Elimination of competing hydrolysis and coupling side reactions of a cyclodextrin glucanotransferase by directed evolution

Ronan M Kelly, Hans Leemhuis, Henriëtte J Rozeboom, Niels van Oosterwijk, Bauke W. Dijkstra, Lubbert Dijkhuizen

Elimination of competing hydrolysis and coupling side reactions of a cyclodextrin glucanotransferase by directed evolution

Ronan M. Kelly[1], Hans Leemhuis[1], Henriëtte J. Rozeboom[2], Niels van Oosterwijk[2], Bauke W. Dijkstra[2] and Lubbert Dijkhuizen[1]*

[1]Microbial Physiology, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Kerklaan 30, 9751 NN Haren, The Netherlands, and Centre for Carbohydrate Bioprocessing, TNO-University of Groningen, Kerklaan 30, 9751 NN Haren, The Netherlands.

[2]Laboratory of Biophysical Chemistry, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands.

Running head: Removing enzyme side reactions.

* Corresponding author. Tel: 31-50-3632150; Fax: 31-50-3632154;
Email: L.Dijkhuizen@rug.nl

[1] Abbreviations: CGTase, cyclodextrin glucanotransferase; epPCR, error-prone polymerase chain reaction; GH13, glycoside hydrolase family 13; MBS, maltose binding site; Tabium, *Thermoanaerobacterium thermosulfurigenes* strain EM1; pNPG7, 4-nitrophenyl-α-D-maltoheptaoside-4-6-O-ethylidene.

Thermoanaerobacterium thermosulfurigenes cyclodextrin glucanotransferase primarily catalyzes the formation of cyclic α-(1,4)-linked oligosaccharides (cyclodextrins) from starch. This enzyme also possesses unusually high hydrolytic activity as a side reaction, thought to be due to partial retention of ancestral enzyme function. This side reaction is undesirable since it produces short saccharides that are responsible for the breakdown of the cyclodextrins formed, thus limiting the yield of cyclodextrins produced. To reduce the competing hydrolysis reaction, while maintaining the cyclization activity, we applied directed evolution, introducing random mutations throughout the *cgt* gene by error-prone PCR. Mutations in two residues, Ser-77 and Trp-239, on the outer region of the active site, lowered the hydrolytic activity up to 15-fold with retention of cyclization activity. In contrast, mutations within the active site could not lower hydrolytic rates, indicating an evolutionary optimized role for cyclodextrin

formation by residues within this region. The crystal structure of the most effective mutant, S77P, showed no alterations to the peptide backbone. However, subtle conformational changes to the side chains of active site residues had occurred, which may explain the increased cyclization/hydrolysis ratio. This indicates that secondary effects of mutations located on the outer regions of the catalytic site are required to lower the rates of competing side reactions, whilst maintaining primary catalytic function. Subsequent functional analysis of various glucanotransferases from the superfamily of glycoside hydrolases also suggests a gradual evolutionary progression of these enzymes from a common 'intermediate-like' ancestor toward specific transglycosylation activity.

Keywords: protein evolution; amylase; CGTase, protein stability; side reactions; saccharide

Introduction

Enzymes are essentially a product of evolution. Continuous evolution of enzymes results in intermediates partially retaining the initial ancestral function while catalysing new function [1]. Over several evolutionary phases, aided by selective pressure, new enzyme function dominates while ancestral function decreases to a mere side reaction [2-4]. This evolutionary phenomenon is clearly evident among members of the glycoside hydrolase family 13 (GH13), retaining a multitude of minor side reaction specificities and activities [5]. Members of GH13 act on starch, glycogen and related oligo- and polysaccharides, representing the largest family of glycoside hydrolases. Their active site regions share high sequence similarity embedded in a TIM $(\beta/\alpha)_8$ barrel structural fold [6], indicating evolutionary diversification from a common enzyme ancestor. All family members either hydrolyze or transglycosylate α-glucosidic linkages to produce α-anomeric mono- and oligo-saccharides; catalysis proceeds via a double displacement mechanism with initial formation of a covalent glycosyl intermediate [7]. It is therefore the type of acceptor substrate utilised in the second half of the reaction which determines enzyme reaction specificity, a water molecule in the case of α-amylases and a hydroxyl group of a sugar substrate for the glucanotransferases. Despite highly specific primary activities of GH13 enzymes, many glucanotransferase enzymes also catalyze hydrolysis as a side reaction. Extensive analysis of cyclodextrin glucanotransferases (CGTase) revealed that this side reaction has a considerable impact on final product yields. CGTases primarily

catalyze the glucanotransferase reaction of cyclization, for the formation of circular α-(1,4)-linked oligosaccharides (cyclodextrins) (Fig. 1) [8]. These enzymes also exhibit, to a lesser extent, α-amylase-like activity, hydrolysing starch into short linear saccharides (Fig. 1). The build-up of short saccharides from starch hydrolysis promotes yet another CGTase side reaction, coupling (Fig. 1). In the coupling reaction cyclodextrin rings are cleaved and transferred to (short) saccharide molecules to yield linear products [9]. While most CGTases generally exhibit low hydrolytic rates, *Thermoanaerobacterium thermosulfurigenes* strain EM1 (Tabium) CGTase displays unusually high hydrolytic activity, but still low compared to α-amylases [10].

By determining the structural/molecular factors responsible for the high hydrolytic rates of Tabium CGTase, we aimed at significantly reducing this side activity. Reduced hydrolytic rates would effectively limit the breakdown of cyclodextrins via the coupling reaction, thus enhancing the enzyme's applicability for industrial applications [11]. Previous protein engineering studies, aimed at improving the cyclodextrin size specificity of CGTases, have been of limited success. A double mutant, Y89D/S146P, at the substrate binding sites of *Bacillus circulans* 251 CGTase, increased α-cyclodextrin production two-fold [12]. Replacement of Ala-233 in *Bacillus clarkii* 7364 CGTase by basic amino acids enhanced γ-cyclodextrin forming activity over 4-fold in the neutral pH range [13], while substitution of Tyr-54 in the amylomaltase from *Thermus aquaticus* reduced hydrolytic activity but also enhanced cyclization activity significantly [14].

To generate CGTase variants with a low hydrolytic activity, we applied directed evolution, introducing random mutations throughout the *cgt* gene by error-prone (ep) PCR, followed by site-specific saturation mutagenesis. Parallel screening for high cyclization and reduced hydrolytic activity on starch yielded three mutations, with no direct interactions with substrate. The crystal structure of the most effective mutant, S77P, revealed subtle conformational changes of a network of amino acids at the active site responsible for the strongly reduced side reaction rates with retention of primary catalytic function.

Experimental Procedures

Bacterial strains, plasmids and protein production – Escherichia coli strain MC1061 was used for DNA manipulations and library screening. CGTase proteins were produced and purified using *Bacillus subtilis* strain DB104A carrying plasmid pCScgt-tt, as described [15]. Plasmid carrying strains were grown on LB medium at 37° C in the presence of kanamycin (50 μg/ml for *E. coli*

and 5 μg/ml for *B. subtilis*). Purity and molecular weight of CGTase proteins were checked by SDS-PAGE. Enzyme concentrations were determined using the Bradford reagent from Bio-Rad (München, Germany) and bovine serum albumin as standard.

