# Tailoring the substrate specificity of the β-glycosidase from the thermophilic archaeon *Sulfolobus solfataricus*[1]

Kevin Corbett[a,b,c,d], Anthony P. Fordham-Skelton[e], John A. Gatehouse[b,c], Benjamin G. Davis[a,c,d,*]

[a] *Department of Chemistry, University of Durham, South Road, Durham DH1 3LE, UK*
[b] *Department of Biological Sciences, University of Durham, South Road, Durham DH1 3LE, UK*
[c] *Research Centre for Biological Chemistry, University of Durham, South Road, Durham DH1 3LE, UK*
[d] *Dyson Perrins Laboratory, University of Oxford, South Parks Road, Oxford OX1 3QY, UK*
[e] *CLRC, Daresbury Laboratory, Warrington, Cheshire WA4 4AD, UK*

Received 19 October 2001; accepted 6 November 2001

First published online 22 November 2001

Edited by Pierre Jolles

**Abstract The substrate specificity of the thermophilic β-glycosidase (*lacS*) from the archaeon *Sulfolobus solfataricus* (SSβG), a member of the glycohydrolase family 1, has been analysed at a molecular level using predictions from known protein sequences and structures and through site-directed mutagenesis. Three critical residues were identified and mutated to create catalysts with altered and broadened specificities for use in glycoside synthesis. The wild-type (WT) and mutated sequences were expressed as recombinant fusion proteins in *Escherichia coli*, with an added $His_6$-tag to allow one-step chromatographic purification. Consistent with side-chain orientation towards OH-6, the single Met439→Cys mutation enhances D-xylosidase specificity 4.7-fold and decreases D-fucosidase activity 2-fold without greatly altering its activity towards other D-glycoside substrates. Glu432→Cys and Trp433→Cys mutations directed towards OH-4 and -3, respectively, more dramatically impair glucose (Glc), galactose (Gal), fucose specificity than for other glycosides, resulting in two glycosidases with greatly broadened substrate specificities. These include the first examples of stereospecificity tailoring in glycosidases (e.g. WT→W433C, $k_{cat}/K_M$ (Gal):$k_{cat}/K_M$ (mannose (Man)) = 29.4:1→1.2:1). The robustness and high utility of these broad specificity SSβG mutants in parallel synthesis were demonstrated by the formation of libraries of β-glycosides of Glc, Gal, xylose, Man in one-pot preparations at 50°C in the presence of organic solvents, that could not be performed by SSβG-WT.** 

*Key words:* β-Glycosidase; Substrate specificity; Site-directed mutagenesis; *Sulfolobus solfataricus* β-glycosidase; Biocatalysis; Enzymatic glycoside synthesis

## 1. Introduction

Recent advances in the development of carbohydrate-based therapeutics [1], and the limitations of present chemical synthetic methods for producing oligosaccharides, have led to more novel approaches to the synthesis of carbohydrates and their conjugates [2]. One approach to this problem is to carry out such syntheses using carbohydrate processing enzymes such as glycosyltransferases or glycosidases, as a valuable source of catalytic activity for the manipulation of unprotected carbohydrates [3–7]. Glycosidases are simple, robust, soluble enzymes, and in general have been preferred for such glycosynthesis [8,9]. Although catalysis of the hydrolysis of glycoside bonds is normally observed, glycosidases may be successfully used to synthesise glycosides through reverse hydrolysis (thermodynamic control) or transglycosylation (kinetic control with activated donors) strategies. Thus far, improvements in glycosidase synthetic utility have largely focused upon developing new strategies for increasing low product yields [10], improving regioselectivity of transfer [11] or characterising available glycosidases for novel activities [8]. For example, a major advance in improving yields has been the development of the glycosynthase [10,12–17] by Withers and co-workers. These nucleophile-less glycosidase mutants are capable of glycosyl transfer in yields of up to 90% using glycosyl fluoride donors but do not hydrolyse glycoside products and they illustrate well the benefits of glycosidase engineering for creating more synthetically useful catalysts.

An area of glycosidase engineering which has thus far been largely neglected is the engineering of new substrate specificities (to the best of our knowledge only three examples of glycosidase specificity alteration exist: 1000-fold increased fucose (Fuc):galactose (Gal) specificity from gene-shuffling induced alteration of six residues [18]; 200-fold increase in glucose (Glc):Fuc specificity from a single point mutation [19]; 5-fold increase in activity towards 6-phospho-β-Gal from a single point mutation [20]; for an excellent review of remaining challenges in glycosidase engineering see [21]). Since the nature of the parent carbohydrate to be coupled to a given acceptor may be determined in synthesis simply through appropriate choice of donor, it is largely the stereoselectivity of a given glycosidase that we wish to exploit. An area of growing interest is that of combinatorial biocatalysis [22,23]: the use of enzyme catalysts in parallel reactions to provide arrays of

*Corresponding author. Fax: (44)-01865-275674.
*E-mail address:* ben.davis@chem.ox.ac.uk (B.G. Davis).

[1] Dedicated to Professor Robert Cummings.

