Chemistry & Biology, Vol. 10, 619–628, July, 2003, ©2003 Elsevier Science Ltd. All rights reserved. DOI 10.1016/S1074-5521(03)00143-1

Anatomy of Glycosynthesis: Structure and Kinetics of the *Humicola insolens* Cel7B E197A and E197S Glycosynthase Mutants

Valérie M-A. Ducros,[1] Chris A. Tarling,[2] David L. Zechel,[2,4] A. Marek Brzozowski,[1] Torben P. Frandsen,[3,5] Ingemar von Ossowski,[3] Martin Schülein,[3,6] Stephen G. Withers,[2] and Gideon J. Davies[1,*]

[1]Structural Biology Laboratory
Department of Chemistry
The University of York
Heslington
York YO10 5YW
United Kingdom

[2]Department of Chemistry
The University of British Columbia
2036 Main Mall
Vancouver V6T 1Z1
British Columbia
Canada

[3]Novozymes A/S
Smoermosevej 25
DK-2880 Bagsvaerd
Denmark

Summary

The formation of glycoconjugates and oligosaccharides remains one of the most challenging chemical syntheses. Chemo-enzymatic routes using retaining glycosidases have been successfully harnessed but require tight kinetic or thermodynamic control. "Glycosynthases," specifically engineered glycosidases that catalyze the formation of glycosidic bonds from glycosyl donor and acceptor alcohol, are an emerging range of synthetic tools in which catalytic nucleophile mutants are harnessed together with glycosyl fluoride donors to generate powerful and versatile catalysts. Here we present the structural and kinetic dissection of the *Humicola insolens* Cel7B glycosynthases in which the nucleophile of the wild-type enzyme is mutated to alanine and serine (E197A and E197S). 3-D structures reveal the acceptor and donor subsites and the basis for substrate inhibition. Kinetic analysis shows that the E197S mutant is considerably more active than the corresponding alanine mutant due to a 40-fold increase in k_{cat}.

Introduction

The synthesis of glycosidic bonds, catalyzed in nature by glycosyltransferases, is central to many biological processes. Simply in terms of quantity, it is the most significant reaction in the biosphere. Carbohydrates and polysaccharides govern a diverse range of cellular functions, including energy storage, cell wall structure, cell-cell interactions and signaling, host-pathogen interactions, and protein glycosylation [1–6] Because these functions, especially those in which carbohydrate moieties act as the cellular language, rely on precise carbohydrate structures that often display an extreme chemical diversity, the biosynthesis of oligosaccharides and polysaccharides may involve the action of hundreds of different and selective glycosyltransferases.

Our ability to investigate, harness, and intervene in carbohydrate-specific processes depends on the availability of specific oligosaccharides. Despite recent advances, the stereo- and regiospecific synthesis of the glycosidic bond still remains a significant chemical challenge. Chemo-enzymatic synthesis of oligosaccharides harnessing recombinant glycosyltransferases is a powerful route to some of these complex products, albeit one limited by the poor availability of the often fragile enzymes [7]. The harnessing of transglycosylation reactions catalyzed by the more robust retaining glycosidases is also a route that has been frequently used (Figure 1A) [8, 9]. Here, the naturally formed covalent glycosyl-enzyme intermediate is intercepted with an acceptor alcohol, resulting in synthesis via transglycosylation. The product is, however, necessarily also a substrate for the hydrolytic reaction, and tight kinetic or thermodynamic control is frequently necessary even for relatively modest yields.

In 1998, a powerful alternative chemo-enzymatic route to carbohydrate synthesis was reported [10]. Site-directed mutation of the enzymatic nucleophilic carboxylate of a retaining *exo*-glycosidase, initially to alanine, generated an enzyme that was hydrolytically inert. When incubated with α-glucosyl fluoride, which mimics the high-energy covalent intermediate (Figure 1A), these mutant glycosidases are able to synthesize product in high yield (Figure 1B). Since the development of the original *Agrobacterium* sp. β-glucosidase mutant, successful "glycosynthases" have been developed on a range of templates from both *endo*- and *exo*-acting β-glycosidases (for example, [6, 11, 12–17]). Recently, the first "α-glycosynthase," which synthesis α-1,4 and α-1,6 bonds, has been described [18], marking a significant expansion of the glycosynthase repertoire.

The *Humicola insolens* endoglucanase Cel7B is a retaining β-glycoside hydrolase that serves a biological role as part of the secreted cellulolytic apparatus of the fungus. As such, its substrates are β-1,4 linked oligo and polysaccharides of D-glucopyranose. The 3-D structure has been determined (Figure 2) and the catalytic nucleophile identified as Glu197 through trapping with 2-deoxy-2-fluorodisaccharides and subsequent analysis by proteolysis and ESI MS/MS tandem mass spectrometry [19]. The Cel7B E197A mutant generates a versatile glycosynthase whose preliminary application to the synthesis of a variety of oligosaccharides [12], including novel bifunctionalized substrates [16], has been described.

Recently, the glycosynthase repertoire has been ex-

*Correspondence davies@ysbl.york.ac.uk
[4]Present address: Biochemishes Institut, Universität Zürich, Winterthurerstrasse 190, CH-8057, Zürich, Switzerland.
[5]Present address: PANTHECO A/S, Danish Science Park, Bøge Alle 3, DK 2970 Hørsholm, Denmark.
[6]Martin Schülein passed away in 2001 and is greatly missed.

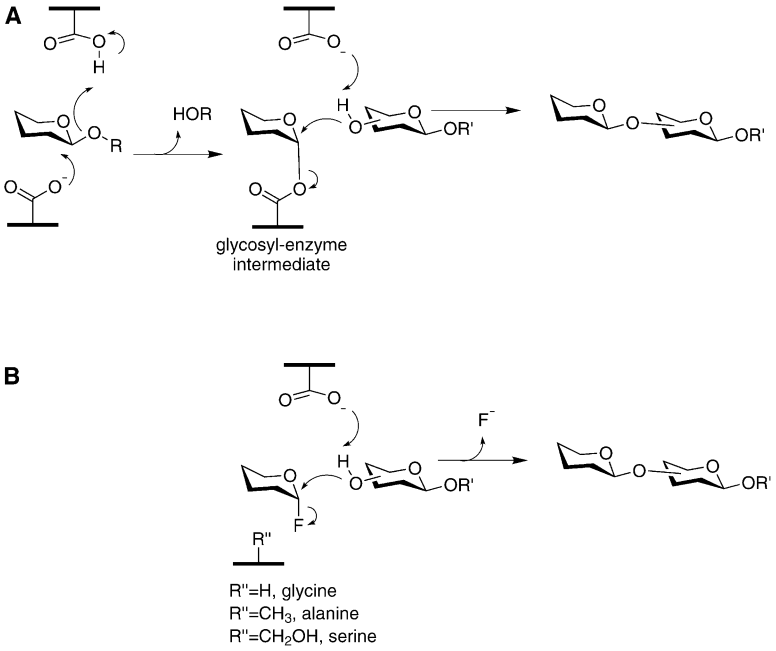

Figure 1. Enzyme-Catalyzed Glycosidic Bond Formation

(A) The transglycosylation reaction catalyzed by retaining glycosidases in which a covalent glycosyl-enzyme intermediate is intercepted by the acceptor alcohol (in competition with water); and (B) the reaction catalyzed by "glycosynthases" whose enzyme-derived nucleophilic carboxylate has been mutated to glycine, serine, or alanine.

