (4), but a weaker inhibitor of the N-terminal domains (4, 10). More recently, studies on a new class of α-glucosidase inhibitors based on the natural extracts of the plant *Salacia reticulata*, including the active compounds salacinol, kotalanol (Fig. 1*B*), and de-*O*-sulfonated kotalanol, have been shown to be stronger inhibitors of ntMGAM compared with acarbose (10–13).

A thorough understanding of the structural basis for substrate specificity in these intestinal α-glucosidases requires the analysis and comparison of crystal structures of the individual catalytic domains. Structural details of the active sites will give insight into how the MGAM and SI domains have evolved their substrate binding sites to accommodate different sugar and starch linkages.

We have previously solved the crystal structure of ntMGAM and analyzed co-crystal structures in complex with α-glucosi-

* This work was supported by Canadian Institutes for Health Research (CIHR) Grant FRN79400 and Heart and Stroke Foundation of Ontario Grant NA-6305.

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

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Figs. 1–3.

[1] Recipient of funding through a postgraduate scholarship from the Natural Sciences and Engineering Research Council of Canada (NSERC).

[2] To whom correspondence should be addressed: Dept. of Biology, University of Waterloo, 200 University Ave. West, Waterloo, Ontario N2L 3G1, Canada. E-mail: drrose@uwaterloo.ca.

[3] The abbreviations used are: MGAM, maltase-glucoamylase; SI, sucrase-isomaltase; ntMGAM, N-terminal domain of maltase-glucoamylase; ntSI, N-terminal domain of sucrase-isomaltase; ctMGAM, C-terminal domain of maltase-glucoamylase; ctSI, C-terminal domain of sucrase-isomaltase; *p*NPG, *p*-nitrophenyl α-D-glucopyranoside; TEV, tobacco etch virus; r.m.s., root mean square; aa, amino acids; Bistris propane, 1,3-bis[tris(hydroxymethyl)methylamino]propane.

dase inhibitors acarbose and salacinol-based compounds (6, 13). From structural studies of the ntMGAM-acarbose complex, the active site of ntMGAM was found to be composed of a shallow substrate-binding pocket comprising −1 and +1 subsites (6). Substrates bind to the pocket via their non-reducing end, with the non-reducing sugar ring interacting with the buried −1 subsite and the reducing ring interacting with the surface-exposed +1 subsite. Substrate cleavage occurs between −1 and +1 subsites and follows a catalytic mechanism, resulting in a net retention of configuration at the anomeric center.

Here we present the crystal structure of the N-terminal domain of human SI in apo form to 3.2 Å and in complex with kotalanol to 2.15 Å resolution and compare its active site and kinetic features with those of ntMGAM. This analysis has identified the active site similarities in the −1 subsite and differences in the +1 subsite that may account for the additional activity of ntSI for α-1,6-linkages.

EXPERIMENTAL PROCEDURES

Cloning and Expression—The coding sequence of ntSI, encompassing residues 62–931 of full-length human SI (GenBank™ accession number NP_001032), was PCR-amplified from the pSG8-SI vector (14) using upstream primer 5′-CCGGAttcgaaGAAAATGTCCAAATGTG-3′ and downstream primer 5′-GGCCAgtttaaacTCACCATTGAACACTAAAG-3′. The PCR amplification introduced a 5′ BstBI and 3′ PmeI restriction site for cloning into a similarly digested pMT-TEVA *Drosophila* expression vector (15) to make a pMT-TEVA-ntSI vector.

ntSI was expressed in *Drosophila* S2 cells using the *Drosophila* Expression System (Invitrogen), following a similar expression protocol as for ntMGAM (10). Stable cell lines were obtained by co-transfecting S2 cells with pMT-TEVA-ntSI and pCopBlast, a blasticidine resistance plasmid. Transfected cells underwent single cell selection, and the best expressing single cell clones were adapted to serum-free medium, scaled up to Fernbach shaker flasks (800 ml/flask), and induced with 4 μM $CdCl_2$, and the medium was harvested after 72 h. The pMT-TEVA-ntSI construct allowed for secretion of ntSI into the culture medium under the control of a metallothionein promoter with an N-terminal hexahistidine tail engineered with a tobacco etch virus (TEV) cut site.

Protein Purification—The secreted ntSI protein was purified from the medium using nickel-Sepharose resin (GE Healthcare) (9 ml of resin/liter of medium). Nickel affinity-purified ntSI ran as a single band (~100 kDa) on a reducing SDS-polyacrylamide gel but as three distinct bands on a non-reducing gel. Further purification using size affinity chromatography (S200) revealed three peaks, each corresponding to a band on the non-reducing gel. Fractions were tested for the ability to cleave *p*-nitrophenyl α-D-glucopyranoside (*p*NPG), and the third peak was found to be the only active peak. Fractions in this peak were pooled and concentrated. The total yield of pure ntSI from expression in *Drosophila* S2 cells was ~10 mg/liter of medium.

Crystallization, Data Collection, and Processing—ntSI was crystallized using the hanging drop vapor diffusion method using 1 μl of protein (~7.5–22 mg/ml) plus 1 μl of reservoir buffer (0.5 M NaCl, 0.1 M Bistris propane, pH 7.0, and 18% polyethylene glycol 4000). These conditions produced clusters of thin crystal plates, which were then used to streak seed equilibrated drops in order to grow single crystal plates. Equilibrated drops were composed of 1.5 μl of protein and 3 μl of reservoir buffer (0.1 M $MgCl_2$, 0.1 M Bistris propane, pH 7.0, 15% polyethylene glycol 4000). For easier crystal handling, crystals were cross-linked with glutaraldehyde (16) before being drawn into a 4:1 ratio of paratone/mineral oil. Crystal complexes were obtained by soaking cross-linked crystals in mother liquor supplemented with 200 μM kotalanol for ~2 h.

Data for the apo-ntSI and ntSI-kotalanol crystals were collected on an ADSC Quantum-270 CCD detector at beamline F1 at the Cornell High Energy Synchrotron Source and were processed with HKL2000 (17). Data collection statistics are given in Table 1.

Phasing and Refinement—apo-ntSI phases were solved by molecular replacement (Phaser (18)), using a trimmed structure of ntMGAM (Protein Data Bank accession code 2QLY) as a model. Four monomers were found in the asymmetric unit, corresponding to a crystal solvent content of 52%. Loop and side-chain rebuilding was carried out in Coot (19) and was alternated with cycles of non-crystallographic symmetry restrained refinement in Refmac (20). The electron density allowed for the building of a $Man_3GlcNAc_2$ chain in two of the monomer active sites.