DNA manipulations – Mutants were constructed in pCScgt-tt as described [16] and verified by DNA sequencing (BaseClear, Leiden, the Netherlands). Single mutants S77P and W239R were constructed using the following oligonucleotides: S77P, 5'-T TGG ATA CCT CAG CCT GTA G-'3; W239R, CAT ATG CCG TTT GGA CGG CAG AAG-'3. For site-specific saturation mutagenesis, Ser-77 and Trp-239 of Tabium CGTase were replaced by all nineteen other amino acids residues, using the following oligonucleotides: S77X 5'-T TGG ATA NNS CAG CCT GTA G-'3; W239X, CAT ATG CCG TTT GGA NNS CAG AAG-'3 and pCScgt-tt as PCR template. The combined saturated mutagenesis library, A231X/F260X, was previously constructed in the same manner [10]. Underlined region of oligo`s indicate where the nucleotide substitution were introduced. N is A+G+C+T, S is G+C and X is any amino acid residue.

Error-prone PCR mutagenesis – The *cgt* gene was amplified from pCScgt-tt with the primers For1 (*Nco*I), 5'- GGC TTT TCA GCC CTG CCC AGG CCA TGG-'3 and Rev1 (*Kpn*I), 5'-GTT TAC AAT TAC GGT ACC TGT ACT AGA-'3. Restriction sites are underlined. PCR mixtures (50 μl) contained: 1x Taq DNA polymerase buffer, 1 mM $MgSO_4$, 0.005 mM or 0.01 mM $MnCl_2$, 0.6 mM of each dNTP, 0.07 μM of each primer, 20 ng template and 2.5 units Taq DNA polymerase (Roche) [10]. PCR reactions were performed for 25 cycles: 30 sec 94° C, 40 sec 54° C, and 2 min 72° C. The PCR products were restricted with *Nco*I and *Kpn*I, and the resulting fragment (2100 bp) was extracted from agarose gel (QIAquick Gel Extraction Kit; Qiagen) and cloned in pCScgt-tt, replacing the wild-type *cgt* gene.

Enzyme assays – All enzyme assays were performed in 10 mM sodium citrate buffer at pH 6.0 and 60º C. β-Cyclodextrin forming activity was determined by incubating 1.3 – 2.6 nM of enzyme with a 2.5% (w/v) solution of partially hydrolysed potato starch with an average degree of polymerisation of 50 (Paselli SA2; AVEBE, Foxhol, The Netherlands). The amount of β-cyclodextrin produced was quantified with phenolphthalein [17]. Starch hydrolysing activity was measured by following the increase in reducing power with dinitrosalicylic acid, using 1% (w/v) soluble potato starch (Sigma-Aldrich) and 26 - 66 nM of enzyme. The disproportionation activity was measured as described [15], using 0.66 – 1.2 nM enzyme, 0.025 - 2 mM

4-nitrophenyl-α-D-maltoheptaoside-4-6-O-ethylidene (pNPG7; Megazyme, Wicklow, Ireland) and 0.5 - 20 mM maltose as donor and acceptor substrates, respectively. Coupling activities were measured as described [15], using α-, β- and γ-cyclodextrin (0.5 - 10; 0.5 - 7.5 and 0.5 - 10 mM, respectively) as donor substrates and 200 mM methyl-α-D-glucopyranoside (MαDG) as acceptor substrate. Linear products formed were converted to glucose with amyloglucosidase (Sigma, Zwijndrecht, the Netherlands) and the amount of glucose formed was determined with the GOD-PAP reagent (Roche, Almere, the Netherlands).

High-throughput screening of CGTase variants - *E. coli* MC1061 transformed with the library were plated on LB agar plates, and the resulting colonies were transferred to 200 µl of LB medium in 96-well microtiter plates using the Q-pix robot (Genetix, New Milton Hamsphire, U.K.) followed by incubation overnight (750 rpm) at 37º C. For starch hydrolysis, 25 µl of each culture was transferred to a second 96 well plate containing 25 µl of bacterial protein extraction reagent (Pierce, Rockford, IL) per well to lyse the cells. Subsequently, 200 µl of 1% (w/v) soluble starch (Lamers and Pleuger, Wijnegen, Belgium) in 10 mM sodium citrate buffer (pH 6.0) was added, and the microtiter plates were incubated at 60 °C. After 6 h the amount of reducing sugars formed was measured using an adapted version of the Nelson-Somogyi assay [18]. In parallel, β-cyclization activity was measured by addition of 50 µl of cell lysate to 200 µl of 1% (w/v) partially hydrolysed potato starch (Paselli SA2). Microtitre plates were then incubated at 60 ºC and the amount of β-cyclodextrin formed was measured after 2 h by the addition of a 10 µl sample to 100 µl phenolphthalein solution [17].

HPLC product analysis - Formation of cyclodextrins, glucose, maltose and maltotriose from 10% (w/w) starch (Paselli SA2) in 10 mM citrate buffer (pH 6.0), was analysed by incubating the starch solution for 58 h at 60 °C with 26 nM of wild-type and mutant proteins (S77P, W239R and W239L). Samples were subsequently boiled for 30 min for enzyme inactivation and products formed were analysed by HPLC, using an Econosphere NH2 5U column (250 by 4.6mm; Alltech Nederland BV, Breda, the Netherlands) linked to a refractive index detector. A mobile phase of acetonitrile/water (60/40, v/v) at a flow rate of 0.5 ml/min was used.

Differential Scanning Caliometery (DSC) – Thermal unfolding of CGTase proteins was measured using the VP-DSC microcalorimeter (MicroCal Inc., Northhampton, Massachusetts, USA). The cell volume was 0.52 ml and the experiments were performed at a scan rate of 1 °C/min at a constant pressure of 2.75 bar.

Samples were degassed prior to the scan. The enzyme concentration used was 6.9 µM in 10 mM sodium acetate buffer at pH 5.5 [19].

Structure determination - Before crystallization the Tabium S77P CGTase sample was concentrated with a Centricon-30K device to 80 – 106 µM in 10 mM sodium acetate, pH 5.5. Large crystals were grown from 17 – 20 % saturated ammonium sulphate at room temperature, 100 mM HEPES buffer, pH 7.6 (or 100 mM EPPS buffer, pH 7.8, or 100 mM Tris-HCl buffer, pH 8.0), like the wild-type enzyme [20]. Before data collection the crystal was soaked in 25% glycerol and directly flash cooled. Another crystal was soaked with 2% (w/v) acarbose, 4% (w/v) maltohexaose and 25% (v/v) glycerol in 20% saturated ammonium sulphate for 25 min followed by flash cooling. Diffraction data were collected at beamline ID14-3, ESRF, Grenoble, France, at 100 K and at a wavelength of 0.931 Å and processed using MOSFLM and SCALA from the CCP4 suite (Collaborative Computational Project Number 4, 1994). Data collection statistics are shown in Table 1. The structure of wild-type Tabium CGTase (PDB code 1A47) with all waters and the maltohexaose inhibitor removed was used as a starting model. Rigid body and restrained refinement were done with REFMAC5 in a standard way [21], followed by TLS refinement [22;23]. Ligands were manually placed in sigmaA-weighted 2Fo-Fc and Fo-Fc electron density maps using the program Coot [24], water molecules were automatically picked with Coot. In the native and sugar bound structures 12 amino acids have been built with double conformations. The quality of the models was checked using the program MolProbity [25]. The atomic coordinates and the structure factor amplitudes have been deposited with the RCSB Protein Data Bank with PDB codes 3BMV and 3BMW, respectively.