panded through substitution of the catalytic nucleophile with residues other than alanine. Serine and glycine mutations generate glycosynthases whose prowess frequently outstrips that provided by the original alanine mutant [15, 20–22]. Improved glycosynthases not only act more rapidly, but as a direct consequence, they are able to transfer to a much wider array of acceptors on a useful timescale, making them considerably more versatile synthetic tools. In the case of the *Agrobacterium* β-glucosidase enzyme, it was proposed that the E358S mutant generates a more powerful enzyme as the serine hydroxyl could hydrogen bond to the departing fluoride atom [20]; indeed, a similar role could be envisaged for solvent water in the glycine glycosynthase variants [23]. Here we present a dissection of the *H. insolens* Cel7B glycosynthases at kinetic and structural levels. The E197S glycosynthase variant displays markedly better catalytic properties than the E197A mutant. The 3-D structure reveals the locations and interactions in both donor (−1, −2) and acceptor (+1, +2) subsites (subsite nomenclature according to [6]), as well as revealing the molecular basis for substrate inhibition. Cel7B E197S exclusively forms β-1-4 linkages with cellobiosyl or lactosyl fluoride donors, yet inspection of the structure reveals that this regioselectivity reflects fine structural tuning in which the acceptor 4-OH is, surprisingly, not the closest atom to the anomeric carbon of the donor. We conclude that both distance and angle criteria are

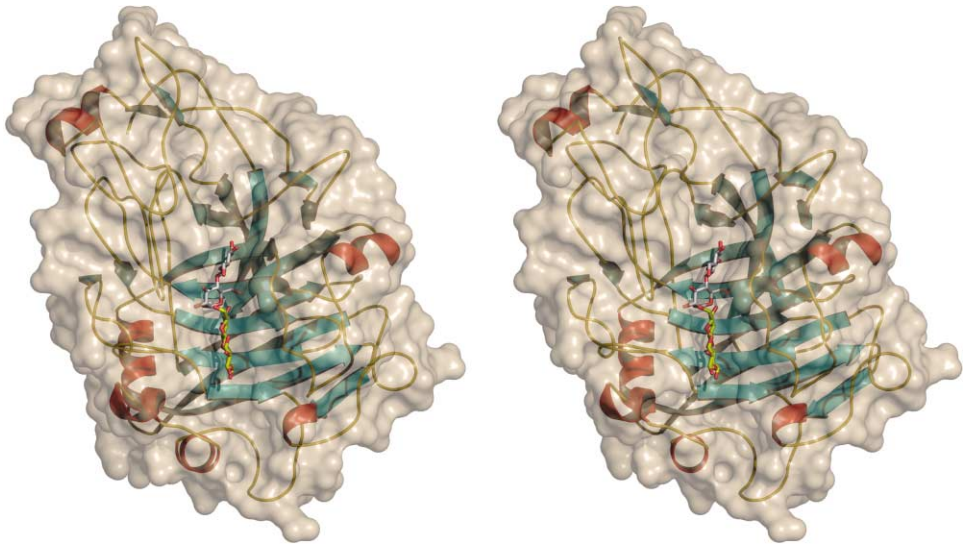

Figure 2. 3-D Structure of the Cellobiose Complex of the *Humicola insolens* Cel7B (E197S) Glycosynthase

The protein topology is shown (helices, red; sheets, blue; coil, yellow) together with the molecular surface (wheat). Donor (−2, −1) sugars are shown in gray, and acceptor (+1, +2) in yellow "licorice". The figure is in divergent ("wall-eyed") stereo.

Table 1. Enzyme Kinetics for Cel7B Nucleophile Mutant Catalyzed Condensation of Lactosyl Fluoride with PNP-Cellobioside

Mutation	k_{cat} (min^{-1})[a]	$K_{M(LacF)}$ (mM)	$K_{M(PNPC)}$	$k_{cat}/K_{M(LacF)}$
E197A	24.2 ± 0.6	0.87 ± 0.05	7.87 ± 0.61	27.8
E197S	1080 ± 23	1.12 ± 0.13	2.26 ± 0.16	964

[a] Apparent k_{cat} derived from fixed LacF donor concentration of 10 mM α-LacF with varying [PNPC].

crucial components of enzyme specificity, opening up possibilities for future broadening of enzyme acceptor range through mutagenesis approaches.

Results and Discussion

The E197A glycosynthase mutant of the *H. insolens* Cel7B had previously been used in conjunction with α-lactosyl fluoride (α-LacF) donor and a variety of acceptors to produce oligosaccharides in 51%–100% yield [12]. Use of α-cellobiosyl fluoride led to polymerization and subsequent precipitation of "cello" (β-1,4 linked) oligosaccharides while the 3-D structure of the free enzyme was determined at 1.75 Å resolution (PDB Code 1DYM). In order to generate a more powerful enzyme with improved kinetics and acceptor range, by analogy with work on the *Agrobacterium* Abg β-glucosidase [20] and the *Cellulomonas fimi* Man2A mannosidase [15], the E197S mutant was constructed and analyzed by kinetic and structural methods.

Reaction Kinetics: E197S versus E197A Glycosynthases

The catalytic constants for the condensation of α-LacF with *p*-nitrophenyl β-cellobioside (PNPC) were determined by monitoring the release of fluoride with a fluoride ion selective electrode. Since no transfer is observed to an acceptor with an axial 4-hydroxyl group, only a single transfer event is assayed, that from α-LacF to the PNPC. On the timescale of the experiment, neither E197S nor E197A catalyzed the transfer of α-LacF to water; no above-background release of fluoride was detected in the absence of acceptor, and the reaction ceased following stoichiometric transfer of donor under conditions of limiting acceptor, all demonstrating that the mutants do not catalyze transfer to water in a significant manner. The E197A mutant transfers α-LacF to PNPC with a k_{cat} of 24.2 min^{-1}, K_M (PNPC) of 7.9 mM, and K_M (α-LacF) of 0.87 mM, with a corresponding catalytic efficiency ($k_{cat}/K_{M(LacF)}$) of 27.8 $min^{-1}mM^{-1}$ (Figure 3A; Table 1). Substrate inhibition is observed at high donor concentrations (Figure 3b) with a K_i of 173 ± 17 mM, presumably reflecting binding of α-LacF to the acceptor subsites, as revealed by X-ray crystallography, below.

The E197S mutant is substantially faster than its alanine counterpart. While K_M values remain similar (Table 1; K_M [PNPC] 2.3 mM, K_M [α-LacF] 1.1 mM), k_{cat} increases to 1080 min^{-1} with a corresponding $k_{cat}/K_{M(LacF)}$ of 964 min^{-1} mM^{-1}. This represents a 34-fold increase in catalytic efficiency for the serine glycosynthase mutant, a similar improvement to that previously observed for both the Abg and Man2A-derived glycosynthases. In the case of the Abg enzyme, the improvement in efficiency appears to be derived solely from K_M [20], whereas with Cel7B, as with Man2A [15], this improvement stems almost exclusively from k_{cat}, which in the case of E197S is increased some 44-fold.