The apo-ntSI structure was then used as a starting model in the refinement of the ntSI-kotalanol structure (2.15 Å). Rigid body refinement was first carried out in Refmac (20), followed by restrained refinement (without non-crystallographic symmetry restraints). The higher resolution data of ntSI-kotalanol allowed for rebuilding of missing side-chain residues in Coot (19) and the addition of solvent molecules using ARP/wARP (21). The electron density in the active site allowed for the modeling of kotalanol molecules in two of the monomer active sites and of $Man_2GlcNAc_2$ glycans and Tris molecules in the remaining two active sites.

B-factor analysis revealed that monomer D had significantly higher overall B-factor (48 $Å^2$) compared with monomers A, B, and C (22, 25, and 27 $Å^2$). TLS refinement (22) was then carried out for 15 cycles using each monomer as a TLS group, followed by 15 cycles of restrained refinement. After TLS refinement, the R/R_{free} dropped from 0.198/0.253 to 0.178/0.223 and resulted in a slight improvement in electron density.

A summary of the refinement statistics is given in Table 1. Inhibitor topologies and restraints were generated using the PRODRG server (23), and protein stereochemistry was validated using the Molprobity server (24). Graphics were generated using PyMOL (available on the World Wide Web).

Enzyme Assay—ntSI and ntMGAM activities against maltose and isomaltose substrates were determined using the glucose oxidase assay (25). Assays were carried out in 96-well microtiter plates containing substrate (2–40 mM maltose or 2.5–400 mM isomaltose), 100 mM MES buffer, pH 6.5, and enzyme (4 nM ntMGAM or 2 nM ntSI in maltose assays and 20 nM ntMGAM or 2 nM ntSI in isomaltose assays). Assay reactions were linear over the entire reaction period; however, ntSI required supplementation of 0.05% Tween 20 to maintain enzyme stability.

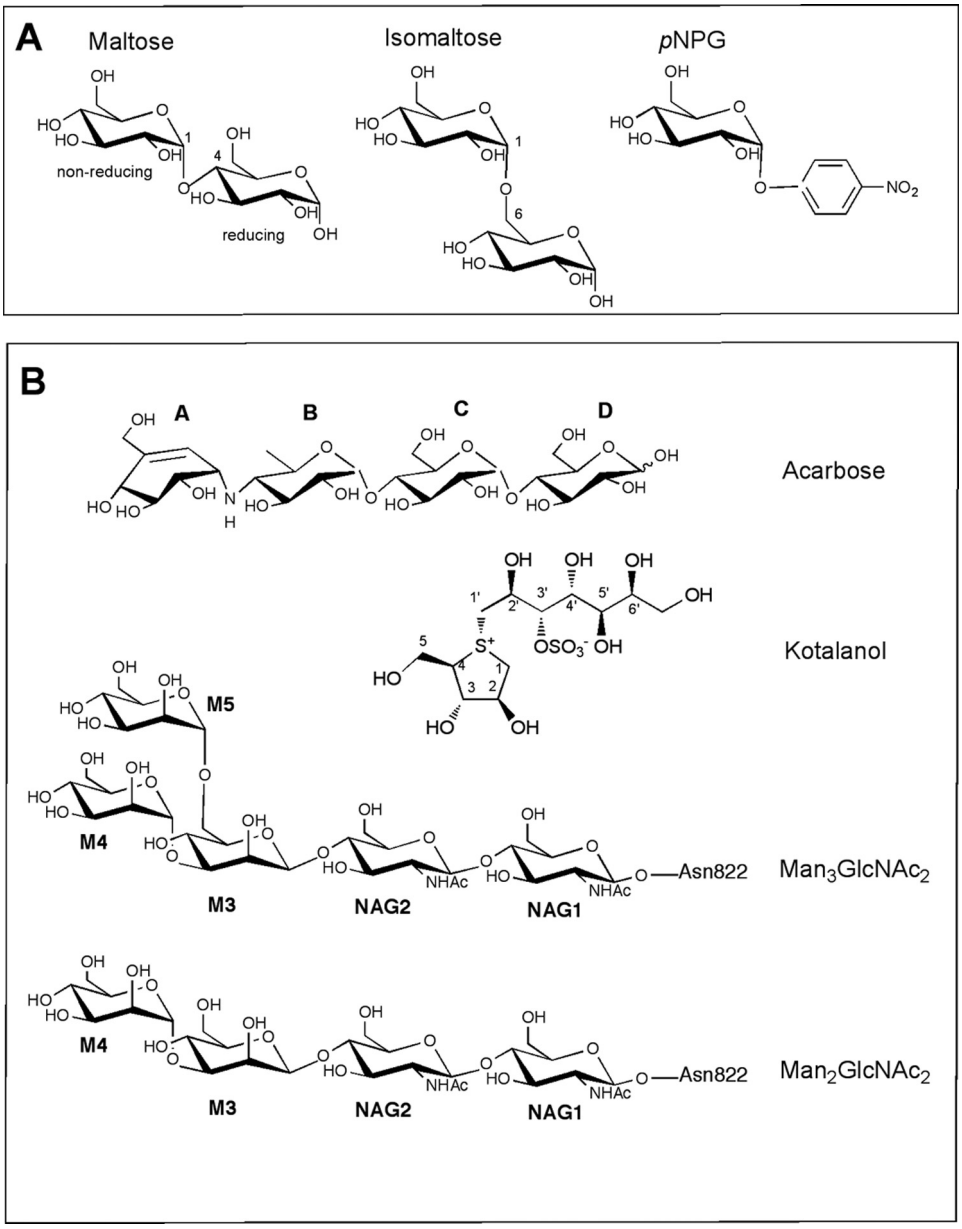

FIGURE 1. **Substrates, inhibitors, and glycan ligands.** *A*, maltose, isomaltose, and *p*NPG substrates. Nonreducing and reducing ends are indicated for the maltose substrate. *B*, α-glucosidase inhibitors (acarbose and kotalanol) and Asn^{822}-linked oligosaccharide ligands ($Man_3GlcNAc_2$ and $Man_2GlcNAc_2$).

Reactions were incubated at 37 °C for 30 min and terminated with Tris-HCl (pH 7) to a final concentration of 1 M. Aliquots of glucose oxidase/peroxidase reagent (125 μl) were added to each well and left to develop at 37 °C for 30 min. Absorbance was measured at 450 nm and compared with a standard glucose curve to determine the amount of glucose released. Each molecule of maltose or isomaltose releases two molecules of glucose upon cleavage. All reactions were performed in triplicate, and absorbance measurements were averaged to give a final result. Inhibition assays of ntMGAM and ntSI with kotalanol and acarbose followed protocols that have been described previously (10, 13).