Results

Generation of low hydrolytic CGTase variants – Genetic diversity was introduced in Tabium CGTase by epPCR. Twelve thousand CGTase clones were subsequently screened in a parallel manner for decreased starch hydrolytic activity and high β-cyclization activity using microtitre based plate assays. Five strongly and three slightly improved single mutants were selected. Sequencing of these mutants revealed a S77P mutation in all five strongly improved variants and a W239R mutation in the three slightly improved variants. Both Ser-77 and Trp-239 residues are strongly conserved amongst bacterial and archaeal CGTases, and are replaced only in two hypothetical CGTases (data not shown). Saturated mutagenesis was subsequently carried out at Ser-77 and Trp-239 in an effort to reduce

hydrolytic activity even further and retain or raise β-cyclization activity. Six hundred clones were screened for both saturated mutant libraries, S77X and W239X. Selection followed by sequencing of five variants of the S77X library revealed the S77P mutation in each case. Sequencing of six selected variants revealed W239R (4x) and W239L (2x) mutations from the W239X library.

Catalytic rates of wild-type and mutant CGTases – Hydrolytic rates of selected variants S77P, W239R and W239L were lowered compared to wild-type CGTase (Table 2). While the W239R and W239L mutants displayed decreases of 27% and 37% in hydrolytic rates, the S77P variant appeared most effective, lowering the hydrolytic activity 15-fold compared to wild-type (Table 2). Initial β-cyclization rates remained similar to wild-type CGTase for the W239 mutants while the S77P variant displayed a 1.8-fold reduction. The overall cyclization/hydrolysis ratio for the S77P mutant increased 8.3-fold compared to wild-type (Table 2). For the disproportionation reaction an approximate 2-fold reduction in k_{cat} was observed for processing the blocked $pNPG_7$ substrate for the S77P and W239R variants (Table 3). An increase in K_M for $pNPG_7$ was also noted for the mutants despite the absence of direct interactions with the blocked substrate. The S77P and W239R mutations reduced catalytic efficiencies for the breakdown of all three cyclodextrins (α, β and γ) in the coupling reaction over 8 and 1.5-fold, respectively, compared to wild-type (Table 4).

Product profiles of wild-type and mutant CGTases – For wild-type CGTase the degradation of β- and α-cyclodextrins surpassed production after 6 h from starch (Fig. 2). The increased rate of cyclodextrin breakdown coincided with an increase in short linear saccharide production (Fig. 2). The W239L variant showed a reduced capacity in the production of short linear saccharides while both α-, β- and γ-cyclodextrin formation remained similar to that of wild-type (Fig. 2). Similarly, for the W239R mutant the amount of total short saccharides did not surpass α-, β- or γ-cyclodextrin formation at any stage. A minor reduction in β- and γ-cyclodextrin breakdown was noted for this variant compared to wild-type. In contrast, the S77P variant showed a continual increase in β-cyclodextrin production till 9 h with no detectable degradation of β-cyclodextrin thereafter (Fig. 2). Despite lower initial cyclodextrin formation, the amount of α-, β- and γ-cyclodextrins produced at 58 h by the S77P variant, was 32 (α), 81 (β) and 57 (γ) % higher than wild-type. Production of short saccharides remained extremely low throughout the 58 h incubation for S77P, 7.5-fold lower than that of wild-type (Fig. 2).

Stability of wild-type and mutant CGTases - To investigate whether the selected mutations affected enzyme stability, the CGTase variants were denatured by heat using differential scanning calorimetry (DSC). Both the wild-type and mutant proteins displayed irreversible thermal unfolding patterns. The S77P mutant displayed a similar apparent melting temperature as the wild-type with a 1 °C difference (Fig. 3). Both the W239L and W239R variants were, however, severely destabilised, lowering the apparent melting temperature by 15 and 21 °C, respectively, compared to wild-type (Fig. 3).

Structures – Two crystal structures of the S77P mutant were determined at a resolution of 1.6 Å (Table 1), one with and one without a heptasaccharide analog bound at the active site. In the acarbose and maltohexaose soaked crystal electron density was present for seven covalently linked sugar residues, at subsites –4 to +3 in the active site, with the valienamine moiety bound at –1 and the 6-deoxyglucose at +1, indicating a complete acarbose positioned from –1 to +3, coupled with a maltotriose at the -2 site (Fig. 4). Addition of sugar residues to acarbose is commonly observed in CGTase crystal structures [7;26;27]. Electron density was also found for a maltose molecule near Tyr-633 at maltose binding site 3 in the E-domain. Replacement of Ser-77 by Pro showed no alteration in the backbone conformation of residues 75 to 117, compared to the liganded wild-type (PDB Code 1A47). The main chain conformation was not affected by the S77P mutation as the phi/psi angles of Ser77 are -58/158, and those of Pro77 are -63/166. There were also no significant changes observed in temperature factors of the wild-type compared to the average temperature factors of the variants (15.8/20.5 $Å^2$ for wild-type enzyme and 7.9/7.3 $Å^2$ for the S77P mutant; the structures of the two variants were determined at different temperatures). However, the Tyr-101 side chain has a double conformation in the unliganded S77P variant. The conformation of the Arg-228 side chain is also altered in the S77P variant, in both structures with and without the ligand, when compared to the native structure.

Discussion

Effects of the S77P mutation on CGTase side reactions - By applying directed evolution to Tabium CGTase, we identified two mutants with single mutations (S77P and W239R) remote from the active site with strongly reduced hydrolysis side reaction rates on starch. Structural analysis of the mutant CGTase S77P helped to elucidate the effect of these distant mutations on the evolution of side reaction specificity of glucanotransferase enzymes.

The S77P residue is positioned at the beginning of a long loop, connecting

residues Trp-75 of strand β2 and Thr-117 of helix α2 of the $(\beta/\alpha)_8$ barrel structure of the catalytic domain (Fig. 4). Ser-77 is located over 10 Å away from the catalytic nucleophile Asp-230 and 8.7 Å from the nearest saccharide bound at the donor subsites (Fig. 5). While there is no direct interaction between S77P and the substrate, the 75-117 loop contains two essential amino acids, Tyr-101 and Trp-102, that stabilize the substrate bound at donor subsites -1/-2 via hydrophobic stacking and hydrogen bonding interactions, respectively [20]. In the wild-type structure, Ser-77 forms a water mediated hydrogen bonding interaction with Tyr-101. This interaction is lost in the S77P mutant and results in a lack of stabilisation of the Tyr-101 side chain. Alteration of the side chain conformation of Tyr-101, as seen from its double conformation in the unliganded S77P variant, directly affects the hydrogen bonding interaction with the neighbouring Arg-228 residue. The Arg-228 side chain attains a new conformation to retain the hydrogen bonding interaction with the distorted Tyr-101 side chain. This in turn affects the orientation of the acid/base catalyst Glu-258 side chain, which has a hydrogen bond interaction with Arg-228, prior to saccharide binding [28] (Fig. 5). From the crystal structure of the unliganded S77P mutant it appears that the Glu-258 side chain is orientated almost 1 Å further away from where the glycosidic oxygen of a substrate would be, than in wild-type (Fig. 5). Slight distortion of the Glu-258 side chain by Arg-228, may therefore require saccharide binding at the acceptor subsites to induce a more catalytically competent conformation of its side chain. This induced side chain repositioning would allow for effective protonation of the leaving saccharide during substrate cleavage and selective deprotonation of the acceptor molecule prior to α-1,4 bond formation [29]. The use of water as an acceptor molecule, however, would likely lack the necessary interactions with the active site residues to force the favourable positioning of the Glu-258 side chain for efficient deprotonation. In agreement with this induced-fit mechanism, recent structural studies of the amylomaltase of *Thermus thermophilus* revealed that the acid/base catalyst Glu-340 only assumes its catalytic competent conformation upon binding of an efficacious saccharide acceptor [30]. As expected this latter glucanotransferase enzyme displays the highest transglycosylation/hydrolysis activity ratio, of 5000:1 (Fig. 6) [31].