3-D Structure of the Cel7B E197S Glycosynthase and Complexes with Cellobiose and Lactose

The structure of free Cel7B E197S was determined by molecular replacement from a hexagonal crystal form at a resolution of 2.15 Å (see Experimental Procedures for details). The structure refined with R_{cryst} and R_{free} of 0.18 and 0.22, respectively (Table 2), and was essentially isomorphous with the free E197A and wild-type enzymes, indicating that the mutation had no significant effect on the 3-D structure of the protein.

Cel7B presents a long, open, active-center "groove" with sufficient space to accommodate approximately eight monosaccharide binding sites, arranged −5 to +3 [6], but maximal hydrolytic activity of the wild-type enzyme is achieved on reduced cellopentaoside substrates, reflecting the significant contribution of four subsites (−2 to +2) to hydrolysis. As seen with many other carbohydrate-protein interactions, the enzyme provides a series of hydrophobic platforms for pyranoside binding, notably Trp347 and Tyr171 in the −2 subsite, Tyr147 in the −1 subsite, and Trp356 in +1. The equivalent hydrophobic platform in the +2 subsite is provided by the aliphatic portion (CA-CB-CG) of the side-chain of Asn237. At the center of this subsite binding canyon lies, in the wild-type enzyme, the catalytic nucleophile Glu197. In the glycosynthase variants, Glu197 is replaced with alanine or serine with no disruption to the overall 3-D structure; the mutants are "isomorphous" with the wild-type protein. In order to dissect the protein-ligand interactions, the E197S variant was studied in complex with both lactose and cellobiose. Relatively poor diffraction from the hexagonal crystal form led us to harness a better diffracting form for complex studies.

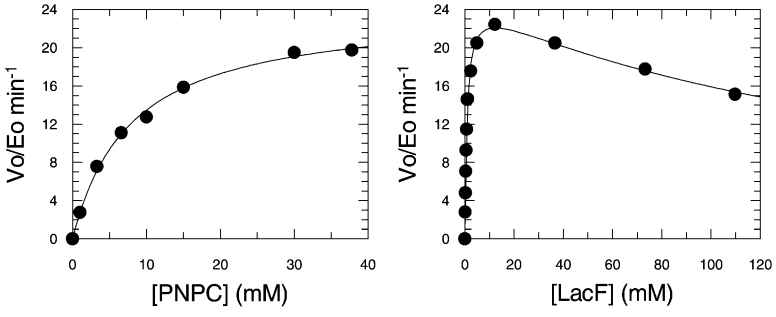

Figure 3. Reaction Kinetics of Cel7B Glycosynthase

(A) PNPC acceptor kinetics and (B) LacF donor kinetics for the Cel7B E197A mutant (see Table 1 for full details).

Table 2. Data and Structure Quality for the *Humicola insolens* Cel7B E197S Glycosynthase and Its Complexes

	E197S Uncomplexed	E197S Cellobiose	E197S Lactose
Data			
Radiation source	ESRF, ID14-4	ESRF, ID14-1	ESRF, ID14-1
Space group	$P6_5$	$P2_1$	$P2_1$
Cell dimensions	a = b = 122.3 Å, c = 82.62 Å	a = 66.3 Å, b = 74.8 Å, c = 86.1 Å, β = 102.6°	a = 66.4 Å, b = 74.8 Å, c = 85.8 Å, β = 102.5°
Number mols. in A.U.	1	2	2
Resolution (Å)	15-2.15 (2.23-2.15)	30-1.5 (1.55-1.5)	15-1.4 (1.45-1.4)
Completeness (%)	100 (99.9)	99.3 (97.4)	99 (96)
R_{merge}	0.09 (0.35)	0.043 (0.32)	0.033 (0.30)
Multiplicity	8.9 (8.7)	3.5 (3.6)	3.6 (3.2)
$I/\sigma I$	21 (7)	24 (3.8)	35 (3.5)
Refinement			
R_{cryst}	0.18	0.16	0.15
R_{free}	0.22	0.19	0.17
Rms on bond distances (Å)	0.018	0.010	0.011
Rms on bond angles (°)	1.68	1.56	1.58
Mean B protein ($Å^2$)	25	16	14
Number waters (mean B/$Å^2$)	330 (30)	793 (22)	642 (21)
Mean B ligand $Å^2$ (−2,−1/+1,+2 subsites)	N/A	9,10/20,26 (Amol) 10,10,16,21 (Bmol)	10,10/30,36 (Amol) 11,10/25,33 (Bmol)
PDB code	1OJI	1OJK	1OJJ

Outer resolution bin statistics are given in parentheses.

A monoclinic crystal form of Cel7B (E197S) ($P2_1$; approximate cell dimensions, a = 66 Å, b = 75 Å, c = 86 Å, β = 103°) diffracting to beyond 1.4 Å resolution proved amenable to ligand binding studies. Synchrotron data were collected to 1.5 Å on Cel7B (E197S) in complex with cellobiose and to 1.4 Å in complex with lactose (Table 2). Both cellobiose and lactose complexes display essentially identical interactions in the donor −2 and −1 subsites. Both −1 and −2 glycosyl moieties bind in their undistorted $^4C_1$ (chair) conformations and are well-ordered, as reflected in low temperature factors of around 10 $Å^2$ for the −2/−1 subsite units of the cellobiose and lactose complexes, respectively (Table 2).

There is no steric hindrance to lactose binding, consistent with the prowess of lactosyl fluoride as a donor with K_M values around 1 mM. Both lactose and cellobiose bind as their α-anomers (Figure 4), mimicking the α-glycosyl fluorides of the glycosynthase reaction. The −2 and −1 subsite interactions feature hydrophobic "stacking" with Trp347 and Tyr147, and in the −1 subsite the 6, 3, and 2 hydroxyls hydrogen-bond to Trp347 NE1, Asp173 OD2, and Gln175 NE2, respectively. This latter interaction, between the amide hydrogen and the 2-OH, is reminiscent of that seen (and known to be important for catalysis) in the structurally unrelated glycosidases from clan "GH-A." Consistent with the "*syn*" protonation trajectory [24], the catalytic acid/base Glu202 interacts with the 6-OH of the −1 subsite glucosyl moiety. All these interactions of the −1 and −2 subsites are essentially identical to that described for the distorted thio-oligosaccharide "Michaelis" complex of the related Cel7B from *Fusarium oxysporum* [25]. The E197S mutant generates sufficient space to accommodate the axial anomeric hydroxyl of an α-glycoside, yet the serine hydroxyl does not make a direct interaction with anomeric hydroxyl, instead lying some 3.8 Å distant. The implications of this for catalysis are discussed below.