Kinetic parameters of ntMGAM and ntSI were also determined using *p*NPG as a substrate. Reactions were carried out in microtiter plates and consisted of 1.5–40 mM *p*NPG, 0.1 M MES, pH 6.5, 20 nM ntMGAM or 10 nM ntSI supplemented with 0.05% Tween 20. Plates were incubated at 37 °C for 45 min before being quenched with 0.5 M sodium carbonate, and absorbance was measured at 405 nm.

RESULTS

Substrate Specificity—The substrate specificities of recombinant human ntMGAM and ntSI for maltose, isomaltose, and the general glucosidase substrate *p*NPG (Fig. 1*A*) were examined, and kinetic parameters were calculated (Table 2). For ntSI, specificity constants (k_{cat}/K_m) were calculated to be 19 ± 4, 9 ± 2, and 13 ± 1 s^{-1} mM^{-1}, for maltose, isomaltose, and *p*NPG substrates, respectively. The constants are on the same order of magnitude, suggesting that ntSI shows no overt preference for one substrate over another. In contrast, ntMGAM displays a clear preference for maltose substrates because its k_{cat}/K_m for maltose (26 ± 8 s^{-1} mM^{-1}) is ~20 times greater than for *p*NPG (1.4 ± 0.1 s^{-1} mM^{-1}) and ~450 times greater than for isomaltose (0.06 ± 0.01 s^{-1} mM^{-1}).

Regarding their substrate specificities against natural substrates, both ntMGAM and ntSI have the ability to hydrolyze the α-1,4 bond of maltose, but only ntSI is able to efficiently hydrolyze the α-1,6 bond of isomaltose. Kinetic results also reveal that *p*NPG is more efficiently cleaved and tightly bound to ntSI because its K_m is ~10-fold less (1.3 ± 0.1 mM) than that of ntMGAM (12.1 ± 1.0 mM).

Overall Structure of ntSI—The crystal structures of recombinant human ntSI in apo form and in complex with the inhibitor kotalanol were solved to 3.2 and 2.15 Å, respectively. Phases were solved by molecular replacement using the structure of ntMGAM (~60% sequence identity; Protein Data Bank code 2QLY) as a model. Due to its higher resolution, the ntSI-kotalanol structure will form the basis of the initial discussion. However, several insights from comparison with the apo-ntSI structure will also be addressed below. Crystal packing analysis of ntSI-kotalanol reveals that each asymmetric unit contains four ntSI molecules (monomers A, B, C, and D, each ~102 kDa) (Fig. 2*A*) with an average root mean square (r.m.s.) deviation of 0.29 ± 0.07 Å between monomer subunits. In all monomers, the N-terminal residues (aa 1–27 (monomer A), 1–29 (mono-

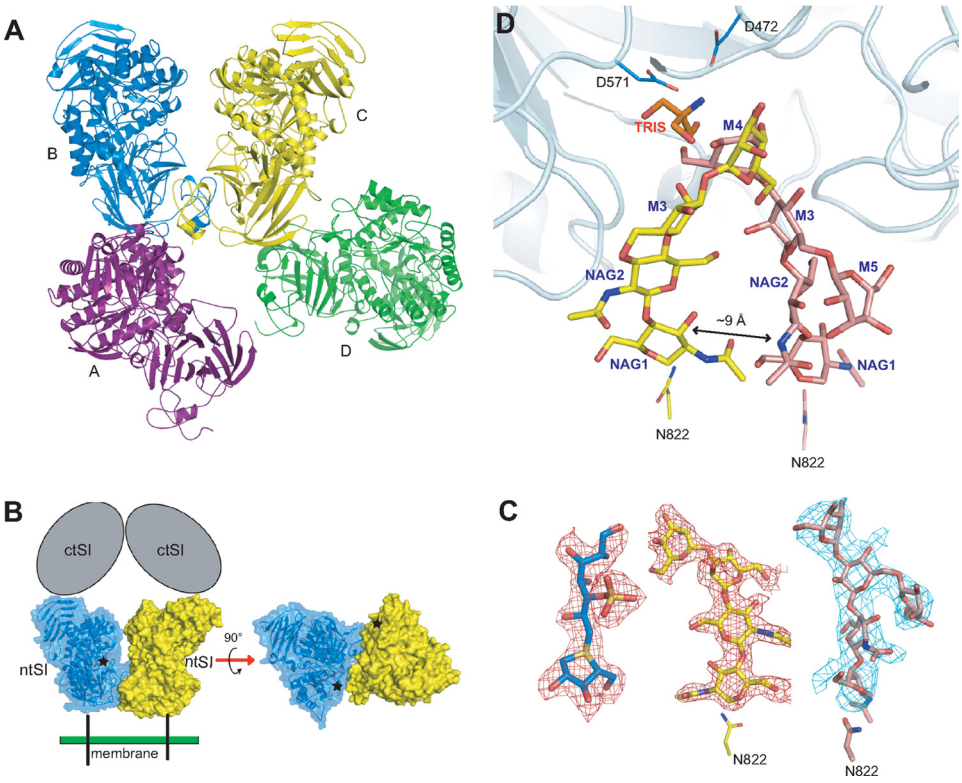

FIGURE 2. **Crystal structure and ligands of ntSI.** *A*, arrangement of ntSI monomers A, B, C, and D in the crystal asymmetric unit. *B*, interaction between monomer B (*blue*) and monomer C (*yellow*) observed in crystal-packing lattice and the proposed assembly of the full-length SI protein (composed of *O*-glycosylated stalk (*thick black line*) and ntSI and ctSI domains) with respect to the membrane. Approximate locations of the active site pockets are indicated by *. *C*, representative $F_o - F_c$ electron density maps of ntSI-kotalanol (shown in *red mesh*) with bound kotalanol (*blue*) and $Man_2GlcNAc_2$ (*yellow*) and of apo-ntSI (shown as *cyan mesh*) with bound $Man_3GlcNAc_2$ (*pink*). Maps are contoured at 2.5σ, and Asn^{822} residues from the ntSI-kotalanol (*yellow*) and apo-ntSI (*pink*) structures are shown attached to $Man_2GlcNAc_2$ and $Man_3GlcNAc_2$, respectively. *D*, active site superposition of apo-ntSI-monomer C with bound $Man_3GlcNAc_2$ glycan (*pink*) and the ntSI-kotalanol monomer A with bound $Man_2GlcNAc_2$ glycan (*yellow*) and Tris (*orange*). The Asn^{822} glycosylation sites attached to the glycan chains are shown as *thin pink* and *yellow sticks*, and the catalytic nucleophile Asp^{472} and acid/base catalyst Asp^{571} are indicated as *thin blue sticks*. A *double-headed arrow* indicates the distance (~9 Å) between the NAG1 rings of the glycan rings (labeled according to Fig. 1*B*).

mer B), 1–27 (monomer C), and 1–33 (monomer D)), including the His_6 tag and engineered TEV cleavage site are disordered. It is interesting to note that the removal of the disordered N-terminal region by TEV cleavage resulted in poorly diffracting crystals (>10 Å resolution).