*Abbreviations:* SSβG, *Sulfolobus solfataricus* β-glycosidase; BPβG, *Bacillus polymyxa* β-glycosidase; WT, wild-type; Man, mannose; Gal, galactose; Glc, glucose; Xyl, xylose; GlcA, glucuronic acid; 4-MU-XXX, 4-methylumbelliferyl β-D-glycoside; SDS–PAGE, sodium dodecyl sulphate–polyacrylamide gel electrophoresis

0014-5793/01/$20.00 
PII: S0014-5793(01)03154-4

related molecules. In particular, the importance of gaining access to diverse arrays of glycoconjugates has recently been highlighted [24]. However, although combinatorial chemistry has revolutionised the approach to traditional chemical synthesis, the development of combinatorial biocatalysis has been hampered by the often stringent substrate specificities of synthetically useful enzymes. As a solution to these problems we have set ourselves the goal of (a) creating glycosidases with the ability to process a broad range of glycosides while maintaining their excellent stereoselectivity in glycoside bond formation; (b) retuning the specificity of glycosidases in alternative directions.

To this end, we have probed the binding domain of the thermophilic, retaining, exo-β-glycosidase, from *Sulfolobus solfataricus* (SSβG, EC 3.2.1.23) using site-directed mutagenesis. The gene encoding this enzyme was originally isolated and sequenced from the *S. solfataricus* strain MT4 [25] and is classified as a member of the glycosyl hydrolase family 1 [26]. This robust, thermophilic enzyme is ideal as it has been routinely expressed in *Escherichia coli* [27], characterised in terms of its properties [28,29], activity and binding specificity [30], has shown initial utility in synthesis [15] and its three-dimensional (3D) structure, a classic $(\alpha/\beta)_8$ TIM [31] barrel containing a radial active site channel in a kink of the fifth α/β repeat, is known [32]. Substrate specificity in this enzyme has been associated with two residues in the binding site, glutamate 432 and methionine 439 which are largely conserved across family 1 glycosyl hydrolases (Fig. 2). For example, E432 (SSβG numbering) has been suggested to be responsible for the rejection of 6-phosphoglycosides and glycosides of glucuronic and galacturonic acid as substrates by the enzyme, due to repulsive charge interactions [32]. Importantly, those family 1 hydrolases in which these residues differ also show altered substrate specificities (vide infra). It follows therefore that alteration of these residues and others in SSβG might result in mutants that will accommodate differing substrates. With this rationale in mind, we have analysed the structure of SSβG and created point mutants in which key residues implicated in specificity determination have been tailored. This results in robust mutant enzymes with altered substrate specificities and enhanced synthetic utility.

## 2. Materials and methods

### 2.1. Reagents, enzymes and bacterial strains

The wild-type (WT) sequence, *lacS*, encoding SSβG, was amplified by PCR from *Sulfolobus* genomic DNA, using the following primers: 5′: CCATGGGACACCACCACCACCACCACTCATTAC; 3′: CTCGAGTTAGTGCCTTTATGGCTTTACTGGAGGTAC.

The 5′ primer introduced an N-terminal *Nco*I site and a 6× His-tag immediately following the ATG initiation codon. The 3′ primer introduced a *Xho*I site after the stop codon. The PCR product was cloned into pCR2.1 (Invitrogen) and individual clones were sequenced to verify that no errors had been introduced.

Electrocompetent *E. coli* strain BL21(DE3) and His-bind nickel resin were obtained from Novagen. 4-Methylumbelliferyl (4-MU) β-D-glycoside substrates were purchased from Sigma. *Pfu*-turbo DNA polymerase was obtained from Stratagene and *Nco*I, *Xho*I restriction endonucleases, T4 DNA ligase from Promega, UK. Oligonucleotide primers were obtained from MWG BioTech GmbH and Cruachem Ltd. DNA sequencing was carried out by the DNA Sequencing Service, Department of Biological Sciences, Durham, UK, using standard protocols on Applied Biosystems DNA Sequencers.

### 2.2. Construction, selection and screening of the single point mutants

Mutations were introduced into the *lacS* gene coding sequence (in pCR2.1) according to the Stratagene QuickChange mutagenesis system, using the supplier's protocol. Oligonucleotide primers used for the generation of the point mutations were: for Glu432→Cys, 5′-TCTAGCTGATAATTAC**TGT**TGGGCTTCAGGATTCT-3′; for Trp433→Cys, 5′-CTAGCTGATAATTACGAA**TGT**GCTTCAGGATTCTC-3′; for Met439→Cys, 5′-GCTTCAGGATTCTCT**TGT**AGGTTTGGTCTG-3′ along with the corresponding complementary primers. Individual point mutations were verified by DNA sequence analysis. WT and mutated coding sequences were cloned into the *Nco*I/*Xho*I sites of expression vector pET-24-d(+) (Novagen) and transformed into *E. coli* BL21(DE3). Putative transformants were identified by colony PCR using the SSβG coding sequence primers. Selected clones were checked by DNA sequencing to confirm the mutation, and the absence of unintended PCR-introduced base changes.

### 2.3. Overexpression and purification of the $His_6$-tagged mutant enzymes

Selected clones were grown in LB medium containing kanamycin (50 μg/ml), at 37°C to an optical density of 0.6 at 600 nm, and the target were proteins induced by the addition of 0.1 M IPTG. Cells were harvested by centrifugation, resuspended in 1/10 volume of column loading buffer (5 mM imidazole, 20 mM Tris, 0.5 M NaCl, pH 7.8), and lysed using a Soniprep 150 Sonicator. The suspension was recentrifuged to pellet cell debris (10 000 rpm, 30 min), and the $His_6$-tagged recombinant proteins were purified from the supernatant using Ni-chelation chromatography (wash buffer: 60 mM imidazole, 20 mM Tris, 0.5 M NaCl, pH 7.8; elution buffer: 300 mM imidazole, 20 mM Tris, 0.5 M NaCl, pH 7.8). The eluted protein peak was dialysed against 50 mM sodium phosphate buffer (pH 6.5), and stored at 4°C. Protein concentration was quantified by the method of Bradford [33] (reagents from Bio-Rad, The Netherlands). Purified proteins were analysed by sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS–PAGE), gel filtration chromatography and ESMS (Micromass LCT, ±8 Da).