In the acceptor subsites, lactoside and cellobioside complexes are not equivalent. Steric factors, perhaps the potential clash between the axial *galacto* 4-OH of lactose with the side-chain of Gln175, prevent lactose binding in the "true" +1 and +2 subsites. Instead, it is rotated slightly and displaced approximately 1.7 Å "out" of the active center, spanning the "+1.5 to +2.5" subsites (Figure 4C). This explains, in part, why lactosides are not acceptor substrates for Cel7B. Given that lactose in this position blocks productive binding in these acceptor subsites, it is extremely likely that this binding mode is responsible for the substrate inhibition ($K_i \sim 170$ mM) observed at high donor concentrations. That the binding is comparatively weak is also reflected in poor electron density (Figure 4B) and consequent temperature factors around 30–40 $Å^2$ for the displaced lactose, compared to 20–25 $Å^2$ for cellobiose bound in the productive +1 and +2 subsites.

Catalysis by Cel7B E197S Glycosynthase

The acceptor and donor site interactions of Cel7B (E197S) with cellobiose occupying the −2/−1 and +1/+2 subsites cast provocative new light on glycosynthase catalysis. The Cel7B glycosynthase forms β-1,4 bonds exclusively using lactosyl and cellobiosyl donors and *gluco*-configured acceptors (regio-selectivity is characterized in [12]), a reaction which demands catalytic base assistance from Glu202. In the Cel7B E197S cellobiose complex, however, the acceptor O4 is not the closest residue to Glu202; indeed, it lies as much as 3.9 Å from this group compared to 3.3 Å for O3, an atom which does not act as an acceptor nucleophile. Furthermore, O4 lies 4.4 Å from the anomeric C1 of the donor, greater than the sum of their van der Waals' contacts (compared to 3.3 Å for the O3 position), and again this looks both suboptimal for catalysis and more suggestive of β-1,3 bond formation (not observed) [12]. One possibility is that the observed position of the acceptor substrate in the crystal structure is not representative of the initial

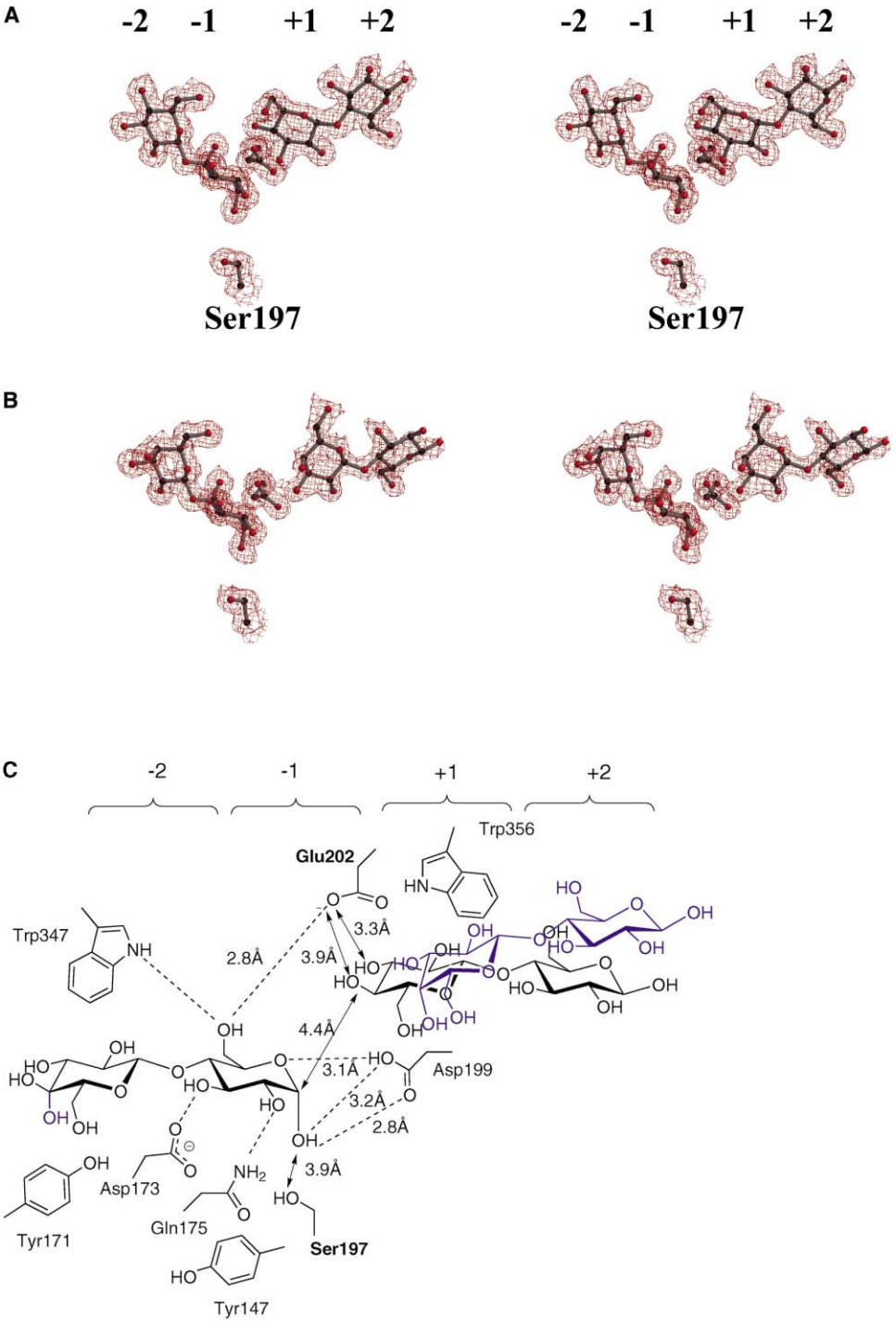

Figure 4. Observed Electron Density and Interactions for the Cel7B (E197S)

(A) Cellobiose and (B) lactose complexes. Electron density shown is a REFMAC maximum likelihood-weighted 2Fobs-Fcalc synthesis contoured at approximately 0.47 and 0.44 electrons/Å^3, respectively. The figures are shown in divergent (wall-eyed) stereo. (C) Schematic diagram of the Cel7B (E197S)-ligand interactions. Only the −1 subsite interactions (from the cellobiose complex) are given in their entirety, and the approximate position of the lactose moieties is shown for reference in blue.

encounter complex between Cel7B and its substrates, although both donor and acceptor positions are similar to the trapped "Michaelis" complex with nonhydrolysable thio-oligosaccharide for the related *F. oxysporum* Cel7B (formally EG1) [25], which would represent an early, distorted, product complex in the direction of synthesis. The donor positions and interactions between these two complexes are similar, save the −1 subsite $^4C_1 \rightarrow {}^1S_3$ distortion of the thio-oligosaccharide complex. In the +1 subsite, the planes of the two glucosyl moieties

of cellobiose are identical, although the glucosyl moiety of the *F. oxysporum* Cel7B thio-oligosaccharide complex is approximately 0.5 Å closer to the catalytic acid than the equivalent glucoside of the Cel7B E197S complex. Indeed, even in the thio-oligosaccharide complex with the *F. oxysporum* Cel7B, the closest +1 subsite atom to both $C1_{donor}$ and Glu202 (acid/base) is the O3 and not the O4 of the acceptor. Given these counterintuitive distances, it would appear that regio-selectivity for 1,4 bond formation appears to stem not solely from distance criteria, but from the angle of nucleophilic attack. The O4…C1-O1 (virtual) bond angle is 173°, close to the 180° required for in-line nucleophilic substitution. While O3 is closer to both the acid/base and C1 of the donor, its potential angle of attack is considerably less optimal than that possible for O4.