The architectural fold of ntSI is identical to that of ntMGAM and the structures were found to superpose with an r.m.s. deviation of 0.5 Å over 776 Cα residues (supplemental Fig. 1). The largest structural deviations are localized to surface loops (*i.e.* ntSI aa 400–408 and ntMGAM aa 372–379) and the trefoil domain. Like ntMGAM, the ntSI monomer is divided into five subdomains, including 1) a type P trefoil subdomain (aa 29–80), 2) an N-terminal subdomain (aa 81–296), 3) a catalytic $(\beta/\alpha)_8$ barrel subdomain (aa 297–681) with variable inserts 1 and 2 (aa 395–445 and 476–521, respectively), 4) a proximal C-terminal subdomain (aa 682–759), and 5) a distal C-terminal subdomain (aa 760–898) (supplemental Fig. 2). Although the type P trefoil is only present in the N- and C-terminal intestinal MGAM and SI enzymes, the remaining folds are characteristic of GH31 enzymes (6, 26, 27). However, the N-terminal subdomains and catalytic inserts 1 and 2 can vary in length from one enzyme to another (6, 26).

Oligomeric Assembly—In studies with full-length SI (encompassing the transmembrane domain, *O*-glycosylation stalk, ntSI, and ctSI), homodimers have been observed in electron microscopy and sedimentation studies and are assembled with N-terminal domains interacting and C-terminal domains interacting with one another (28, 29). However, the observance of SI dimers seems to be highly dependent on experimental conditions because other cross-linking studies have suggested that SI exists as a monomer (30). In our studies, ntSI was found to elute as an active monomer during gel filtration chromatography (data not shown). However, an interesting interaction between monomers B and C was observed in the crystal-packing lattice (Fig. 2*B*). These monomers are related by a non-crystallographic 2-fold axis of symmetry and share an interface composed of residues from the trefoil subdomain and catalytic insert 2. This interaction is also observed between monomer A and a symmetry-related copy of monomer D. Analysis using the Protein Interfaces, Surfaces, and Assemblies (PISA) server (31) reveals that this interface has a buried surface area of ~800 $Å^2$/interface and contains extensive hydrophobic interactions and 10–12 hydrogen bonds. Although the dimer interface area is small for a protein of this size (32), it might represent a low affinity physiological interaction, as will be discussed below (see "Discussion").

ntSI Active Site—Prior to freezing and data collection, ntSI crystals were soaked in mother liquor supplemented with a 200 μM concentration of the inhibitor kotalanol. $F_o - F_c$ maps clearly reveal the presence of the kotalanol inhibitor bound to active sites of monomers B and D (Fig. 2*C*). In comparison, a Tris molecule (from the buffer used during purification) and a $Man_2GlcNAc_2$ glycan chain (originating from glycosylation site Asn^{822} of a crystal symmetry-related molecule) simultaneously occupy the active sites of monomers A and C (Fig. 2, *C* and *D*). It was noted that active sites of monomers A, B, C, and D are structurally identical to one another and are not affected by the kotalanol, Tris, or $Man_2GlcNAc_2$ binding. In addition, the Tris and glycan chains and the kotalanol molecules bind identically in the A and C monomers and the B and D monomers, respectively (data not shown).

The $Man_2GlcNAc_2$ glycan interacts with the ntSI active site primarily through the terminal M4 ring and a to a lesser extent with the NAG2 ring. These interactions include hydrogen

TABLE 1
Data collection and refinement statistics

	Apo-ntSI	ntSI-kotalanol
Crystal parameters		
Space group	$P2_12_12_1$	$P2_12_12_1$
Unit cell dimensions (Å)		
a	76.34	69.40
b	172.59	165.76
c	343.87	341.40
Data collection statistics[a]		
Resolution range (Å)	50–3.2	20–2.15
Reflections (unique)	76,095	20,2547
Redundancy	4.2 (4.1)	4.4 (3.1)
$I/\sigma I$	7.7 (2.6)	8.1 (2.5)
Completeness (%)	99.7 (100)	95.0 (84.6)
R_{sym}	0.18 (0.58)	0.17 (0.47)
Refinement statistics		
R/R_{free}	0.226/0.252	0.177/0.223
Maximum Resolution (Å)	3.2	2.15
Ramachandram plot (%)[b]	93.3/99.1/0.9	95.5/99.7/0.3
r.m.s. bonds (Å)	0.011	0.011
r.m.s. angles (degrees)	1.34	1.30
Estimated coordinate error (Å)[c]	0.51	0.20
Residue range built (monomer A/B/C/D)	29–898/29–898/29–898/29–898	28–898/30–898/28–898/34–898
Active site ligand(s)		
A	$Man_3GlcNAc_2$	$Man_2GlcNAc_2$ + Tris
B		Kotalanol
C	$Man_3GlcNAc_2$	$Man_2GlcNAc_2$ + Tris
D		Kotalanol
No. of water molecules	0	1727
Protein Data Bank codes	3LPO	3LPP

[a] Values in parentheses refer to the highest resolution shell.
[b] Residues in the favored/allowed/disallowed regions, as determined by MolProbity (24).
[c] Based on R_{free}, reported in REFMAC.

bonding between the side chain hydroxyls of Asp^{231} with the C3-OH and C4-OH groups of M4 and between the side chain of Asp^{571} and the C4-OH M4 group via a water molecule (supplemental Fig. 3). Additional stabilization of the $Man_2GlcNAc_2$ may result from the hydrophobic interactions with Leu^{233}, Trp^{327}, Trp^{435}, Phe^{479}, Val^{605}, and Tyr^{634}, which line the ntSI active site.