The Trp-239 residue is essential for both CGTase activity and stabilisation - The second mutant identified from our screening, residue W239R, is located at the start of the fourth α-helix of the $(\beta/\alpha)_8$ barrel (Fig. 4), following the loop (residues 230–236) containing the catalytic nucleophile Asp-230

and two residues, Ala-231 and Lys-233, critical in substrate binding and reaction specificity [18;32]. While the guanidinium group of the arginine residue at position 239 may take the place of two water molecules, modelling shows that the space occupied by the aromatic side chain of tryptophan is not completely filled by arginine or leucine, thus reducing the structural stability of these variants (Fig. 3). Additionally, the W239R/L mutations also destroy a hydrogen-bonding interaction between the side chains of Asp-209 and Trp-239, compromising the stability of these variants even further. The creation of greater space within this region may also change the packing of a hydrophobic cavity close to the active site, altering the orientation of the catalytic residues for a more selective deprotonation of the acceptor saccharide than the water molecule.

Thus, it appears that subtle conformational rearrangements at the active site induced by second shell mutants, (S77P, W239R/L) are essential for significant reduction in side reaction rates, with retention or even increase in primary catalytic function rates (Tables 2, 3 & 4) [33-36].

Effect of active site mutations on transglycosylation/hydrolysis activity - As the nature of the acceptor utilised determines reaction type, it seems logical that the residues at the acceptor subsites play a large role in determining reaction specificity. Previously it has been shown that the acceptor subsite residues of CGTases are important for the transglycosylation specificity of the enzyme [10;16]. Most mutations at the acceptor subsites, close to the catalytic nucleophile Asp-230 of Tabium CGTase, increased the hydrolytic activity significantly and lowered the transglycosylating activity of the enzyme (Fig. 6). To investigate whether mutations at the acceptor subsites, directly interacting with substrate, could enhance the transglycosylation specificity of CGTase, by lowering the competing hydrolysis reaction, we screened a saturated mutagenesis library of Tabium A231X/F260X CGTase, randomising two residues at the +1 and +2 acceptor subsites (Fig. 4). This library has been made previously [10] and yielded a mutant CGTase with an 11-fold higher hydrolytic activity than wild-type. Library analysis of 1,500 clones revealed that over 80% of the variants had strongly impaired β-cyclization activity with no variant retaining over 5% wild-type β-cyclization activity while reducing hydrolytic activity. Mutations at the acceptor sites were therefore not capable of increasing the cyclization versus hydrolysis performance of CGTase. Direct substitution of residues within this area resulted in either a strong increase in hydrolysis or a dramatic loss of overall activity (Fig. 6). Thus residues of the acceptor subsites have optimally evolved to

favour the use of saccharide acceptors over water, lowering the ancestral activity of hydrolysis while favouring transglycosylation activity (Fig. 6). This is a further indication that in order to lower enzyme side reactions, mutations of second shell active site residues are required to induce subtle conformational changes within the active site.

Specialization of glucanotransferase function from evolutionary intermediates within GH13 - Many glucanotransferase enzymes of GH13 display large variations in hydrolysis side reaction rates (Fig. 6). As continuous evolution of enzyme intermediates results in new specialized function with partial retention of initial ancestral function as a side reaction [37], we propose that retention of hydrolysis by these enzymes is a remnant of ancestral function. The specialized glucanotransferases of GH13 most likely evolved from 'intermediate enzymes' similar to that of maltogenic amylase and neopullulanase. Both enzymes produce similar amounts of end products from both hydrolase and transglycosylase activity on starch and maltodextrin (Fig. 6) [38;39]. However, due to growing competition between microbes for available carbon source from starch and glucans, increased pressure was placed upon the evolution of these 'intermediate enzymes' toward transglycosylation activity, for selective uptake and storage of the available saccharide substrate. For instance, certain bacteria secrete CGTases to monopolise on the starch substrate, converting it into cyclodextrins which cannot be used by competing organisms [40;41]. Gradually these glucanotransferase enzymes became more specialized till hydrolysis decreased to a mere side reaction or disappeared completely. Branching enzymes of GH13 which catalyze the formation of α-1,6 bonds, display no detectable hydrolytic activity on the amylose substrate [42]. These enzymes play a vital role in defining the structural and physical properties of the storage compound glycogen, in bacteria, archaea, fungi and animals.

As all members of GH13, including amylases and transglycosylases, share both an identical catalytic machinery and mechanism, with highly similar active site architecture, the question remains as to how these glucanotransferases became more selective in their catalysis of transglycosylase reactions. A likely explanation may come from the recent elucidation of the induced-fit mechanism in the catalytic cycle of amylomaltase from GH77 [30]. As previously mentioned, short saccharide binding at the acceptor subsites during the final stages of catalysis is required for reorientation of active site residues to allow the acid/base catalyst side chain to attain a catalytically competent

conformation. This ensures the preference for saccharide as acceptor over water for the transglycosylation reaction (Fig. 6). Mutations introducing this selective induced fit in GH13 glucanotransferases are most likely to have arisen on the outer regions of the enzymes active site. As our studies revealed, a single mutation on the outer regions of the active site is sufficient to induce subtle conformational changes to the catalytic residues resulting in increased preference for transglycosylation over hydrolysis activity. Introducing mutations neighbouring the catalytic residues may have too large a disruptive effect on the overall catalytic mechanism to increase primary catalytic function, as shown by investigation of our acceptor subsite library.

To conclude, continuous enzyme evolution within the functionally diverse GH13 family has resulted in intermediate enzymes catalysing new function while partially retaining ancestral function as a side reaction. By applying directed evolution to our Tabium CGTase enzyme we have successfully lowered the side reactions of hydrolysis and coupling while retaining the primary cyclization activity. Both effective mutations generated, S77P and W239R were identified outside the active site region, indicating that subtle conformational changes are sufficient to retain primed pre-existing catalytic function, while removing side reactions.

† HL acknowledges financial support from the Netherlands Organization for Scientific Research (NWO).

References

1 Aharoni, A., Gaidukov, L., Khersonsky, O., McQ Gould, S., Roodveldt, C., & Tawfik, D.S. (2005) The 'evolvability' of promiscuous protein functions. *Nat. Genet.* **37,** 73-76.

2 Kazlauskas, R.J. (2005) Enhancing catalytic promiscuity for biocatalysis. *Curr. Opin. Chem. Biol.* **9,** 195-201.

3 Bornscheuer, U.T. & Kazlauskas, R.J. (2004) Catalytic promiscuity in biocatalysis: using old enzymes to form new bonds and follow new pathways. *Angew. Chem. Int. Ed Engl.* **43,** 6032-6040.