### 2.4. Characterisation of the kinetic properties of enzymes

Parameters were determined by the method of initial rates. Activity was measured in time course assays of the hydrolysis of 4-MU β-D-glycosides (β-D-gluco, β-D-galacto, β-D-fuco, β-D-manno, β-D-xylo, β-D-glucurono) at 5–15 concentrations (0.001–1.5 mM) incubated at 80°C in 50 mM sodium phosphate buffer, pH 6.5. Reactions were terminated at 2, 5, 10 and 15 min by the addition of 100 μl of ice cold 1 M $Na_2CO_3$, pH 10 and analysed (Labsystems Fluoroscan Ascent plate reader, excitation 460 nm, emission 355 nm). $K_M$ and $k_{cat}$ were derived by fitting the initial rates to Michaelis–Menten curves using GraFit 4 (Erithacus Software Ltd., Staines, UK).

### 2.5. Sequence analysis

Sequence alignment was performed using ClustalW based on a BLOSUM42 matrix. Enzymes of interest were determined by their sequence similarity using PSI-BLAST searches of SwissProt and TREMBL to BGAL_SULSO (SSβG) [25], including *Pyrococcus furiosus* β-glucosidase (*CelB*) [34] used for molecular mechanics analysis. In this way several glycosidases were also identified with both altered substrate specificity and differences in the residues occupying positions 432, 433 and 439 (SSβG numbering): *Dalbergia cochinchinensis* dalcochinin-8′-*O*-β-glucoside β-glucosidase [35]; *Costus speciosus* furostanol-26-*O*-β-glycoside hydrolase [36]; LPH_HUMAN, human lactase phlorizin hydrolase [37]; MY3_SINAL, myrosinase from *Sinapis alba* [38]; LACG_STAAU (6-PBG), *Staphylococcus aureus* 6-phosphogalactosidase [39].

### 2.6. Molecular mechanics and docking analysis

The X-ray structure of SSβG (RCSB-PDB entry 1gow) was used as the starting point for calculations. The enzyme setup was performed with Insight II, version 2.3.0 (Accelerys Inc., San Diego, CA, USA). To create initial coordinates for the minimisation, hydrogens were added at the pH used for kinetic measurements (6.5). The model system was solvated with a 5 Å layer of water molecules. Energy simulations were performed with the DISCOVER module within Cerius2, Version 3.8 on a Silicon Graphics Indigo computer, using the consistent valence force field function. A non-bonded cutoff distance of 18 Å with a switching distance of 2 Å was employed. The non-bonded pair list was updated every 20 cycles and a dielectric constant of 1 was used in all calculations. Docked structures were generated

using the Builder module, and aligned within the active site using appropriate bump, hydrogen bonding and docking interaction monitors. The enzyme was then minimised in stages, with initially only the water molecules being allowed to move, followed by water molecules and the amino acid side chains, and then finally the entire enzyme. The β-D-Glc*p* was free to move throughout all stages of the minimisation. Each stage of energy minimisation was conducted by means of the method of steepest descends without Morse or cross terms until the derivative of energy with respect to structural perturbation was less than 5.0 kcal/Å; then the method of conjugate gradients, without Morse or cross terms until the derivative of energy with respect to structural perturbation was less than 1.0 kcal/Å; and finally the method of conjugate gradients, with Morse and cross terms until the final derivative of energy with respect to structural perturbation was less than 0.1 kcal/Å.

### 2.7. Glycoside synthesis

Enzyme (WT, W443C or E432C, 1 mg) was added to a mixed solution (1 ml) of *para*-nitrophenyl (pNP) β-D-manno-, galacto-, gluco- and xylo-pyranosides (0.03 mmol of each) in 1:9 MeOH:phosphate buffer (pH 6.5) and incubated at 50°C for 45 min (WT), 4 h (WT), 8 h (W433C, E432C). After this time the solutions were extracted with EtOAc to remove *para*-nitrophenol and passed through short Sephadex and Celite:Graphite (1:1) columns to remove protein, pNP-glycoside and remaining *para*-nitrophenol. Solvent was removed and product mixtures were analysed by $^1$H nuclear magnetic resonance (NMR) and ESMS. Yields based on donor were calculated from integration of anomeric proton resonances in $^1$H NMR ($D_2O$, 500 MHz): α-Gal ($\delta$ 5.12, d, $J$ 4.0 Hz), α-Glc ($\delta$ 5.08, d, $J$ 3.8 Hz), α-xylose (Xyl) ($\delta$ 5.04, d, $J$ 3.4 Hz), α-mannose (Man) ($\delta$ 5.03, d, $J$ 1.8 Hz), β-Man ($\delta$ 4.75, s), β-Glc ($\delta$ 4.49, d, $J$ 8.0 Hz), β-Gal ($\delta$ 4.45, d, $J$ 7.9 Hz), Me-β-Man ($\delta$ 4.44, s), β-Xyl ($\delta$ 4.42, d, $J$ 7.8 Hz), Me-β-Glc ($\delta$ 4.23, d, $J$ 8.1 Hz), Me-β-Xyl ($\delta$ 4.18, d, $J$ 7.8 Hz), Me-β-Gal ($\delta$ 4.17, d, $J$ 8.0 Hz).