In a 3.2 Å resolution structure of apo-ntSI, a $Man_3GlcNAc_2$ chain attached to a symmetry-related Asn^{822} residue is observed in the active sites of monomers A and C, whereas monomers B and D remain empty. The electron density of the apo-ntSI structure was clear enough to position the five $Man_3GlcNAc_2$ rings within the density (Fig. 2*C*) but not detailed enough to assign hydrogen bonding interactions. The exact location of glycan binding, however, is not fixed and is dependent on the arrangement of monomers in the asymmetric unit. The structure of apo-ntSI has slightly different unit cell dimensions (76.34 × 172.59 × 343.87 Å) compared with the ntSI-kotalanol structure (69.4 × 165.8 × 341.4 Å) (Table 1), which results in a slight rearrangement of ntSI monomers in the asymmetric unit. Comparing the $Man_2GlcNAc_2$ binding in the 2.15 Å structure with the $Man_3GlcNAc_2$ binding in the 3.2 Å structure, the terminal mannose residues (M4) overlap, but the rest of the chain diverges from there, resulting in an ~9-Å separation between the *N*-linked NAG1 residues (Fig. 2*D*). Further detail about the $Man_2GlcNAc_2$ glycan, including its location in relations to the substrate binding subsites, will be discussed below in context with active site comparison with ntMGAM.

Comparison of Kotalanol Binding in ntSI and ntMGAM—To characterize the ntSI active site architecture relative to that of ntMGAM and to determine the extent of conservation in the −1 and +1 subsites, the active site of ntSI with bound kotalanol was superimposed with the recently solved structure of ntMGAM-kotalanol (Protein Data Bank code 3L4V). Although monomers B and D both contain kotalanol in their active sites, monomer B is of higher structural quality (lower *B* factors) and was used as the basis for structural comparison with ntMGAM.

Structural comparison of the ntMGAM and ntSI complexed structures reveals identical binding of the kotalanol ring group in the −1 subsite and the conservation of most of the −1 subsite residues (Fig. 3*A*). These residues include catalytic nucleophiles Asp^{443} and Asp^{472} and acid/base catalysts Asp^{542} and Asp^{571} in ntMGAM and ntSI, respectively. Substrate binding residues His^{600}, Asp^{327}, Arg^{526}, Asp^{203}, Trp^{406}, and Phe^{450} in ntMGAM are also structurally conserved with ntSI residues His^{629}, Asp^{355}, Arg^{555}, Asp^{231}, Trp^{435}, and Phe^{479}. The most obvious structural change in the −1 site is attributed to the substitution of ntMGAM Tyr^{299} with ntSI Trp^{327}, which results in the extension of the Trp side chain toward the +1 subsite (Fig. 3*A*).

In the +1 subsite, there are notable structural differences between ntMGAM and ntSI. These include the substitution of ntMGAM Thr^{205} with ntSI Leu^{233}, which positions the Leu side chain closer to the +1 subsite and the substitution of ntMGAM residue Ala^{570} by ntSI residue Val^{605} (Fig. 3*A*). There is also a 5.3-Å shift of the ntSI Lys^{509} side chain toward the +1 subsite compared with the side chain of ntMGAM Lys^{480}. These structural differences seem to affect the binding of the polyhydroxylated chain of kotalanol in the +1 subsite. The sulfate group in the ntSI-bound kotalanol is displaced by ~1–1.5 Å from its ntMGAM-bound position to accommodate the bulky Trp^{327} ntSI group. Additionally, the hydrogen bonding interaction between Thr^{205} and the C5′-OH group (via a water molecule) in the ntMGAM-bound kotalanol structure is lost in the ntSI-kotalanol structure. Instead, hydrogen bonding interactions