4 Afriat, L., Roodveldt, C., Manco, G., & Tawfik, D.S. (2006) The latent promiscuity of newly identified microbial lactonases is linked to a recently diverged phosphotriesterase. *Biochemistry* **45,** 13677-13686.

5 Stam, M.R., Danchin, E.G., Rancurel, C., Coutinho, P.M., & Henrissat, B. (2006) Dividing the large glycoside hydrolase family 13 into subfamilies: towards improved functional annotations of alpha-amylase related proteins. *Protein Eng Des Sel.* **19,** 555-562.

6 Janecek, S. (1997) α-amylase family: molecular biology and evolution. *Prog. Biophys. Molec. Biol.* **25,** 67-97.

7 Uitdehaag, J.C.M., Mosi, R., Kalk, K.H., van der Veen, B.A., Dijkhuizen, L., Withers, S.G., & Dijkstra, B.W. (1999) X-ray structures along the reaction pathway of cyclodextrin glycosyltransferase elucidate catalysis in the alpha-amylase family. *Nature Struct. Biol.* **6,** 432-436.

8 Uitdehaag, J.C.M., van der Veen, B.A., Dijkhuizen, L., Elber, R., & Dijkstra, B.W. (2001) Enzymatic circularization of a malto-octaose linear chain studied by stochastic reaction path calculations on cyclodextrin glycosyltransferase. *Proteins:Struct. Funct. Genet.* **43,** 327-335.

9 van der Veen, B.A., Uitdehaag, J.C.M., Dijkstra, B.W., & Dijkhuizen, L. (2000) Engineering of cyclodextrin glycosyltransferase reaction and product specificity. *Biochim. Biophys. Acta* **1543,** 336-360.

10 Kelly, R.M., Leemhuis, H., & Dijkhuizen, L. (2007) Conversion of a cyclodextrin glucanotransferase into an alpha-amylase: assessment of directed evolution strategies. *Biochemistry* **46,** 11216-11222.

11 Qi, Q. & Zimmermann, W. (2005) Cyclodextrin glucanotransferase: from gene to applications. *Appl. Microbiol. Biotechnol.* **66,** 475-485.

12 van der Veen,B.A., Uitdehaag,J.C.M., Penninga,D., Van Alebeek,G.J., Smith,L.M., Dijkstra,B.W., & Dijkhuizen,L. (2000) Rational design of cyclodextrin glycosyltransferase from *Bacillus circulans* strain 251 to increase alpha-cyclodextrin production. *J. Mol. Biol.* **296,** 1027-1038.

13 Nakagawa,Y., Takada,M., Ogawa,K., Hatada,Y., & Horikoshi,K. (2006) Site-directed mutations in Alanine 223 and Glycine 255 in the acceptor site of gamma-Cyclodextrin glucanotransferase from Alkalophilic *Bacillus clarkii* 7364 affect cyclodextrin production. *J. Biochem.* **140,** 329-336.

14 Fujii,K., Minagawa,H., Terada,Y., Takaha,T., Kuriki,T., Shimada,J., & Kaneko,H. (2005) Use of random and saturation mutageneses to improve the properties of *Thermus aquaticus* amylomaltase for efficient production of cycloamyloses. *Appl. Environ. Microbiol.* **71,** 5823-5827.

15 Leemhuis, H., Dijkstra, B.W., & Dijkhuizen, L. (2003) *Thermoanaerobacterium thermosulfurigenes* cyclodextrin glycosyltransferase: mechanism and kinetics of inhibition by acarbose and cyclodextrins. *Eur. J. Biochem.* **270,** 155-162.

16 Leemhuis, H., Dijkstra, B.W., & Dijkhuizen, L. (2002) Mutations converting cyclodextrin glycosyltransferase from a transglycosylase into a starch hydrolase. *FEBS Lett.* **514,** 189-192.

17 Vikmon, M. (1982) Rapid and simple spectrophotometric method for determination of microamounts of cyclodextrins. (Szejlti, J.,ed) pp.64-74, Dordrecht, The Netherlands, Reidel Publishing Co. Proceedings of the first international symposium on cyclodextrins.

18 Leemhuis, H., Rozeboom, H.J., Wilbrink, M., Euverink, G.J.W., Dijkstra, B.W., & Dijkhuizen, L. (2003) Conversion of cyclodextrin glycosyltransferase into a starch hydrolase by directed evolution: the role of Ala230 in acceptor subsite +1. *Biochemistry* **42,** 7518-7526.

19 Leemhuis, H., Rozeboom, H.J., Dijkstra, B.W., & Dijkhuizen, L. (2004) Improved thermostability of *Bacillus circulans* cyclodextrin glycosyltransferase by the introduction of a salt bridge. *Proteins* **54,** 128-134.

20 Wind, R.D., Uitdehaag, J.C.M., Buitelaar, R.M., Dijkstra, B.W., & Dijkhuizen, L. (1998) Engineering of cyclodextrin product specificity and pH optima of the thermostable cyclodextrin glycosyltransferase from *Thermoanaerobacterium thermosulfurigenes* EM1. *J. Biol. Chem.* **273,** 5771-5779.

21 Vagin, A.A., Steiner, R.A., Lebedev, A.A., Potterton, L., McNicholas, S., Long, F., & Murshudov, G.N. (2004) REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr. D. Biol. Crystallogr.* **60,** 2184-2195.

22 Painter, J. & Merritt, E.A. (2006) Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr. D. Biol. Crystallogr.* **62,** 439-450.

23 Painter, J. & Merritt, E.A. (2007) *TLSMD* web server for the generation of multi-group TLS models. *J. Appl. Cryst.* **39,** 109-111.

24 Emsley, P. & Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D. Biol. Crystallogr.* **60,** 2126-2132.

25 Davis, I.W., Leaver-Fay, A., Chen, V.B., Block, J.N., Kapral, G.J., Wang, X., Murray, L.W., Arendall, W.B., Snoeyink, J., Richardson, J.S., & Richardson, D.C. (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* **35,** 375-383.

26 Strokopytov, B., Knegtel, R.M., Penninga, D., Rozeboom, H.J., Kalk, K.H., Dijkhuizen, L., & Dijkstra, B.W. (1996) Structure of cyclodextrin glycosyltransferase complexed with a maltononaose inhibitor at 2.6 angstrom resolution. Implications for product specificity. *Biochemistry* **35,** 4241-4249.

27 Mosi, R., Sham, H., Uitdehaag, J.C.M., Ruiterkamp, R., Dijkstra, B.W., & Withers, S.G. (1998) Reassessment of acarbose as a transition state analogue inhibitor of cyclodextrin glycosyltransferase. *Biochemistry* **37,** 17192-17198.

28 Leemhuis, H., Rozeboom, H.J., Dijkstra, B.W., & Dijkhuizen, L. (2003) The fully conserved Asp residue in conserved sequence region I of the α-amylase family is crucial for the catalytic site architecture and activity. *FEBS Lett.* **541,** 47-51.

29 Uitdehaag, J.C.M., van der Veen, B.A., Dijkhuizen, L., & Dijkstra, B.W. (2002) Catalytic mechanism and product specificity of cyclodextrin glycosyltransferase, a prototypical transglycosylase from the alpha-amylase family. *Enzyme Microb. Technol.* **30,** 295-304.