In the donor site, where the serine glycosynthase mutant is approximately 35 times more efficient, there is no direct interaction between the serine hydroxyl and the axial O1 of the donor. If the extra benefit of the serine mutant does come from the potential of a stabilizing interaction between departing fluoride and serine hydroxyl, as has been proposed [20], then this feature may be expressed solely at the transition state, which will indeed be of a highly dissociative character and may well place the fluoride closer to the serine hydroxyl. Alternatively, the position of the glycosyl fluoride may not be appropriately mimicked by the cellobiosyl moiety. Cel7B (E197S) appears to be the most efficient glycosynthase described to-date, with a catalytic efficiency around 1000 $min^{-1}mM^{-1}$; but even with the benefit of structural hindsight, it is difficult to dissect reactivity contributions that must be expressed at the transition state as opposed to the ground state. Additional kinetic assistance to an axial leaving group of the donor might also be provided by Asp199 in Cel7B, a group likely to be appropriately protonated given its interaction with the O5 atom of the pyranoside ring in these glycosynthase complexes and its normal hydrogen bond donation to the carboxylate nucleophile of the wild-type family GH-7 enzymes. Intriguingly, this aspartate completes, along with the nucleophile and acid/base, a cluster of three carboxylates also found in similar location in family GH-16, which provides the powerful *Bacillus lichiniformis* β-1,3-1,4 glucanase-derived glycosynthase [11]. Given these interactions, it is pertinent to consider what makes a good glycosynthase and whether one can predict such activity a priori?

What Makes an Efficient Glycosynthase?

It has been observed a number of times that simple replacement of the catalytic nucleophile of a retaining glycosidase, particularly if the mutation is to alanine, does not always produce a glycosynthase. This is true even when the wild-type enzyme shows strong transglycosylating properties. The reasons are unclear, but some insight into the glycosynthase potential of a wild-type retaining glycosidase may be gleaned by examination of the kinetic parameters for turnover of the glycosyl-enzyme intermediate, most easily achieved following trapping as the 2-fluoro glycosyl enzyme. Reactivation of the trapped covalent intermediate, which may be regarded as a model of the glycosynthase bound α-glycosyl fluoride, can occur by transfer of the 2-fluoro glycosyl moiety to a competing acceptor glycoside (k_{trans}) (Figure 1A) or by hydrolysis (k_{H2O}). Table 3 lists reactivation parameters for a number of retaining glycosidases and the results of glycosynthase conversions attempted thus far. Although this is only a partial survey, glycosidases that yield active glycosynthases appear to have two reactivation characteristics: high rate constants for transfer to an acceptor ($k_{trans} \geq 10^{-2}\ min^{-1}$) and also high selectivity for transfer to acceptor over water ($k_{trans}/k_{H2O} > 20$).

Abg and Cel7B clearly display these two characteristics for reactivation, and both enzymes were converted to effective glycosynthases upon mutating the nucleophiles to alanine. The *Streptomyces lividans* CelB [38] and *Cellulomonas fimi* Man2A [15] are borderline cases, for while the absolute rates for transfer to acceptors are high ($>10^{-2}\ min^{-1}$), the selectivities for transfer to an acceptor over water are relatively modest (k_{trans}/k_{H2O} = 10–20). The corresponding alanine nucleophile mutants of Man2A and CelB display weak glycosynthase activity, but in both cases the serine mutants are more effective catalysts. In contrast, wild-type *Thermosporum saccharolyticum* β-xylosidase [39] and *Bacillus circulans* β-galactosidase [40] have impressive transglycosylation activities, but conversion to the alanine or serine nucleophile mutants was unsuccessful in producing glycosynthase (D. Vocadlo and S.G.W., unpublished data). In the case of the xylosidase, high selectivity is observed with disaccharide acceptors (k_{trans}/k_{H2O} = 80 with xylobiose), but the low rate constants for transfer to an acceptor (k_{react} = 10^{-5} to $10^{-3}\ min^{-1}$) appear to legislate against glycosynthase activity. This example raises an unavoidable question. How are impressive transglycosylation qualities in the wild-type enzyme lost upon trapping the intermediate as a 2-fluoro glycosyl enzyme or mutating the nucleophile to alanine? This may be a reflection of the relative reactivities of the glycosyl donor in each case (glycosyl enzyme, 2-fluoro-glycosyl enzyme, and α-glycosyl fluoride). Another feature evident is the high binding constants measured for many acceptors. Yet, despite poor binding, these acceptors often greatly accelerate turnover of the 2-fluoro-glycosyl enzyme intermediates. Indeed, the turnover of the Abg intermediate increases nearly 3000-fold with PNP β-glucoside as acceptor. The important role of the acceptor at the transition state is also responsible for the reduced transfer efficiency to water displayed by these mutants. Potential hydrolytic reactions in glycosynthases resulting from direct attack of water on the glycosyl fluoride are sufficiently compromised that they are insignificant compared to the preferred transglycosylation reaction. The latter process is facilitated by the binding interactions with the acceptor sugar in the +1 site in much the same way that binding interactions at that site are important for the cleavage of chemically unreactive oligosaccharides. These interactions result in the stabilization of the transition state for glycosyl transfer.

Significance

Chemical glycobiology is one of the most rapidly expanding fields of modern science, yet such work and

Table 3. Kinetic Parameters for Reactivation of 2-Fluoro Glycosyl-Enzyme Intermediates