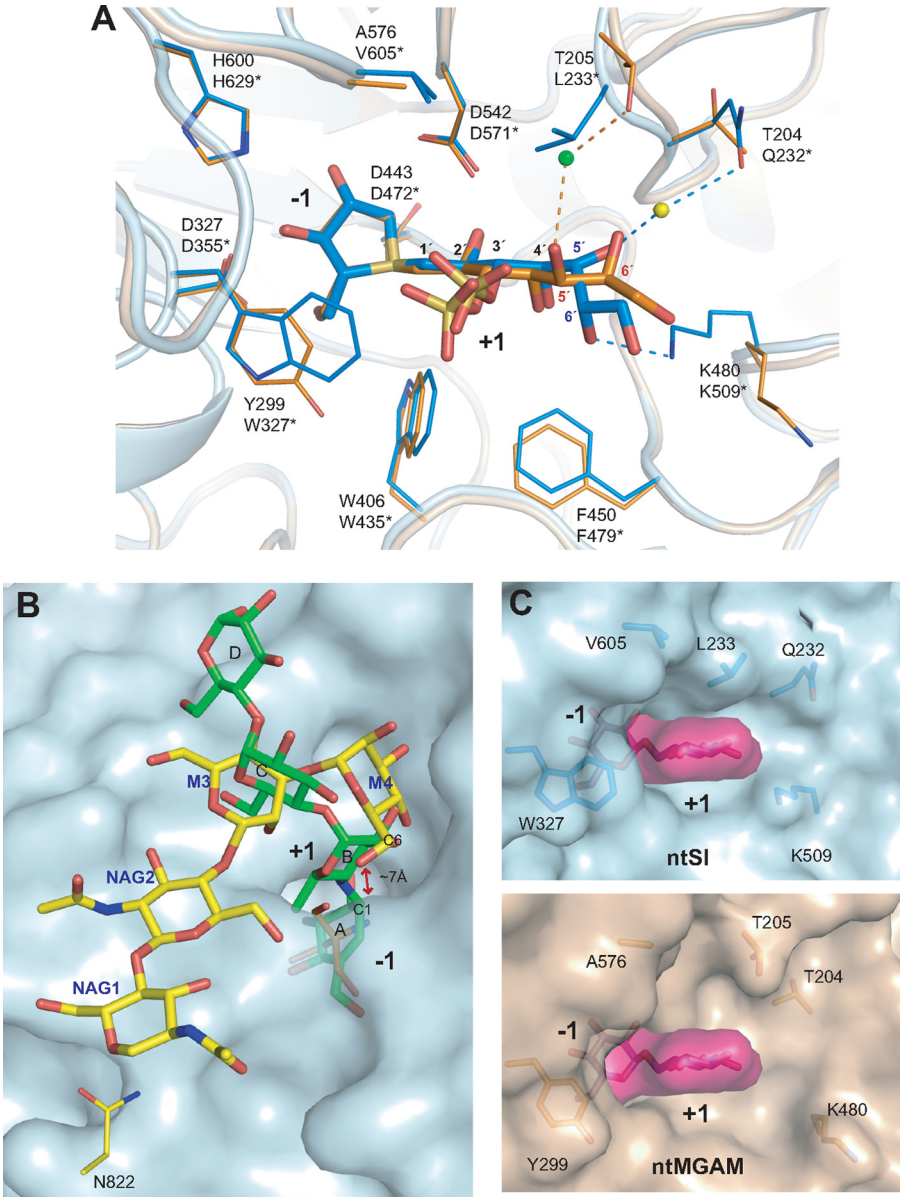

FIGURE 3. **Comparison of ntMGAM and ntSI active sites.** For *A–C*, the approximate locations of the −1 and +1 subsites are labeled. *A*, superposition of ntMGAM-kotalanol (*orange*) and ntSI-kotalanol (*blue*) active sites. Kotalanol and ntMGAM/ntSI residues are represented as *thick sticks* and *thin sticks*, respectively. Selected bound waters in the ntMGAM and ntSI structures are shown in *yellow* and *green*, respectively. ntMGAM residues are labeled with an *asterisk*. *B*, superposition of $Man_2GlcNAc_2$ (*yellow*) and Tris (*orange*) from ntSI-kotalanol with acarbose (*green*) from ntMGAM-acarbose in the ntSI active site (*blue surface representation*). A *double-headed arrow* indicates the distance (~7 Å) between the C6-OH of the M4 ring and C1 of acarbose ring A. *C*, *surface representation* of the ntSI (*top*) and ntMGAM (*bottom*) active sites with non-structurally conserved residues displayed as *blue* and *orange sticks*, respectively. One possible conformation of isomaltose (*pink*, *surface-rendered*) is modeled into the −1 and +1 subsites.

appear between Lys^{509} and C6′-OH and between Gln^{232} and C5′-OH (via a water molecule) (Fig. 3*A*).

Comparison of $Man_2GlcNAc_2$ and Acarbose Binding in ntSI and ntMGAM—As was described above, $Man_2GlcNAc_2$ and $Man_3GlcNAc_2$ Asn^{822}-linked glycans were found in the monomer A and C active sites in the 2.15 and 3.2 Å ntSI structures, respectively. Because the terminal M4 residues of both glycans overlap (Fig. 2*D*) and were found to interact with acid/base catalyst Asp^{571} and +1 subsite residue Asp^{231} (supplemental Fig. 3), we wanted to determine whether the M4 binding could represent a biologically relevant interaction occurring between substrate and subsite. Although there is no precedence for mannose binding in ntSI, it is a C2 epimer of glucose and thus very similar in structure. Because an ntSI crystal structure complexed with a substrate analogue or inhibitor, such as acarbose, has not been determined, we compared the ntSI structure with that of ntMGAM-acarbose (Protein Data Bank code 2QMJ) in order to map the location of the $Man_2GlcNAc_2$ binding in relation to the proposed −1 and +1 subsites.

Comparison of the structures reveals that acarbose and $Man_2GlcNAc_2$ groups bind differently from one another with essentially no overlap between rings (Fig. 3*B*). The ntMGAM −1 subsite is occupied by the non-reducing ring A of acarbose, whereas in ntSI, it is occupied by a Tris molecule. The binding of the terminal M4 group in $Man_2GlcNAc_2$ reveals that the ring is positioned ~4 Å away from the +1 subsite delineated by the B ring of acarbose and that its C6-OH arm is oriented toward the −1 subsite (Fig. 3*B*). We considered the possibility that the binding site occupied by M4 marks the +1 subsite for the reducing ring of isomaltose substrates. However, the distance between its C6-OH arm and the C1 group in acarbose ring A is too large (~7 Å) to accept an α-1,6 bond without a significant movement in M4.

DISCUSSION

In this study, we report the crystal structure of human ntSI. Its overall architectural fold is very similar to that of the paralogous enzyme ntMGAM (r.m.s. deviation = 0.5 Å). From crystal packing analysis, the largest interface between two ntSI monomers suggests a possible mode of dimerization mediated through the ntSI trefoil and catalytic insert 2 subdomains. This arrangement is consistent with the assembly of full-length SI, with *O*-glycosylation stalks stemming from ntSI toward the membrane and the continuation of the ctSI domains on the distal side away from the membrane attachment (Fig. 2*B*), similar to the assembly observed in electron microscopy studies (28). However, it is important to realize that this interaction occurs under the high protein concentrations required

TABLE 2
Kinetic parameters for α-1,4 and α-1,6 substrate hydrolysis by ntSI and ntMGAM

Substrate	ntSI			ntMGAM		
	K_m	k_{cat}	k_{cat}/K_m	K_m	k_{cat}	k_{cat}/K_m
	m_M	s^{-1}	$s^{-1}\ m_M^{-1}$	m_M	s^{-1}	$s^{-1}\ m_M^{-1}$
Maltose (α-1,4)	7.1 ± 1.3	137 ± 4	19 ± 4	4.3 ± 1.2	111 ± 10	26 ± 8
Isomaltose (α-1,6)	11.1 ± 1.3	97 ± 16	9 ± 2	227 ± 24	13 ± 1	0.06 ± 0.01
*p*NPG	1.3 ± 0.1	17 ± 1	13 ± 1	12.1 ± 1.0	17 ± 1	1.4 ± 0.1

for crystallization and that ntSI is observed only in the monomeric state during our purification studies. Nevertheless, because full-length SI dimers have been observed in other sedimentation and electron microscopy studies, it is possible that interactions between ntSI monomers are weak, at least under our experimental conditions and in the absence of membrane, *O*-glycosylation stalk, and C-terminal SI domain. The physiological relevance of a putative dimerization is still not clear, and further biochemical tests are needed to investigate the presence of this interface. However, many brush-border enzymes exist as dimers, and dimerization is believed to be relevant for intracellular trafficking (29). In any case, because we have observed active monomers in our kinetic studies, dimerization is not crucial for catalytic activity. Moreover, because the active site pockets on the ntSI monomers are exposed and are not composed of elements from adjacent monomers (Fig. 2*B*), we predict that they act independently of one another.