## 3. Results and discussion

### 3.1. Analysis of the −1 binding site of SSβG

In an attempt to dissect the specificity determining interactions of SSβG with its substrates we examined the 3D structures of SSβG (RCSB-PDB 1gow) and the close structural homologue *Bacillus polymyxa* β-glycosidase (BPβG). Valuably, 3D structures of BPβG containing D-gluconate bound as a substrate mimic (1bgg) and a 2-deoxy-2-fluoro-α-D-glu-

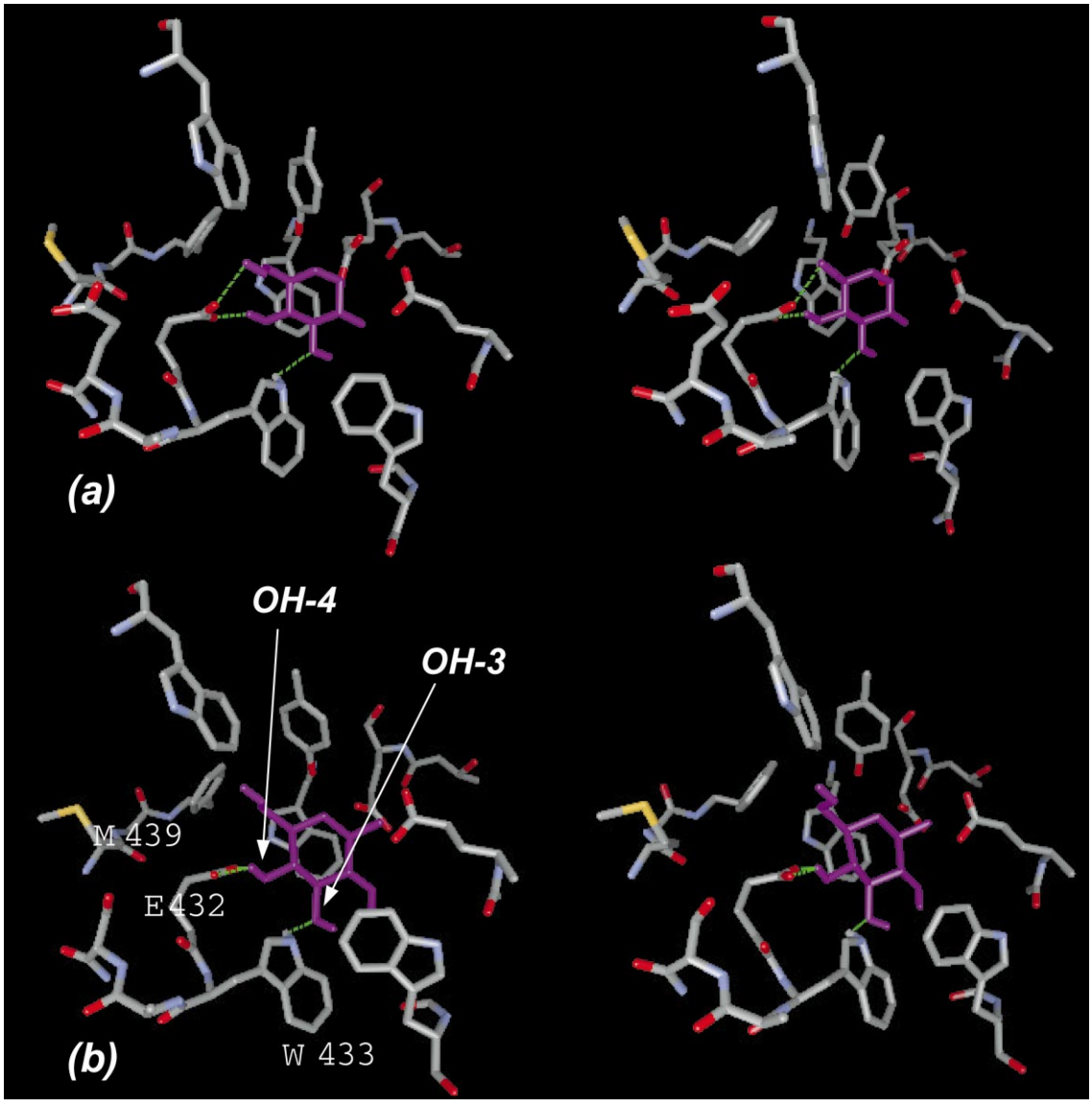



Fig. 1. Stereo view of structures indicating the critical binding of substrates in family 1 glycosyl hydrolases (a) BPβG-2-deoxy-2-fluoro-α-D-glucopyranosyl enzyme intermediate (1e4i); (b) SSβG (1gow) containing docked and minimised β-D-glucopyranose.