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

31 Kaper, T., Leemhuis, H., Uitdehaag, J.C., van der Veen, B.A., Dijkstra, B.W., van der Maarel, M.J., & Dijkhuizen, L. (2007) Identification of acceptor substrate binding subsites +2 and +3 in the amylomaltase from *Thermus thermophilus* HB8. *Biochemistry.* **46,** 5261-5269.

32 van der Veen, B.A., Leemhuis, H., Kralj, S., Uitdehaag, J.C.M., Dijkstra, B.W., & Dijkhuizen, L. (2001) Hydrophobic amino acid residues in the acceptor binding site are main determinants for reaction mechanism and specificity of cyclodextrin glycosyltransferase. *J. Biol. Chem.* **276,** 44557-44562.

33 Hinz, S.W., Doeswijk-Voragen, C.H., Schipperus, R., van den Broek, L.A., Vincken, J.P., & Voragen, A.G. (2006) Increasing the transglycosylation activity of alpha-galactosidase from *Bifidobacterium adolescentis* DSM 20083 by site-directed mutagenesis. *Biotechnol. Bioeng.* **93,** 122-131.

34 Kuriki, T., Kaneko, H., Yanase, M., Takata, H., Shimada, J., Handa, S., Takada, T., Umeyama, H., & Okada, S. (1996) Controlling substrate preference and transglycosylation activity of neopullulanase by manipulating steric constraint and hydrophobicity in active center. *J Biol. Chem.* **271,** 17321-17329.

35 Feng, H.Y., Drone, J., Hoffmann, L., Tran, V., Tellier, C., Rabiller, C., & Dion, M. (2005) Converting a {beta}-glycosidase into a {beta}-transglycosidase by directed evolution. *J. Biol. Chem.* **280,** 37088-37097.

36 Morley, K.L. & Kazlauskas, R.J. (2005) Improving enzyme properties: when are closer mutations better? *Trends Biotechnol.* **23,** 231-237.

37 Matsumura, I. & Ellington, A.D. (2001) In vitro evolution of beta-glucuronidase into a beta-galactosidase proceeds through non-specific intermediates. *J. Mol. Biol.* **305,** 331-339.

38 Kim, T.J., Kim, M.J., Kim, B.C., Kim, J.C., Cheong, T.K., Kim, J.W., & Park, K.H. (1999) Modes of action of acarbose hydrolysis and transglycosylation catalyzed by a

thermostable maltogenic amylase, the gene for which was cloned from a *Thermus* strain. *Appl. Environ. Microbiol.* **65,** 1644-1651.

39 Takata, H., Kuriki, T., Okada, S., Takesada, Y., Iizuka, M., Minamiura, N., & Imanaka, T. (1992) Action of neopullulanase. Neopullulanase catalyzes both hydrolysis and transglycosylation at alpha-(1-4)- and alpha-(1-6)-glucosidic linkages. *J. Biol. Chem.* **267,**18447-18452.

40 Hashimoto,Y., Yamamoto,T., Fujiwara,S., Takagi,M., & Imanaka,T. (2001) Extracellular synthesis, specific recognition, and intracellular degradation of cyclomaltodextrins by the hyperthermophilic archaeon *Thermococcus* sp. strain B1001. *J. Bacteriol.* **183,** 5050-5057.

41 Fiedler, G., Pajatsch, M., & Bock, A. (1996) Genetics of a novel starch utilisation pathway present in *Klebsiella oxytoca*. *J. Mol. Biol.* **256,** 279-291.

42 Thiemann, V., Saake, B., Vollstedt, A., Schafer, T., Puls, J., Bertoldo, C., Freudl, R., & Antranikian, G. (2006) Heterologous expression and characterization of a novel branching enzyme from the thermoalkaliphilic anaerobic bacterium *Anaerobranca gottschalkii*. *Appl. Microbiol. Biotechnol.* **72,** 60-71.

43 DeLano,W.L. The PyMOL Molecular Graphics System. 2002.

44 Cha, H.J., Yoon, H.G., Kim, Y.W., Lee, H.S., Kim, J.W., Kweon, K.S., Oh, B.H., & Park, K.H. (1998) Molecular and enzymatic characterization of a maltogenic amylase that hydrolyzes and transglycosylates acarbose. *Eur. J. Biochem.* **253,** 251-262.

45 Albenne, C., Skov, L.K., Mirza, O., Gajhede, M., Feller, G., D'Amico, S., Andre, G., Potocki-Veronese, G., van der Veen, B.A., Monsan, P., & Remaud-Simeon, M. (2004) Molecular basis of the amylose-like polymer formation catalyzed by *Neisseria polysaccharea* amylosucrase. *J. Biol. Chem.* **279,** 726-734.

46 Kelly,R.M., Leemhuis,H., Gatjen,L., & Dijkhuizen,L. (2008) Evolution toward small molecule inhibitor resistance affects native enzyme function and stability: generating acarbose insensitive cyclodextrin glucanotransferase variants. *J. Biol. Chem.* Doi: 10.1074/jbc.M709287200.

47 Roujeinikova, A., Raasch, C., Burke, J., Baker, P.J., Liebl, W., & Rice, D.W. (2001) The crystal structure of *Thermotoga maritima* maltosyltransferase and its implications for the molecular basis of the novel transfer specificity. *J. Mol. Biol.* **312,** 119-131.

48 Kaper, T., Talik, B., Ettema, T.J., Bos, H., van der Maarel, M.J., & Dijkhuizen, L. (2005) Amylomaltase of *Pyrobaculum aerophilum* IM2 produces thermoreversible starch gels. *Appl. Environ. Microbiol.* **71,** 5098-5106.

TABLES & FIGURES

Table 1. Data collection statistics and quality of the *Thermoanaerobacterium thermosulfurigenes* EM1 CGTase mutant S77P.

	S77P	S77P with inhibitor
Data collection		
Spacegroup	$P2_12_12_1$	$P2_12_12_1$
Cell axes a, b, c (Å)	73.6, 96.0, 114.2	73.8, 96.6, 115.2
Resolution range (Å)	50.0 – 1.6	50.0 – 1.6
Reflections total (unique)	353324 (95798)	384566 (102676)
Completeness (%)[a]	89.9 (57.2)	94.2 (72.7)
$<I/\sigma(I)>$[a]	14.2 (5.2)	11.6 (3.6)
R_{merge} (%)[a]	7.6 (18.4)	10.0 (19.0)
Refinement stastics		
No. of amino acids	683 (all)	683 (all)
No. of Ca^{2+} ions	2	2
No. of SO_4^{2-} ions	5	3
No. of Cl^- ions	-	1
No. of glycerol molecules	4	7
No. of water molecules	1023	1160
Active site ligand	-	maltotriose-acarbose
MBS3 ligand	-	maltose
Final R-factor (%)	14.2	14.4
Final free R-factor (%)	16.3	16.5
Root mean square deviation (rmsd) from ideal geometry		
Bond lengths (Å)	0.007	0.007
Angles (°)	1.1	1.1
Ramachandran plot (%)		
favored	96.9	97.2
allowed	3.1	2.8
disallowed	0.0	0.0

[a] Highest resolution shell in parentheses

Table 2. The β-cyclization and hydrolysing activities of wild-type and mutant Tabium CGTases at 60 °C and pH 6.0.