From kinetic studies, it was observed that maltose could be hydrolyzed by both ntMGAM and ntSI but that isomaltose could only be hydrolyzed efficiently by ntSI (Table 2). The differential substrate specificities of ntMGAM and ntSI seem consistent with their roles in terminal starch digestion; the redundancy in α-1,4 activity seen in both enzymes (also in ctMGAM and ctSI) is beneficial due the higher distribution of α-1,4 linkages (95%) in starch molecules. In contrast, α-1,6 linkages make up a small fraction of the linkages (5%) in starch molecules. Thus, it is logical that only one subunit (ntSI) is responsible for cleavage of this particular bond. To investigate the structural basis for differential substrate specificity observed between ntMGAM and ntSI, their active sites were compared.

Structural superposition of the ntMGAM and ntSI kotalanol bound structures revealed strong structural conservation of −1 subsite residues (Fig. 3*A*). This is quite expected because the −1 subsite binds the non-reducing ring of substrates, and this ring is identical in maltose and isomaltose (Fig. 1*A*). Substrate discrimination is most likely mediated by the +1 subsite, and there are accordingly structural differences observed between ntSI and ntMGAM in this region. A significant difference between ntMGAM and ntSI subsite architecture is that the ntMGAM +1 subsite is wide and open compared with the narrow groove observed in the ntSI +1 subsite (Fig. 3*C*). This narrowing results from the presence at analogous positions of smaller residues in ntMGAM (Tyr^{299}, Ala^{576}, Thr^{205}, and Thr^{204}) as compared with larger residues in ntSI (Trp^{327}, Val^{605}, Leu^{233}, and Gln^{232}) and the displacement of side chain Lys^{509} closer to the +1 subsite of ntSI. Sequence comparison of the four MGAM and SI domains reveals that Trp^{327} in ntSI may be important in conferring α-1,6 specificity because it is conserved as Tyr in the other ntMGAM, ctMGAM, and ctSI domains (6). However, it seems counterintuitive that the narrow +1 subsite of ntSI would accommodate both its α-1,4 and α-1,6 substrates, whereas a wider +1 subsite in ntMGAM would account for specificity for the α-1,4 substrate.

Unfortunately, there is currently no crystal structure of ntSI or ntMGAM in complex with an α-1,6-linked substrate or inhibitor analogue that can be used to predict isomaltose binding in ntSI. Instead, isomaltose was modeled by hand into the −1 subsites of ntMGAM and ntSI by aligning the non-reducing glucose ring of isomaltose to that of acarbose (using the ntMGAM-acarbose structure superposed on ntSI as a reference) (Fig. 3*C*). The non-reducing ring of isomaltose was then fixed in place while rotating torsion angles of the α-1,6 linkage to avoid steric clashes between the reducing ring and the +1 subsite residues. From this modeling exercise, it was evident that the number of conformations that the reducing isomaltose ring can adopt in the ntSI structure is limited due to narrowness of the +1 subsite. In comparison, the ntMGAM +1 architecture is wide and open, and supports the modeling of multiple conformations at the reducing isomaltose ring. Moreover, conformational studies have indicated that isomaltose is a more flexible disaccharide compared with maltose (33, 34). This is due to the three-bond α-1,6 linkage in isomaltose, in comparison with the two-bond α-1,4 linkage in maltose, which results in less interaction between reducing and non-reducing rings. We speculate that the narrow ntSI +1 subsite may be necessary for spatially constraining the flexible α-1,6 bond in place during substrate binding. Moreover, the presence of large hydrophobic residues Trp^{327}, Leu^{233}, and Val^{605}, in the +1 subsite, which are absent in ntMGAM, may help greatly in binding by interacting with hydrophobic patches on the reducing ring of isomaltose.

To gauge the hydrophobic environment of the +1 subsite, ntSI and ntMGAM were assayed with *p*NPG (Fig. 1*A*), a general glucosidase substrate that lacks hydrogen-bonding capability in its leaving group. With the assumption that the glucose moiety of *p*NPG binds identically in both ntMGAM and ntSI −1 subsites, any interactions with the +1 leaving group should be mostly hydrophobic. From these studies, *p*NPG was found to bind ~10 times tighter to ntSI compared with ntMGAM (Table 2), supporting the idea of a more hydrophobic +1 ntSI subsite. We propose that because the +1 subsite is spatially constricted and probably limits the variety of hydrogen bonding possibilities, nonspecific hydrophobic interactions contribute greatly toward binding of the reducing ring of substrates and broader substrate specificity of ntSI.

Moreover, although the binding of the $Man_2GlcNAc_2$ glycan to the ntSI active site (Figs. 2*D* and 3*B*) is not biologically relevant, the binding of the M4 mannose moiety close to the +1 subsite is a further indication of the nonspecificity of the ntSI

TABLE 3
Inhibition kinetics of ntMGAM and ntSI by α-glucosidase inhibitors

Inhibitor	K_i	
	ntSI	ntMGAM
	μM	
Acarbose	14 ± 1	62 ± 13[a]
Kotalanol	0.60 ± 0.06	0.19 ± 0.03[b]

[a] Reported in Ref. 10.
[b] Reported in Ref. 13.

substrate-binding site. However, for further validation of the hydrophobic environment of the +1 ntSI subsite and the entropic changes upon isomaltose binding, additional calorimetric studies as described in Ref. 35 will be pursued in future studies.

Inhibition studies with α-glucosidase inhibitors revealed that in comparison with acarbose, kotalanol is a ~20 times and ~300 times better inhibitor of ntSI and ntMGAM, respectively (Table 3). Structural comparison of kotalanol in the ntSI and ntMGAM active sites revealed similarity in binding, especially in the −1 subsite. Although there are differences in the binding of the polyhydroxylated chain of kotalanol in the +1 subsite, this does not seem to affect significantly its inhibition properties toward ntSI (K_i = 0.60 ± 0.06 μM) and ntMGAM (K_i = 0.19 ± 0.03 μM). This is most likely due to the structural conservation of the C2′ and C4′ chain hydroxyls (Fig. 3*A*). From previous inhibitor and structural studies, it was determined that stereochemistry of the C2′ and C4′ substituents is crucial for the activity of kotalanol because important hydrogen bonds are formed with the active site residues (11, 13).

The structural comparison between ntMGAM and ntSI provides preliminary insights into the structural basis for their complementary activities toward α-1,4- and α-1,6-linked substrates. The crystal structures will also be helpful in the design of specific inhibitors that exploit differences between MGAM and SI domains. Kotalanol was shown to be a specific inhibitor of both ntMGAM and ntSI, with no selectivity for either. However, removal of the essential hydrogen bonding interactions required for its specificity (through the removal of key kotalanol hydroxyls C2′-OH or C4′-OH) and the addition of a hydrophobic extension that could be stabilized by the narrow hydrophobic +1 ntSI subsite might result in a modified kotalanol inhibitor that is more specific for ntSI. The design of domain-specific inhibitors will greatly aid in the characterization of these domains and is of interest as a means of addressing ntSI and ntMGAM activities *in vivo*, in order to better define the complementary roles of MGAM and SI in the process of terminal starch digestion.