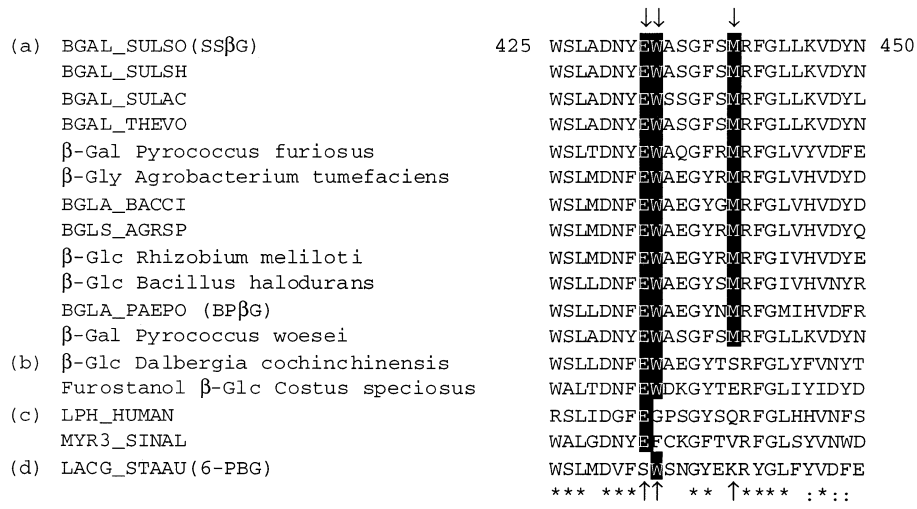



Fig. 2. Partial sequence alignment of the −1 binding pocket motif of SSβG [25] with high sequence similarity (left hand column gives SwissProt or TREMBL annotation, numbering is that of SSβG); glycosidases with similar substrate specificity (a) to SSβG and glycosidases with different and/or broadened specificities in which E432 (d), W433 (c) and M439 (b–d) differ (marked with arrow and highlighted) (*D. cochinchinensis* β-glucosidase [35]; *C. speciosus* furostanol-β-glycoside hydrolase [36]; LPH_HUMAN, human lactase phlorizin hydrolase [37]; MY3_SINAL, myrosinase [38]; LACG_STAAU (6-PBG), *S. aureus* 6-phosphogalactosidase [39]).

cosyl-enzyme intermediate (1e4i, Fig. 1a) have recently been reported. This allowed homology modelling and docking analysis of SSβG to create a minimum energy structure through molecular mechanics containing β-D-glucopyranose as a substrate mimic (Fig. 1b). Both the structures of BPβG and SSβG showed that the conserved (Fig. 2) residues E432 and W433 (SSβG numbering) create vital hydrogen bonds to the OH-4 and -3, respectively, of their substrates (Fig. 1). Furthermore, M439 sits at the base of the small side pocket that lies in close proximity to OH-6. Gratifyingly, sequence analysis (Fig. 2) supports the identification of the potential of these residues in specificity determination: e.g. S432 (SSβG numbering) rather than E432 in the phosphogalactosidase (EC 3.2.1.85) from *S. aureus*, [39] and G433 rather than W433 in the broad specificity glycosidase/cerebrosidase human lactase phlorizin hydrolase (EC 3.2.1.62) [37].

We therefore selected E432, W433 and M439 for mutagenesis as potentially critical active site residues for determining substrate specificity. Cysteine was chosen as the target residue for mutations, as a single flexible residue that could play a variety of roles. C behaves in proteins similarly to W and M [40], is structurally close to S but would alter some of the key interactions identified in Fig. 1 (e.g. abolish hydrogen bonding) in a conservative, informative manner.

Table 1
Kinetic parameters for SSβG enzymes at 80°C, pH 6.5

| Substrate | Enzyme, SSβG- | $K_M$ (mM) | $k_{cat}$ ($s^{-1}$) | $k_{cat}/K_M$ ($s^{-1}$ $mM^{-1}$) |
|---|---|---|---|---|
| 4-MUGlc | WT | 0.046 ± 0.017 | 140 ± 20 | 2900 |
| | E432C | 0.34 ± 0.07 | 5.1 ± 0.5 | 15 |
| | W433C | 1.61 ± 0.35 | 33 ± 5 | 20 |
| | M439C | 0.068 ± 0.028 | 190 ± 40 | 2900 |
| 4-MUGal | WT | 0.066 ± 0.017 | 98 ± 7 | 1490 |
| | E432C | 0.47 ± 0.14 | 5.4 ± 0.8 | 11 |
| | W433C | 2.2 ± 1.2 | 14 ± 6 | 6.3 |
| | M439C | 0.083 ± 0.016 | 94 ± 11 | 1130 |
| 4-MUFuc | WT | 0.011 ± 0.002 | 80 ± 2 | 7300 |
| | E432C | 0.34 ± 0.04 | 18 ± 1 | 53 |
| | W433C | 0.41 ± 0.09 | 31 ± 3 | 76 |
| | M439C | 0.023 ± 0.005 | 91 ± 8 | 4000 |
| 4-MUMan | WT | 0.036 ± 0.009 | 1.8 ± 0.2 | 50 |
| | E432C | 0.90 ± 0.26 | 2.8 ± 0.7 | 3.2 |
| | W433C | 0.18 ± 0.02 | 0.92 ± 0.05 | 5.1 |
| | M439C | 0.042 ± 0.015 | 2.3 ± 0.4 | 53 |
| 4-MUXyl | WT | 0.13 ± 0.03 | 3.8 ± 0.3 | 30 |
| | E432C | 1.26 ± 0.21 | 2.8 ± 0.3 | 2.2 |
| | W433C | 0.59 ± 0.19 | 1.5 ± 0.3 | 2.5 |
| | M439C | 0.068 ± 0.007 | 9.3 ± 0.2 | 136 |
| 4-MUGlcA | WT | 1.3 ± 0.4 | 0.81 ± 0.18 | 0.60 |
| | E432C | NAD[a] | NAD | NAD |
| | W433C | NAD | NAD | NAD |
| | M439C | 1.4 ± 0.6 | 1.3 ± 0.4 | 0.92 |

[a]No activity detected.