Glycosidase (GH Family)	Inactivator	k_{H2O} (min^{-1})	Acceptor	k_{trans} (min^{-1})	K_{trans} (mM)	k_{trans}/k_{H2O}	Nucleophile Mutant (Glycosynthase) Activity[b]	Reference
Abg (1)	DNP 2F-β-glucoside[a]	1×10^{-5}	β-glucosyl benzene	5×10^{-3}	59	500	E358A ++, E358S ++++, E358G +++++	[20, 21, 26]
Abg (1)	DNP 2F-β-glucoside	1×10^{-5}	PNPGlu (20 mM)	2.7×10^{-2}	–	2700		[20, 21, 26]
Abg (1)	2F-β-mannosyl fluoride	1×10^{-3}	β-glucosyl benzene	6×10^{-3}	64	6		[20, 21, 26]
Abg (1)	DNP 2F-β-galactoside	5×10^{-3}	β-glucosyl benzene	0.355	69	71		[20, 21, 26]
E. coli, lacZ β-galactosidase (2)	DNP 2F-β-galactoside	2.7×10^{-4}	glucose	4.8×10^{-3}	460	18	E537S +, E357S/G794D ++	[22, 27, 28]
Human β-glucoronidase (2)	2F-β-glucuronic acid fluoride	6.7×10^{-5}	chitobiose (50 mM)	1.2×10^{-4}	–	2	No data	[29]
Cellulomonas fimi Man2A (2)	2F-β-mannosyl fluoride	2×10^{-3}	gentiobiose	4×10^{-2}	78	20	E519A +, E519S ++	[15, 30]
Candida albicans endoglucanase (5)	DNP 2F-β-glucoside	1.9×10^{-3}	benzyl thio-β-glucoside	0.024	56	13	No data	[31]
Humicola insolens Cel7B (7)	DNP 2F-β-cellobioside	3×10^{-4}	cellobiose (20 mM)	1.5×10^{-2}	–	50	E197A +++, E197S ++++++	[12, 19], this study
Fusarium oxysporum Cel7B (7)	DNP 2F-β-cellobioside	1.1×10^{-4}	cellobiose (19.5 mM)	4.8×10^{-3}	–	44	No data	[32]
Cellulomonas fimi xylanase (Cex) (10)	DNP 2F-β-cellobioside	8.5×10^{-6}	cellobiose (55 mM)	1.9×10^{-5}	–	2.2	Ala inactive	[33]
Bacillus circulans xylanase (Bcx) (11)	DNP 2F-β-xylobioside	2.1×10^{-3}	benzyl thio β-xylobioside	3×10^{-2}	46	14	Ala inactive	[34]
Streptomyces lividans CelB (12)	DNP 2F-β-cellobioside	2.2×10^{-3}	cellobiose	2.5×10^{-2}	54	11	E120G inactive, E120A +, E120S ++	[35]
Xanthomonas manihotis β-galactosidase (35)	DNP 2F-β-galactoside I	3×10^{-4}	phenyl β-Glc (31 mM)	1.2×10^{-3}	–	4	Ala inactive Ser inactive	[36] (S.G.W. and J. Blanchard, unpublished data)
Bacillus circulans β-galactosidase (35)	DNP 2F-β-galactoside	2×10^{-5}	2-deoxy-glucose (34 mM)	5.5×10^{-5}	–	3	Ala inactive Ser inactive	[36]
Thermoanaerobacterium saccharolyticum β-xylosidase (39)	DNP 2F-β-xyloside	1×10^{-5}	xylose (45 mM)	5×10^{-5}	–	5	Ala inactive Ser inactive	[37] (S.G.W. and D. Vocadlo, unpublished data)
Thermoanaerobacterium saccharolyticum β-xylosidase (39)			xylobiose (45 mM)	8×10^{-4}	–	80		

[a] DNP = 2,4-dinitrophenyl

[b] Plus signs (+) are a qualitative measure of glycosynthase activity and not directly proportional.

biomedical investigation require access to large quantities of specifically designed and synthesized oligosaccharides. Enzyme-catalyzed synthesis using "glycosynthase" mutants provides a powerful tool for oligosaccharide synthesis. The 3-D structure and catalytic dissection of the *H. insolens* Cel7B enzyme reveals the fine structural tuning that contributes to regio-selectivity, opening up the possibilities for future tailoring of enzyme specificity through additional mutation. Wider evidence suggests that the qualities of a "good" acceptor appear to stem from kinetic (or reactivity) effects that are manifested primarily in the transition state, rather than ground state binding. Given that the structure of a good acceptor is difficult to predict, a future challenge is to develop "acceptor screens" for high-yielding transglycosylation reactions [41]. Such approaches will also be necessary to discover new glycosynthase mutants [21, 22] that raise the reactivity of the α-glycosyl fluoride donor to that of the wild-type glycosyl-enzyme intermediate and allow full harnessing of engineered variants for glycosidic bond synthesis on a large scale. Structural analyses, in combination with kinetic data, will inform future engineering of glycosynthases that expand the synthetic repertoire and stimulate research in glycobiology.

Experimental Procedures

Mutagenesis and Protein Production

The construction and purification of the E197A mutant has been described elsewhere [12]. For the E197S mutant, the mutational changes (GAG→TCG) necessary to introduce the E197S substitution in *Humicola insolens* Cel7B were produced by the overlap extension PCR (OE-PCR) method [42, 43]. The plasmid pHW704eg1 served as the DNA template for the OE-PCR mutagenesis. This plasmid is an *Escherichia coli-Aspergillus* shuttle expression vector that carries the *H. insolens* Cel7B coding region, the DNA sequence for its own secretion signal peptide, as well as an *Aspergillus* α-amylase promoter and glucoamylase terminator for transcriptional control of the wild-type Cel7B gene [44]. Two pairs of oligonucleotide primers, 5′-GAGGGCAAGGGCTCGTGCTGCAACTCGATGGATATCTGGGAG (Ser codon underlined) and 5′-CCCCATCCTTTAACTATAGCGAAA TGG as well as 5′-CGACAACATCACATCAAGCTCTCC and 5′-GCA GCACGAGCCCTTGCCCTCG, were used to generate an OE-PCR fragment that was first cleaved with BamHI and BstEII restriction endonucleases and then reinserted into the pHW704eg1 vector devoid of the wild-type Cel7B gene. After DNA sequencing had verified the mutational changes, the E197S mutant plasmid and an acetamidase selection plasmid (pTOC202) were transformed into *Aspergillus oryzae* JaL228 as described earlier [45]. Secreted E197S mutant protein was recovered from the spent fermentation broths and then purified by column chromatography essentially as described [12]. Standard protocols were employed in all DNA manipulations as described by Sambrook and colleagues [46].

Kinetics

Reaction kinetics were performed using α-lactosyl fluoride (α-LacF) as donor (prepared by fluorination of per-*O*-acetylated lactose with HF/pyridine according to Jünnemann and colleagues [47] followed by deprotection with catalytic sodium methoxide in methanol) and *p*-nitrophenyl β-cellobioside (PNPC) as acceptor. Assays (500 μl) were performed in 150 mM sodium phosphate (pH 7.0). Fluoride detection was performed using an Orion fluoride ion selective electrode (model 96-09BN) interfaced with a Fischer Scientific Accumet 925 pH/ion meter, essentially as described previously [20]. Enzymatic rates were corrected for spontaneous hydrolysis of α-LacF. Initial studies used a fixed (20.6 mM) concentration of PNPC in order to determine a rate profile for α-LacF and to determine a concentration for α-LacF that exhibited minimal substrate inhibition (10 mM). Subsequent experiments used a fixed 10 mM concentration of α-LacF, with [PNPC] varying between 1 and 37.8 mM. For donor kinetics, [PNPC] was fixed at 20.6 mM, with [α-LacF] varying between 0.1 and 122 mM. Curve fitting, taking into account substrate inhibition where appropriate, was performed with GraFit 5.0 [48].

Crystallization, Data Collection, and Processing

Both hexagonal rods and monoclinic plates of Cel7B grow from similar conditions in 2–4 μl hanging drops with 20 mg/ml protein, 20 mM TRIS-HCl buffer (pH 7–8.5), and 15%–30% polyethylene glycol 4000 as precipitant. The hexagonal form grows as single rods, but diffracts more weakly. The monoclinic form, used for complex studies, diffracts to higher resolution but frequently grew as clusters of plates, and inclusion of 5% DMSO helped obtain more single plates. Complex structures were obtained by soaking crystals in 50 mM of the appropriate ligand for 1 hr prior to cryo-storage. Single crystals were frozen with the inclusion of 20% v/v glycerol as cryoprotectant, and data was collected on ESRF ID14-EH1 beamline (Mar CCD detector) for the lactose and cellobiose soaks and ID14-EH4 (ADSC Quantum 4 CCD detector) for the E197S uncomplexed rods. All data were processed and reduced using the HKL suite of programs [49].