Acknowledgments—The pMT-TEVA vector was a generous gift from Dr. Peter L. Davies (Department of Biochemistry, Queens University, Kingston, Ontario, Canada). This work is based upon research conducted at the Cornell High Energy Synchrotron Source (CHESS), which is supported by the National Science Foundation (NSF) and NIGMS, National Institutes of Health (NIH), under NSF award DMR-0225180, using the Macromolecular Diffraction at CHESS (MacCHESS) facility, which is supported by National Center for Research Resources, National Institutes of Health Grant RR-01646.

REFERENCES

1. Van Beers, E. H., Büller, H. A., Grand, R. J., Einerhand, A. W., and Dekker, J. (1995) *Crit. Rev. Biochem. Mol. Biol.* **30,** 197–262
2. Semenza, G. (1986) *Ann. Rev. Cell Biol.* **2,** 255–313
3. Nichols, B. L., Avery, S., Sen, P., Swallow, D. M., Hahn, D., and Sterchi, E. (2003) *Proc. Natl. Acad. Sci. U.S.A.* **100,** 1432–1437
4. Quezada-Calvillo, R., Sim, L., Ao, Z., Hamaker, B. R., Quaroni, A., Brayer, G. D., Sterchi, E. E., Robayo-Torres, C. C., Rose, D. R., and Nichols, B. L. (2008) *J. Nutr.* **138,** 685–692
5. Heymann, H., Breitmeier, D., and Günther, S. (1995) *Biol. Chem. Hoppe Seyler* **376,** 249–253
6. Sim, L., Quezada-Calvillo, R., Sterchi, E. E., Nichols, B. L., and Rose, D. R. (2008) *J. Mol. Biol.* **375,** 782–792
7. Gray, G. M., Lally, B. C., and Conklin, K. A. (1979) *J. Biol. Chem.* **254,** 6038–6043
8. Asano, N. (2003) *Glycobiology* **13,** 93R–104R
9. Li, C., Begum, A., Numao, S., Park, K. H., Withers, S. G., and Brayer, G. D. (2005) *Biochemistry* **44,** 3347–3357
10. Rossi, E. J., Sim, L., Kuntz, D. A., Hahn, D., Johnston, B. D., Ghavami, A., Szczepina, M. G., Kumar, N. S., Sterchi, E. E., Nichols, B. L., Pinto, B. M., and Rose, D. R. (2006) *FEBS J.* **273,** 2673–2683
11. Mohan, S., and Pinto, B. M. (2007) *Carbohydr. Res.* **342,** 1551–1580
12. Mohan, S., and Pinto, B. M. (2009) *Collect. Czech. Chem. Commun.* **74,** 1117–1136
13. Sim, L., Jayakanthan, K., Mohan, S., Nasi, R., Johnston, B. D., Pinto, B. M., and Rose, D. R. (2010) *Biochemistry* **49,** 443–451
14. Ouwendijk, J., Moolenaar, C. E., Peters, W. J., Hollenberg, C. P., Ginsel, L. A., Fransen, J. A., and Naim, H. Y. (1996) *J. Clin. Invest.* **97,** 633–641
15. Scotter, A. J., Kuntz, D. A., Saul, M., Graham, L. A., Davies, P. L., and Rose, D. R. (2006) *Protein Expr. Purif.* **47,** 374–383
16. Heras, B., and Martin, J. L. (2005) *Acta Crystallogr. D Biol. Crystallogr.* **61,** 1173–1180
17. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* **276,** 307–326
18. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) *J. Appl. Crystallogr.* **40,** 658–674
19. Emsley, P., and Cowtan, K. (2004) *Acta Crystallogr. D Biol. Crystallogr.* **60,** 2126–2132
20. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta Crystallogr. D Biol. Crystallogr.* **53,** 240–255
21. Perrakis, A., Harkiolaki, M., Wilson, K. S., and Lamzin, V. S. (2001) *Acta Crystallogr. D Biol. Crystallogr.* **57,** 1445–1450
22. Winn, M. D., Isupov, M. N., and Murshudov, G. N. (2001) *Acta Crystallogr. D Biol. Crystallogr.* **57,** 122–133
23. Schüttelkopf, A. W., and van Aalten, D. M. (2004) *Acta Crystallogr. D Biol. Crystallogr.* **60,** 1355–1363
24. Davis, I. W., Leaver-Fay, A., Chen, V. B., Block, J. N., Kapral, G. J., Wang, X., Murray, L. W., Arendall, W. B., 3rd, Snoeyink, J., Richardson, J. S., and Richardson, D. C. (2007) *Nucleic Acids Res.* **35**, W375–W383
25. Dahlqvist, A. (1964) *Anal. Biochem.* **7,** 18–25
26. Ernst, H. A., Lo Leggio, L., Willemoës, M., Leonard, G., Blum, P., and Larsen, S. (2006) *J. Mol. Biol.* **358,** 1106–1124
27. Lovering, A. L., Lee, S. S., Kim, Y. W., Withers, S. G., and Strynadka, N. C. (2005) *J. Biol. Chem.* **280,** 2105–2115
28. Cowell, G. M., Tranum-Jensen, J., Sjöström, H., and Norén, O. (1986) *Biochem. J.* **237,** 455–461
29. Danielsen, E. M. (1994) *Biochemistry* **33,** 1599–1605
30. Jascur, T., Matter, K., and Hauri, H. P. (1991) *Biochemistry* **30,** 1908–1915
31. Krissinel, E., and Henrick, K. (2007) *J. Mol. Biol.* **372,** 774–797
32. Miller, S., Lesk, A. M., Janin, J., and Chothia, C. (1987) *Nature* **328,** 834–836
33. Coutinho, P. M., Dowd, M. K., and Reilly, P. J. (1997) *Carbohydr. Res.* **297,** 309–324
34. Dowd, M. K., Reilly, P. J., and French, A. D. (1994) *Biopolymers* **34,** 625–638
35. Sigurskjold, B. W., and Bundle, D. R. (1992) *J. Biol. Chem.* **267,** 8371–8376
36. Henrissat, B., and Davies, G. J. (1997) *Curr. Opin. Struct. Biol.* **7,** 637–644