Enzyme	k_{cat} (s^{-1})		Cyclization/ Hydrolysis	Fold-reduction[a]
	β-cyclization	Hydrolysis		
Wild-type	324 ± 6	40 ± 2	8.1	1
S77P	182 ± 4	2.7 ± 0.3	67.4	15
W239L	317 ± 0.1	29 ± 2	10.9	1.4
W239R	247 ± 2	25 ± 3	9.9	1.6

[a]Fold-reduction in hydrolytic activity of mutants compared to wild-type.

Table 3. Kinetic parameters of the disproportionation reaction of wild-type and mutant Tabium CGTases at 60 °C and pH 6.0.

Enzyme	k_{cat} (s^{-1})	$K_{M,\ pNPG7}$ (mM)	$K_{M,\ maltose}$ (mM)	$k_{cat}/K_{M,\ pNPG7}$ ($s^{-1}\ mM^{-1}$)
Wild-type	1244	0.15	3.3	8293
S77P	558	0.24	2.4	2325
W239L	1082	0.27	4.2	4007
W239R	644	0.2	4.1	3220

Table 4. Kinetic parameters of the coupling reaction of wild-type and S77P mutant Tabium CGTases at 60 °C and pH 6.0.

Enzyme	k_{cat} (s^{-1})	K_M (mM)	k_{cat} / K_M (s^{-1} mM^{-1})
α-CD			
Wild-type	1408 ± 40	1.6 ± 0.2	880
S77P	110 ± 3	1 ± 0.1	110
W239L	1285 ± 49	1.8 ± 0.1	713
W239R	644 ± 29	1.3 ± 0.2	495
β-CD			
Wild-type	285 ± 8	0.9 ± 0.07	317
S77P	26 ± 1	0.7 ± 0.06	37
W239L	295 ±10	0.8 ± 0.1	369
W239R	160 ± 4	0.8 ± 0.1	200
γ-CD			
Wild-type	321 ± 13	1.3 ± 0.2	247
S77P	44 ± 3	1.7 ± 0.2	26
W239L	472 ±19	1.8 ± 0.3	262
W239R	254 ± 8	1.6 ± 0.2	159

Figure 1. Schematic representation of the four reactions catalysed by CGTase. A saccharide binds to the substrate binding subsites of CGTase followed by bond cleavage to yield an enzyme-glycosyl intermediate. The second step of the reaction involves the transfer of the covalently bound saccharide to an acceptor molecule. (A) In the primary cyclization reaction the OH-4 group of the non-reducing end saccharide is used as acceptor. (B) A short saccharide acts as acceptor for disproportionation. (C) For the side reaction of hydrolysis a water molecule replaces the saccharide as acceptor for shorter saccharide production. (D) Coupling, the reverse reaction of cyclization, involves binding of cyclodextrin, breakdown of ring structure, linearization of saccharide before subsequent breakdown of enzyme-glycosyl intermediate by attacking of an acceptor saccharide.

Figure 2. Production of cyclodextrins and the short linear saccharides (glucose, maltose and maltotriose) in g/L from 10% (w/v) Paselli SA2 starch by wild-type and mutant CGTases at pH 6.0 and 60 °C. α, β and γ-cyclodextrins are indicated by black diamonds, circles, and squares, respectively, with short saccharides represented by open circles.

Figure 3. Thermal denaturation curves of wild-type and mutant CGTases as measured with differential scanning calorimetry. The panel gives the apparent melting temperatures of variants.

Figure 4. Structural representation of Tabium CGTase mutant, S77P, with heptasaccharide analog bound at the –4 to +3 subsites (Protein Data Bank 3BMW). Location of the S77P, W239, A231 and F260 residues within the protein structure are indicated by red spheres. The lower panel displays a close-up view of the catalytic core region from the above structure. Figure 4 was created using PyMOL [43].

Figure 5. Molecular modelling of the location and effect of S77P mutation on the Tabium CGTase active site conformation. (A) Superimposition of the wild-type and S77P active sites. (B) Superimposition of the wild-type and S77P active sites with a G6 or G7 ligand bound, respectively. Wild-type Tabium CGTase residues are coloured white and S77P mutant residues dark grey. Hydrogen bond interactions are indicated as black dashed lines. Active site subsites shown are numbered -1 to +2. Figure 5 was created using PyMOL [43].

Figure 6. Development of specialized enzyme function from a common evolutionary ancestor. (A) Both transglycosylation/hydrolase activities and ratios are shown for members of family GH13. Also included are amylomaltases from GH77. 1, neopullulanase [34]; 2, maltogenic amylase [44]; 3, Tabium CGTase; 4, amylosucrase [45]; 5, CGTase mutant S77P; 6, *Bacillus circulans* 251 CGTase [46]; 7, maltosyltransferase [47]; 8, amylomaltase [31]; 9, amylomaltase [48]; 10, branching enzyme. Branching enzyme displays no detectable hydrolytic activity on the amylose substrate [39]. (B) The effect of mutations in the donor and acceptor subsites of Tabium CGTase on the initial cyclization and hydrolytic rates with starch [10;16;20]. S_A and S_D indicate mutated residues located in or closest to the acceptor and donor subsites, respectively.