### 3.2. Construction and kinetic characterisation of WT and mutant enzymes

SSβG-WT, -E432C, -W433C and -M439C enzymes were expressed in *E. coli* as recombinant proteins containing an N-terminal $His_6$-tag to avoid interfering with the critical multimer-forming interactions of the C-terminus of the protein [36]. Yields of recombinant protein were of the order of 15 mg per l of culture. The purified, recombinant WT and mutated SSβG proteins gave single bands on SDS–PAGE at an indicated approximate molecular weight of 57 000, and gave a single peak on analysis by gel filtration under non-denaturing conditions, of an indicated molecular weight consistent with the formation of dimeric molecules (data not presented). Exact masses were confirmed by ESMS (±8 Da). Both WT and mutant recombinant SSβGs were >95% pure by these analyses.

Determination of the Michaelis–Menten parameters for the WT and mutant enzymes was performed at pH 6.5 at 80°C for a broad range of representative, fluorophore-containing 4-MU glycoside substrates, which allowed activities to be determined with a high degree of sensitivity (Table 1). Under these optimised assay conditions, the glucoside (Glc), galactoside (Gal) and fucoside (Fuc) substrates were hydrolysed well by SSβG-WT, but the xyloside (Xyl) substrate was hydrolysed relatively poorly (approximately 3% of turnover as determined by $k_{cat}$ compared with β-D-glucoside). Interestingly, low levels of previously undetected [30] β-D-mannoside (Man) and β-D-glucuronide (glucuronic acid, GlcA) activities (approximately 1% and 0.5% of turnover towards β-D-glucoside) were observed. In all cases the absolute D-stereospecificity and β-stereoselectivity of SSβG were maintained and no activity was detected towards L- or α-glycoside substrates.

It is apparent that the E432C and W433C mutations have a dramatic effect upon activity towards certain substrates. Glc $k_{cat}/K_M$ is reduced 200-fold and 140-fold, and Gal $k_{cat}/K_M$ is reduced 130-fold and 230-fold for E432C and W433C, respectively. However, although Man, Xyl activities were also reduced, these reductions were far less marked to $k_{cat}/K_M$ values only 10–16-fold lower than WT for E432C and W433C. Consistent with the prediction from Fig. 1 that hydrogen bonds to OH-4 (E432C) and OH-3 (W433C) are abolished in these mutants, these $k_{cat}/K_M$ decreases correspond to a loss of affinity of approximately 4.5–10.5 kJ/mol [41]. These reductions in $k_{cat}/K_M$ were largely manifested in reductions in ground state binding with $K_M$ values generally increased up to 37-fold; the greatest $K_M$ increases in both W433C and E432C were observed for Glc, Gal, Fuc. Variations in $k_{cat}$ in the mutants E432C and W433C were less uniform; there were large overall reductions in Gal, Glc turnover ($k_{cat}$ decreased by approximately 5–30-fold), whereas $k_{cat}$ values for Fuc, Xyl, Man in E432C and W433C are essentially similar to those for SSβG-WT (2-fold increased $k_{cat}$ (Man) for E432C to only 2.9-fold lowered $k_{cat}$ (Xyl) for W433C). This indicates that an additional transition state destabilisation is induced by mutation in E432C and W433C that essentially affects Gal, Glc only.

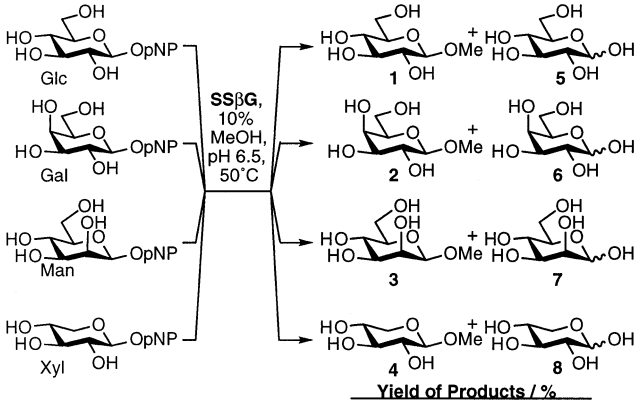
We were pleased to discover that as a result of the varying alterations in $k_{cat}/K_M$ for different substrates the specificities of E432C and W433C were remarkably more broad than SSβG-WT. For example, the variation of $k_{cat}/K_M$ for Glc:Gal:Xyl:Man moves from a restrictive 100-fold specificity range for WT to a broad 8-fold range for W433C (WT, 100:52:1:2 → W433C, 8.1:2.5:1:2).

The M439C mutation has a more subtle effect on specificity than the E432C and W433C mutations. Consistent with the ability of M439 to modulate substrate C-6 substituent specificity suggested by molecular modelling (Fig. 1b), the level of $k_{cat}/K_M$ alteration caused by mutation differs according to the C-6 structure. M439C shows almost identical values to WT for Gal, Glc, Man substrates in which the $CH_2OH$ at C-6 is unaltered. However, $k_{cat}/K_M$ for Fuc, which instead bears a $CH_3$ at C-6, is 1.8-fold lower than WT and excitingly, $k_{cat}/K_M$ for Xyl, which bears no C-6 substituent, is 4.7-fold higher than WT. It should also be noted that the mutation has the effect of increasing $k_{cat}$ for all the substrates, suggesting that a general stabilisation of the transition states is occurring.