Refinement

The different forms of Cel7B were solved by molecular replacement using the program AMORE from the CCP4 suite [50] with an outer radius of Patterson integration of 25 Å and data between 20 and 3.5 Å. The E197A mutant (PDB code 1dym.pdb) was used as the search model. All further computing was performed using the CCP4 suite unless otherwise stated. For the refinement of each structure, 5% of the observations were immediately set aside for cross validation analysis [51] and were used to monitor various refinement strategies such as the weighting of geometrical and temperature factor restraint and the insertion of solvent water during maximum likelihood refinement using REFMAC program [52]. Manual corrections of the model using the X-FIT routines of the program QUANTA (Accelrys, San Diego, CA. USA) were interspersed with cycles of maximum likelihood refinement. "Riding" hydrogen atoms were included for structures with data beyond 1.5 Å resolution and only when their positions were definable. Figure 2 was drawn with PyMOL [53] and Figure 4 with BOBSCRIPT [54].

Acknowledgments

This work was supported by the Wellcome Trust and by the provision of a Peter Wall Institute for Advanced Studies "catalytic visitor" award to G.J.D. The Protein Engineering Network of Centres of Excellence (PENCE) and the Biotechnology and Biological Sciences Research Council are thanked for funding. G.J.D is a Royal Society University Research Fellow.

Received: April 4, 2003
Revised: May 13, 2003
Accepted: May 14, 2003
Published: July 18, 2003

References

1. Dwek, R.A. (1996). Glycobiology: toward understanding the function of sugars. Chem. Rev. *96*, 683–720.
2. Bertozzi, C.R., and Kiessling, L.L. (2001). Chemical glycobiology. Science *291*, 2357–2364.
3. Hart, G.W., Gao, Y., Wells, L., Comer, F.I., Iyer, S., Zachara, N., Vosseller, K., Li, Z.Y., and Kamemura, K. (2001). Glycobiology: a road less traveled. Glycobiology *11*, 1.
4. Wells, L., Vosseller, K., and Hart, G.W. (2001). Glycosylation of nucleocytoplasmic proteins: signal transduction and O-GlcNAc. Science *291*, 2376–2378.
5. Wells, L., Gao, Y., Mahoney, J.A., Vosseller, K., Chen, C., Rosen, A., and Hart, G.W. (2002). Dynamic O-glycosylation of nuclear and cytosolic proteins: further characterization of the nucleocy-