Figure 1

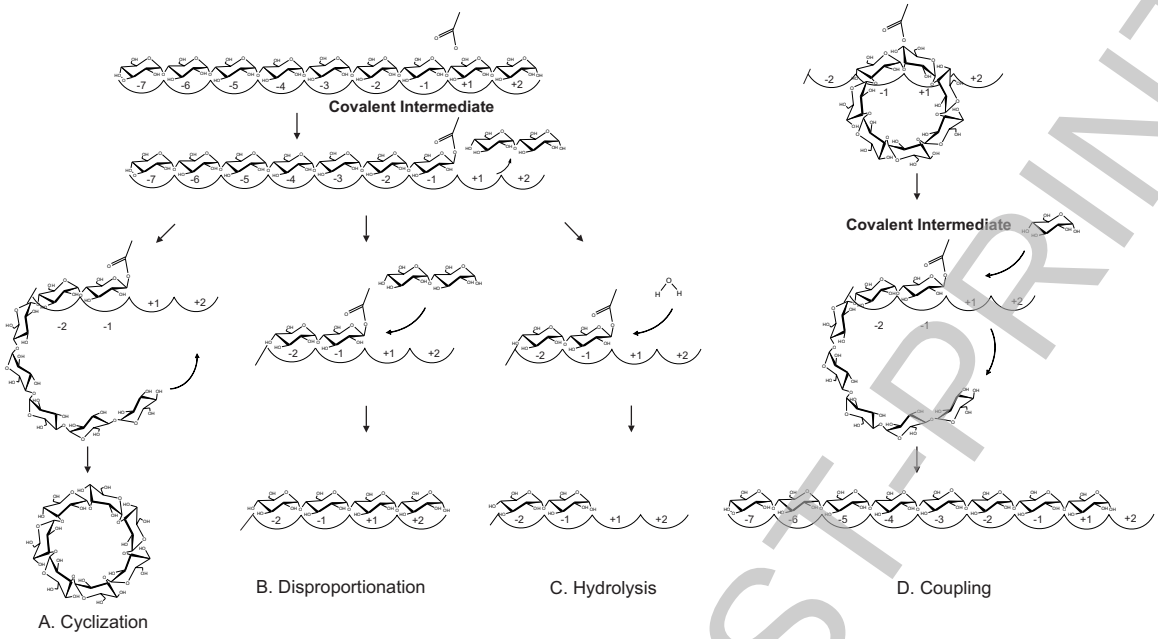


Figure 2

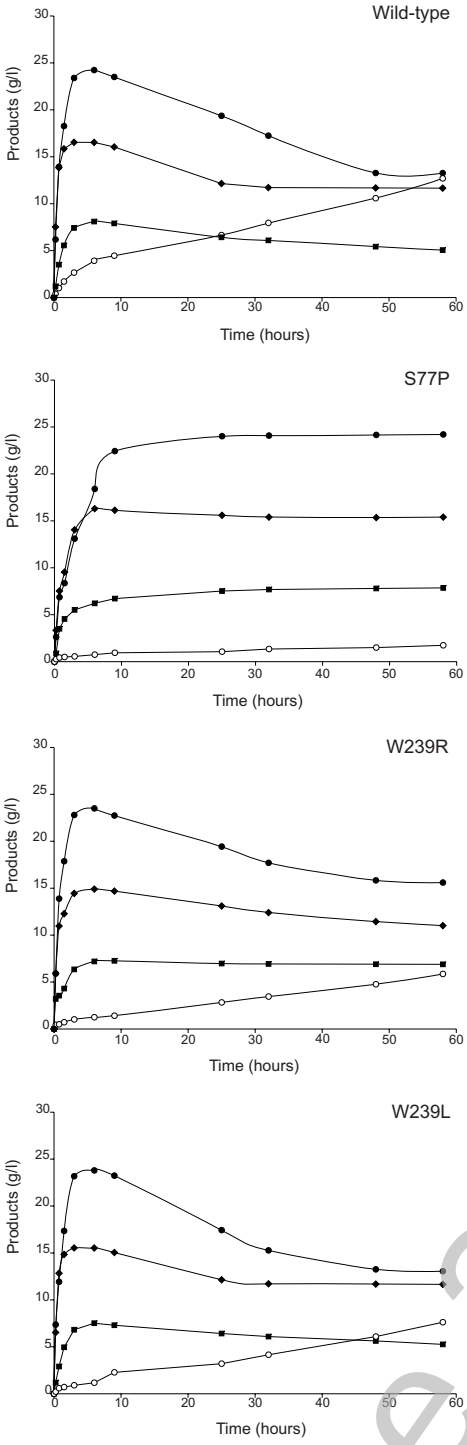


Figure 3

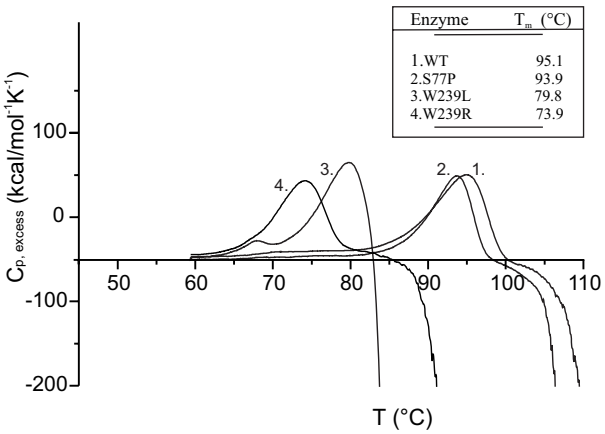
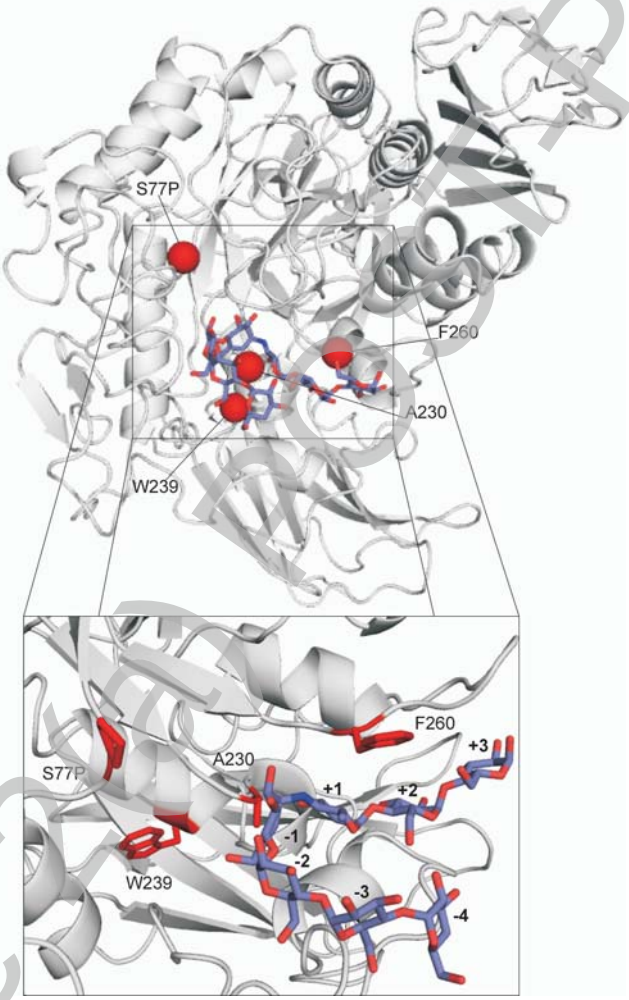


Figure 4



Biochemical Journal Immediate Publication. Published on 21 Apr 2008 as manuscript BJ20080353

Figure 5

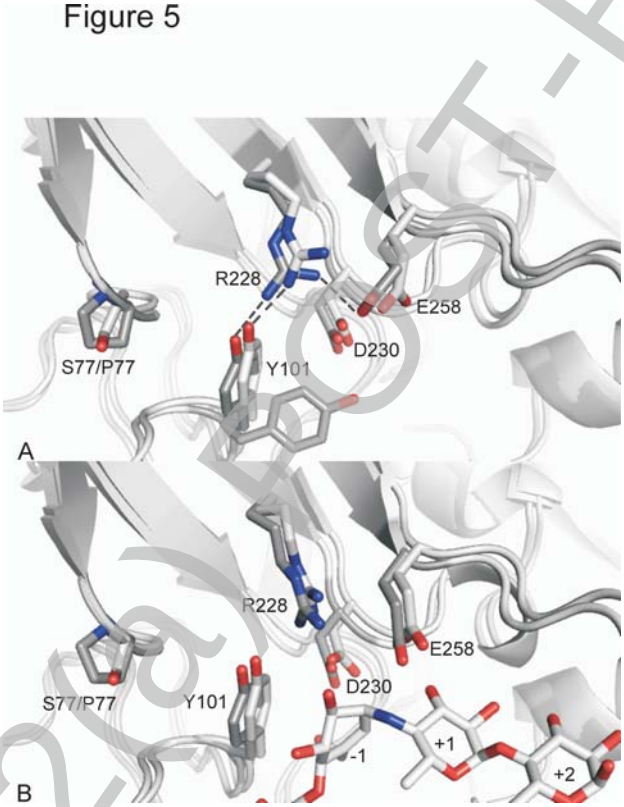


Figure 6