It has been proposed previously that in other family 1 glycosidases the position corresponding to E432 in SSβG is responsible for the modulation of carbohydrate substrate O-6 substituent binding and in particular the rejection of negatively charged substituents [32]. In contrast to this prediction, the E432C mutant has no detectable activity towards GlcA, which at pH 6.5 bears a negative charge at C-6. In contrast, M439C shows slightly enhanced $k_{cat}/K_M$ values for GlcA (1.5-fold higher than WT), also consistent with modulation of C-6 substituent binding by M439.

### 3.3. Improved biocatalytic breadth of SSβG- and -W433C

Valuably, SSβG-WT's very high initial activity at 80°C resulted in enzymes that were still usefully active even after overall reductions in $k_{cat}/K_M$ caused by mutation to E432C and W433C. For example, W433C displays a $k_{cat}/K_M$ towards β-Gal substrates (6.3 $s^{-1}$ $mM^{-1}$) that compares well with the activity of recently described enhanced glycosynthases ($k_{cat}/K_M$ 0.013 $s^{-1}$ $mM^{-1}$) [12]. This activity coupled with greatly broadened specificity resulted in a synthetic utility for W433C




| SSβG | Time | Yield of Products / % 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|---|
| WT | 45 min | 40 | 33 | 2 | 3 | 14 | 8 | - | - |
| WT | 4h | - | - | 8 | 20 | 33 | 30 | 4 | 5 |
| E432C | 8h | 22 | 24 | 14 | 17 | 7 | 9 | 3 | 4 |
| W433C | 8h | 36 | 18 | 8 | 13 | 11 | 6 | 3 | 5 |

Scheme 1. Parallel glycoside syntheses using SSβG-WT, -E432C and -W433C as catalysts. The corresponding yields of products (each compound formed is labelled **1**–**8**) are shown in the table. These show that E432C and W433C mutants of SSβG, in which substrate specificity has been tailored, successfully produced balanced libraries of the four, desired β-glycosides of Glc (**1**), Gal (**2**), Man (**3**) and Xyl (**4**). Such balanced libraries are not produced by SSβG-WT even under varying reaction times.

and E432C that was demonstrated by the parallel synthesis of β-glycosides of Glc, Gal, Xyl, Man within in a one-pot mixture (Scheme 1). SSβG-WT was robust enough to catalyse transglycosylation at 50°C in 1:9 MeOH:buffer solutions, to form β-glycosides. However, its stringent specificity meant that after short periods (45 min) only glucoside **1** and galactoside **2** were formed and although small amounts of mannoside **3** and xyloside **4** were observed after extended periods (4 h), by this time all initially formed **1** and **2** had been hydrolysed. SSβG-WT is therefore incapable of creating libraries of glycosides in this way. We were therefore delighted to find that both W433C and E432C yielded mixtures of methyl Glc, Gal, Xyl, Man glycosides **1**–**4**. Indeed, the tailoring of E432C's specificity is so successful that it catalyses the formation of a small library of **1**–**4** in which each component is present in near equal amount. This balanced and similar yield of each of **1**–**4** mirrors the very similar $k_{cat}$ values (2.8–5.4 $s^{-1}$) of E432C for Glc, Gal, Xyl, Man substrates, an observation that is consistent with the high ($> K_M$) concentrations of substrates used in these reactions.

## 4. Conclusions

We have succeeded in tailoring the specificity of SSβG to create catalysts of broad synthetic utility. The handful of previous examples of substrate specificity alterations in glycosidases have only involved tailoring towards or away from functional groups such as $CH_2OH$ [18,19] or phosphate [20]. Excitingly, our results suggest that tailoring of stereospecificity is also possible. For example, alteration of a single residue W433→C effectively broadened the Gal:Man stereospecificity 25-fold from 29.4:1 in SSβG-WT to 1.2:1 in SSβG-W433C. Similarly, in the M439C mutant the sum of specificity alteration effects, including a 5-fold absolute increase in Xyl activity, causes a 10-fold increase in Xyl over Fuc specificity. The power of these mutant enzymes was further demonstrated by their utility in one-pot parallel syntheses of small arrays of glycosides that could not be accomplished with WT enzyme. Further applications of this enhanced synthetic utility including their apparent potential in combinatorial biocatalysis are currently being investigated.

*Acknowledgements:* We would like to thank Dr Hans Peter Klenk, University of Göttingen, for *S. solfataricus* MT4 genomic DNA, the EPSRC for the award of a quota studentship (KPC) and the BBSRC for JREI funding of mass spectrometry facilities.