toplasmic beta-N-acetylglucosaminidase, O-GlcNAcase. J. Biol. Chem. *277*, 1755–1761.
6. Davies, G.J., Wilson, K.S., and Henrissat, B. (1997). Nomenclature for sugar-binding subsites in glycosyl hydrolases. Biochem. J. *321*, 557–559.
7. Palcic, M.M. (1999). Biocatalytic synthesis of oligosaccharides. Curr. Opin. Biotechnol. *10*, 616–624.
8. Crout, H.G., and Vic, G. (1998). Glycosidases and glycosyl transferases in glycoside and oligosaccharide synthesis. Curr. Opin. Chem. Biol. *2*, 98–111.
9. Vocadlo, D.J., and Withers, S.G. (2000). Glycosidase catalysed oligosaccharide synthesis. In Carbohydrates in Chemistry and Biology, Volume 2, B. Ernst, G.W. Hart, and P. Sinay, eds. (Weinheim, Germany: Wiley-VCH GmbH), pp. 723–844.
10. Mackenzie, L.F., Wang, Q., Warren, R.A.J., and Withers, S.G. (1998). Glycosynthases: mutant glycosidases for oligosaccharide synthesis. J. Am. Chem. Soc. *120*, 5583–5584.
11. Malet, C., and Planas, A. (1998). From β-glucanase to β-glucansynthase: glycosyl transfer to α-glycosyl fluorides by a mutant endoglucanase lacking its catalytic nucleophile. FEBS Lett. *440*, 208–212.
12. Fort, S., Boyer, V., Greffe, L., Davies, G.J., Moroz, O., Christiansen, L., Schülein, M., Cottaz, S., and Driguez, H. (2000). Highly efficient synthesis of β(1 4)-oligo- and -polysaccharides using a mutant cellulase. J. Am. Chem. Soc. *122*, 5429–5437.
13. Trincone, A., Perugino, G., Rossi, M., and Moracci, M. (2000). A novel thermophilic glycosynthase that effects branching glycosylation. Bioorg. Med. Chem. Lett. *10*, 365–368.
14. Faijes, M., Fairweather, J.K., Driguez, H., and Planas, A. (2001). Oligosaccharide synthesis by coupled endo-glycosynthases of different specificity: a straightforward preparation of two mixed-linkage hexasaccharide substrates of 1,3/1,4-beta-glucanases. Chemistry *7*, 4651–4655.
15. Nashiru, O., Zechel, D.L., Stoll, D., Mohammadzadeh, T., Warren, R.A.J., and Withers, S.G. (2001). β-Mannosynthase: synthesis of β-mannoside linkages with a mutant β-mannosidase. Angew. Chem. Int. Ed. Engl. *40*, 417–419.
16. Boyer, V., Fort, S., Frandsen, T.P., Schulein, M., Cottaz, S., and Driguez, H. (2002). Chemoenzymatic synthesis of a bifunctionalized cellohexaoside as a specific substrate for the sensitive assay of cellulase by fluorescence quenching. Chemistry *8*, 1389–1394.
17. Hrmova, M., Imai, T., Rutten, S.J., Fairweather, J.K., Pelosi, L., Bulone, V., Driguez, H., and Fincher, G.B. (2002). Mutated barley (1,3)-beta-D-glucan endohydrolases synthesize crystalline (1,3)-beta-D-glucans. J. Biol. Chem. *277*, 30102–30111.
18. Okuyama, M., Mori, H., Watanabe, K., Kimura, A., and Chiba, S. (2002). alpha-glucosidase mutant catalyzes "alpha-glycosynthase"-type reaction. Biosci. Biotechnol. Biochem. *66*, 928–933.
19. Mackenzie, L.F., Sulzenbacher, G., Divne, C., Jones, T.A., Woldike, H.F., Schülein, M., Withers, S.G., and Davies, G.J. (1998). Crystal structure of the family 7 endoglucanase I (Cel7B) from Humicola insolens at 2.2 Å resolution and identification of the catalytic nucleophile by trapping of the covalent glycosyl-enzyme intermediate. Biochem. J. *335*, 409–416.
20. Mayer, C., Zechel, D.L., Reid, S.P., Warren, R.A.J., and Withers, S.G. (2000). The E358S mutant of Agrobacterium sp. β-glucosidase is a greatly improved glycosynthase. FEBS Lett. *466*, 40–44.
21. Mayer, C., Jakeman, D.L., Mah, M., Karjala, G., Gal, L., Warren, R.A.J., and Withers, S.G. (2001). Screening for new and improved glycosynthases. Chem. Biol. *8*, 437–443.
22. Jakeman, D.L., and Withers, S.G. (2002). On expanding the repertoire of glycosynthases: mutant β-galactosidases forming β-(1,6)-linkages. Can. J. Chem. *80*, 866–870.
23. Jahn, M., Stoll, D., Warren, R.A.J., Szabó, L., Singh, P., Gilbert, H.J., Ducros, V.M.-A., Davies, G.J., and Withers, S.G. (2003). Expansion of the glycosynthase repertoire to produce defined manno-oligosaccharides. Chem. Commun., 1327–1329.
24. Heightman, T.D., and Vasella, A.T. (1999). Recent insights into inhibition, structure, and mechanism of configuration-retaining glycosidases. Angew. Chem. Int. Ed. Engl. *38*, 750–770.
25. Sulzenbacher, G., Driguez, H., Henrissat, B., Schulein, M., and Davies, G.J. (1996). Structure of the fusarium oxysporum endoglucanase i with a nonhydrolyzable substrate analogue: substrate distortion gives rise to the preferred axial orientation for the leaving group. Biochemistry *35*, 15280–15287.
26. Street, I.P., Kempton, J.B., and Withers, S.G. (1992). Inactivation of a β-glucosidase through the accumulation of a stable 2-deoxy-2-fluoro-α-D-glucopyranosyl-enzyme intermediate: a detailed investigation. Biochemistry *31*, 9970–9978.
27. Gebler, J.C., Aebersold, R., and Withers, S.G. (1992). Glu-537, Not Glu-461, is the nucleophile in the active site of (lac Z) β-galactosidase from *E. coli*. J. Biol. Chem. *267*, 11126–11130.
28. McCarter, J., Adam, M., and Withers, S.G. (1992). Binding energy and catalysis: fluorinated and deoxygenated glycosides as mechanistic probes of *Escherichia coli* (lac Z) β-galactosidase. Biochem. J. *286*, 721–727.
29. Wong, A.W., He, S., Grubb, J.H., Sly, W.S., and Withers, S.G. (1998). Identification of Glu-540 as the catalytic nucleophile of human beta-glucuronidase using electrospray mass spectrometry. J. Biol. Chem. *273*, 34057–34062.
30. Stoll, D., He, S., Withers, S.G., and Warren, R.A.J. (2000). Identification of Glu-519 as the catalytic nucleophile in β-mannosidase 2A from *Cellulomonas fimi*. Biochem. J. *351*, 833–838.
31. Mackenzie, L.F., Brooke, G.S., Cutfield, J.F., Sullivan, P.A., and Withers, S.G. (1997). Identification of Glu-330 as the catalytic nucleophile of *Candida albicans* exo-beta-(1,3)-glucanase. J. Biol. Chem. *272*, 3161–3167.
32. Mackenzie, L.F., Davies, G.J., Schulein, M., and Withers, S.G. (1997). Identification of the catalytic nucleophile of endoglucanase I from *Fusarium oxysporum* by mass spectrometry. Biochemistry *36*, 5893–5901.
33. Tull, D., and Withers, S.G. (1994). Mechanisms of cellulases and xylanases: a detailed kinetic study of the exo-beta-1,4-glycanase from *Cellulomonas fimi*. Biochemistry *33*, 6363–6370.
34. Miao, S., Ziser, L., Aebersold, R., and Withers, S.G. (1994). Identification of glutamic acid 78 as the active site nucleophile in *Bacillus subtilis* xylanase using electrospray tandem mass spectrometry. Biochemistry *33*, 7027–7032.
35. Zechel, D.L., He, S., Dupont, C., and Withers, S.G. (1998). Identification of Glu-120 as the catalytic nucleophile in *Streptomyces lividans* endoglucanase CelB. Biochem. J. *336*, 139–145.
36. Blanchard, J.E., Gal, L., He, S., Foisy, J., Warren, R.A., and Withers, S.G. (2001). The identification of the catalytic nucleophiles of two beta-galactosidases from glycoside hydrolase family 35. Carbohydr. Res. *333*, 7–17.
37. Vocadlo, D.J., Mackenzie, L.F., He, S., Zeikus, G.J., and Withers, S.G. (1998). Identification of Glu-277 as the catalytic nucleophile of *Thermoanaerobacterium saccharolyticum* β-xylosidase using electrospray MS. Biochem. J. *335*, 449–455.
38. Zechel, D.L. (2001). Observation and catalytic promiscuity of retaining glycosidase intermediates. PhD thesis, University of British Columbia, Vancouver, British Columbia.
39. Armand, S., Vieille, C., Gey, C., Heyraud, A., Zeikus, J.G., and Henrissat, B. (1996). Stereochemical course and reaction products of the action of beta-xylosidase from *Thermoanaerobacterium saccharolyticum* strain B6A-RI. Eur. J. Biochem. *236*, 706–713.
40. Fujimoto, H., Miyasato, M., Ito, Y., Sasaki, T., and Ajisaka, K. (1998). Purification and properties of recombinant beta-galactosidase from Bacillus circulans. Glycoconj. J. *15*, 155–160.
41. Blanchard, J.E., and Withers, S.G. (2001). Rapid screening of the aglycone specificity of glycosidases: applications to enzymatic synthesis of oligosaccharides. Chem. Biol. *8*, 627–633.
42. Ho, S.N., Hunt, H.D., Horton, R.M., Pullen, J.K., and Pease, L.R. (1989). Site-directed mutagenesis by overlap extension using the polymerase chain reaction. Gene *77*, 51–59.
43. Higuchi, R., Krummel, B., and Saiki, R.K. (1988). A general method of in vitro preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions. Nucleic Acids Res. *16*, 7351–7367.
44. Christensen, T., Wöldike, H., Boel, E., Mortensen, S.B., Hjortshøj, K., Thim, L., and Hansen, M.T. (1988). High level expression of recombinant genes in *Aspergillus oryzae*. Biotechnology *6*, 1419–1422.
45. Kelly, J.M., and Hynes, M.J. (1985). Transformation of Aspergil-

lus niger by the amdS gene of *Aspergillus nidulans*. EMBO J. *4*, 475–479.
46. Sambrook, J., Fritsch, E., and Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual, 2nd Edition (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press).
47. Jünnemann, J., Thiem, J., and Pedersen, C. (1993). Facile synthesis of acetylated glycosyl fluorides derived from di- and trisaccharides. Carbohydr. Res. *249*, 91–94.
48. Leatherbarrow, R.J. (2001). GraFit version 5.0 (London: Erithacus Software, Ltd).
49. Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. *276*, 307–326.
50. CCP4 (Collaborative Computational Project Number 4) (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D *50*, 760–763.
51. Brünger, A.T. (1992). Free R-value: a novel statistical quantity for assessing the accuracy of crystal-structures. Nature *355*, 472–475.
52. Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. D *53*, 240–255.
53. DeLano, W.L. The PyMOL Molecular Graphics System (San Carlos, CA: DeLano Scientific LLC) http://www.pymol.org.
54. Esnouf, R.M. (1999). Further additions to MolScript version 1.4, including reading and contouring of electron-density maps. Acta Crystallogr. D *55*, 938–940.

Accession Numbers

Coordinates have been deposited with the Protein Data Bank via the Molecular Structures Database at URL http://www.ebi.ac.uk/msd/ under codes 1OJI, 1OJJ, and 1OJK.