## References

[1] Koeller, K.M. and Wong, C.H. (2000) Nat. Biotechnol. 18, 835–841.
[2] Davis, B.G. (2000) J. Chem. Soc. Perkin Trans. 1, 2137.
[3] Crout, D.H.G. and Vic, G. (1998) Curr. Opin. Chem. Biol. 2, 98–111.
[4] Wymer, N. and Toone, E.J. (2000) Curr. Opin. Chem. Biol. 4, 110–119.
[5] Watt, G.M., Lowden, P.A.S. and Flitsch, S.L. (1997) Curr. Opin. Chem. Biol. 7, 652–660.
[6] Kren, V. and Thiem, J. (1997) Chem. Soc. Rev. 26, 463–473.
[7] Palcic, M.M. (1999) Curr. Opin. Biotechnol. 10, 616–624.
[8] Scigelova, M., Singh, S. and Crout, D.H.G. (1999) J. Mol. Catal. B Enzym. 6, 483–494.
[9] Van Rantwijk, F., Woudenberg Van Oosterom, M. and Sheldon, R.A. (1999) J. Mol. Catal. B Enzym. 6, 511–532.
[10] Mackenzie, L.F., Wang, Q., Warren, R.A.J. and Withers, S.G. (1998) J. Am. Chem. Soc. 120, 5583–5584.
[11] Prade, H., Mackenzie, L.F. and Withers, S.G. (1998) Carbohydr. Res. 305, 371–381.
[12] Mayer, C., Zechel, D.L., Reid, S.P., Warren, R.A.J. and Withers, S.G. (2000) FEBS Lett. 466, 40–44.
[13] Malet, C. and Planas, A. (1998) FEBS Lett. 440, 208–212.
[14] Moracci, M., Trincone, A., Perugino, G., Ciaramella, M. and Rossi, M. (1998) Biochemistry 37, 17262–17270.
[15] Trincone, A. and Perugino, G. (2000) Bioorg. Med. Chem. Lett. 10, 365–368.
[16] Fort, S., Boyer, V., Greffe, L., Davies, G.J., Moroz, O., Christiansen, L., Schülein, M., Cottaz, S. and Driguez, H. (2000) J. Am. Chem. Soc. 122, 5429–5437.
[17] Nashiru, O., Zechel, D.L., Stoll, D., Mohammadzadeh, T., Warren, R.A.J. and Withers, S.G. (2001) Angew. Chem. Int. Ed. 40, 417–420.
[18] Zhang, J.-H., Dawes, G. and Stemmer, W.P.C. (1997) Proc. Natl. Acad. Sci. USA 94, 4504–4509.
[19] Andrews, S.R., Charnock, S.J., Lakey, J.H., Davies, G.J., Claeyssens, M., Nerinckx, W., Underwood, M., Sinnott, M.L., Warren, R.A.J. and Gilbert, H.J. (2000) J. Biol. Chem. 275, 23027–23033.
[20] Kaper, T., Lebbink, J.H.G., Pouwels, J., Kopp, J., Schulz, G.E., van der Oost, J. and de Vos, W.M. (2000) Biochemistry 39, 4963–4970.
[21] Rye, C.S. and Withers, S.G. (2000) Curr. Opin. Chem. Biol. 4, 573–580.
[22] Michels, P.C., Khmelnitsky, Y.L., Dordick, J.S. and Clark, D.S. (1998) Trends Biotechnol. 16, 210–215.
[23] Krstenansky, J.L. and Khmelnitsky, Y. (1999) Bioorg. Med. Chem. 7, 2157–2162.
[24] Barton, W.A., Lesniak, J., Biggins, J.B., Jeffrey, P.D., Jiang, J.Q., Rajashankar, K.R., Thorson, J.S. and Nikolov, D.B. (2001) Nat. Struct. Biol. 8, 545–551.
[25] Cubellis, M.V., Rozzo, C., Montecucchi, P. and Rossi, M. (1990) Gene 94, 89–94.
[26] Henrissat, B. (1991) Biochem. J. 280, 309–316.
[27] Moracci, M., Nucci, R., Febbraio, F., Vaccaro, C., Vespa, N., Cara, F.L. and Rossi, M. (1995) Enzym. Microb. Technol. 17, 992–997.
[28] Pisani, M.F., Rella, R., Raia, A.C., Rozzo, C., Nucci, R., Gambacorta, A., DeRosa, M. and Rossi, M. (1990) Eur. J. Biochem. 187, 321–328.
[29] Moracci, M., Capalbo, L., Ciaramella, M. and Rossi, M. (1996) Protein Eng. 9, 1191–1195.
[30] Nucci, R., Moraccci, M., Vaccaro, C., Vespa, N. and Rossi, M. (1993) Biotechnol. Appl. Biochem. 17, 239–250.
[31] Banner, D.W., Bloomer, A.C., Petsko, G.A., Phillips, D.C., Pogson, C.I., Wilson, I.A., Corran, P.H., Furth, A.J., Milman, J.D., Offord, R.E., Priddle, J.D. and Waley, S.G. (1975) Nature 255, 609–614.
[32] Aguilar, C., Sanderson, I., Moracci, M., Ciaramella, M., Nucci, R., Rossi, R. and Pearl, L.H. (1997) J. Mol. Biol. 271, 789–802.
[33] Bradford, M.M. (1976) Anal. Biochem. 151, 196–204.
[34] Voorhorst, W.G., Eggen, R.I., Luesink, E.J. and de Vos, W.M. (1995) J. Bacteriol. 177, 7105–7111.
[35] Cairns, J.R.K., Champattanachai, V., Srisomsap, C. and Svasti, J., TREMBL accession no. Q9SPK3.
[36] Inoue, K., Shibuya, M., Yamamoto, K. and Ebizuka, Y. (1996) FEBS Lett. 389, 273–277.
[37] Mantei, N., Villa, M., Enzler, T., Wacker, H., Boll, W., James, P., Hunziker, W. and Semenza, G. (1988) EMBO J. 7, 2705–2713.
[38] Xue, J., Lenman, M., Falk, A. and Rask, L. (1992) Plant Mol. Biol. 18, 387–398.
[39] Breidt, F. and Stewart, G.C. (1987) Appl. Environ. Microbiol. 53, 969–973.
[40] Nagano, N., Ota, M. and Nishikawa, K. (1999) FEBS Lett. 458, 69–71.
[41] Mader, M.M. and Bartlett, P.A. (1997) Chem. Rev. 97, 1281–1301.