

# **S<sub>N</sub>i from S<sub>N</sub>2: a Front-Face Mechanism ‘Synthase’**

## **Engineered from a Retaining Hydrolase**

*Javier Iglesias-Fernández<sup>1,2</sup>ψΦ, Susan M. Hancock<sup>3</sup>ψ, Seung Seo Lee<sup>3</sup>ψ¶, Moala Khan<sup>3</sup>, Jo Kirkpatrick<sup>3</sup>ς, Neil J. Oldham<sup>3</sup>‡, Katherine McAuley<sup>4</sup>, Anthony Fordham-Skelton<sup>5</sup>†, Carme Rovira<sup>1,2,6\*</sup> and Benjamin G. Davis<sup>3\*</sup>*

<sup>1</sup> Departament de Química Orgànica, Universitat de Barcelona, Martí i Franquès 1, 08028  
Barcelona, Spain

<sup>2</sup> Institut de Química Teòrica i Computacional (IQTCUB), Universitat de Barcelona, Martí i  
Franquès 1, 08028 Barcelona, Spain

<sup>3</sup> Department of Chemistry, University of Oxford, Chemistry Research Laboratory, Mansfield  
Road, Oxford OX1 3TA, UK

<sup>4</sup> Diamond Light Source, Diamond House, Harwell Science & Innovation Campus, Didcot,  
Oxfordshire, OX11 0DE, UK

<sup>5</sup> CLRC, Daresbury Laboratory, Warrington, Cheshire, UK

<sup>6</sup> Institució Catalana de Recerca i Estudis Avançats (ICREA), Passeig Lluís Companys 23,  
08010 Barcelona, Spain.

\* c.rovira@ub.edu; ben.davis@chem.ox.ac.uk

ψ These authors contributed equally.

† Deceased.

Φ Current address: Department of Chemistry, King's College London, London SE1 1DB, UK

¶ Current address: School of Chemistry, University of Southampton, Highfield, Southampton, SO17 1BJ UK.

‡ Department of Chemistry, University Park, Nottingham, NG7 2RD, UK

§ Current address: Leibniz Institute on Aging - Fritz Lipmann Institute (FLI), Beutenbergstraße 11, 07745 Jena, Germany

## Abstract (~350 words)

$S_{Ni}$  or  $S_{Ni}$ -like mechanisms, in which leaving group departure and nucleophile approach occur on the same 'front' face, have been observed previously experimentally and computationally in both the chemical and enzymatic (glycosyltransferase) substitution reactions of  $\alpha$ -glycosyl electrophiles. Given the availability of often energetically comparable competing pathways for substitution ( $S_{Ni}$  vs  $S_{N1}$  vs  $S_{N2}$ ) the precise modulation of this archetypal reaction type should be feasible. Here, we show that the drastic engineering of a protein that catalyzes substitution, a retaining  $\beta$ -glycosidase (from *Sulfolobus solfataricus* SS $\beta$ G), apparently changes the mode of reaction from " $S_{N2}$ " to " $S_{Ni}$ ". Destruction of the nucleophilic Glu387 of SS $\beta$ G-WT through Glu387Tyr mutation (E387Y) created a catalyst (SS $\beta$ G-E387Y) with lowered but clear transglycosylation substitution activity with activated substrates, altered substrate and reaction preferences and hence useful synthetic ('synthase') utility by virtue of its low hydrolytic activity with unactivated substrates. Strikingly, the catalyst still displayed retaining  $\beta$ -stereoselectivity, despite lacking a suitable nucleophile; pH-activity profile, mechanism-based inactivators and mutational analyses suggest that SS $\beta$ G-E387Y operates without either the use of nucleophile or general acid/base residues, consistent with an  $S_{Ni}$  or  $S_{Ni}$ -like mechanism. An x-ray structure of SS $\beta$ G-E387Y and subsequent metadynamics simulation suggest recruitment of substrates aided by a  $\pi$ -sugar interaction with the introduced Tyr387 and reveal a QM/MM free energy landscape for the substitution reaction catalyzed by this unnatural enzyme similar to those of known natural,  $S_{Ni}$ -like glycosyltransferase (GT) enzymes. Proton flight from the putative hydroxyl nucleophile to the developing *p*-nitrophenoxide leaving group of the substituted molecule in the reactant complex creates a hydrogen bond that appears to crucially facilitate the mechanism, mimicking the natural mechanism of  $S_{Ni}$ -GTs. An oxocarbenium ion-pair

minimum along the reaction pathway suggests a step-wise  $S_{Ni}$ -like  $D_N^*A_{Nss}$  rather than a concerted  $S_{Ni} D_N A_N$  mechanism. This first observation of a front face mechanism in a  $\beta$ -retaining glycosyl transfer enzyme highlights, not only that unusual  $S_{Ni}$  reaction pathways may be accessed through direct engineering of catalysts with suitable environments, but also suggests that ' $\beta$ - $S_{Ni}$ ' reactions are also feasible for glycosyl transfer enzymes and the more widespread existence of  $S_{Ni}$  or  $S_{Ni}$ -like mechanism in nature.

## Introduction (~500 words)

Since Sinnott and Jencks seminally demonstrated front-side (same face) nucleophilic attack in chemical,  $\alpha$ -glycosyl transfer substitution,<sup>1</sup> the possibility of the wider existence of such an unusual mechanism has been rarely but carefully considered.<sup>2, 3</sup> Such a front-side mechanism has been invoked to explain the seemingly unusual behavior of retaining glycosyltransferases (GTs).<sup>4</sup> Most retaining GTs do not contain obvious, conserved, functional nucleophiles and/or acid/base residues required to operate the double-displacement mechanism<sup>5</sup> that is found in glycoside hydrolases (GHs).<sup>4</sup> Whilst typically-observed 'chemical' nucleophilic substitution involves likely intermediacy of solvent exposed and accessible reactions centers, even for such reactions,  $S_Ni$ -like mechanisms, facilitated by assisted delivery of the nucleophile to the electrophile, are observed.<sup>6, 7</sup> In proteins, more constrained environments (and possible alternative pathways) exist. Structures of several retaining GTs<sup>8-11</sup> show positioning of substrates, leaving group and nucleophile in positions suitable for front-face mechanisms.<sup>2, 12</sup>

Recently, we have provided experimental evidence that supports the operation of a front-face mechanism in the retaining GT trehalose-6-phosphate synthase (OtsA)<sup>13</sup> consistent with detailed computational QM/MM metadynamics simulations.<sup>14</sup> These were followed by an experimental and computational study of glycosyl transfer in solution chemistry, indicating that the solvolysis of  $\alpha$ -glucosyl fluoride in hexafluoro-2-propanol, a non-nucleophilic environment, also follows a front-face mechanism.<sup>7</sup> Subsequent QM/MM studies on the retaining GTs lipopolysaccharyl  $\alpha$ -galactosyltransferase C (LgtC),<sup>15</sup>  $\alpha$ -1,2-mannosyltransferase Kre2p/Mnt1p,<sup>16</sup> polypeptide GalNAc-transferase T2 (GalNAc-T2)<sup>17, 18</sup> and glucosyl-3-phosphoglycerate synthase (GpgS)<sup>19</sup> have further contributed to disentangle the molecular details of the frontal face mechanism for these  $\alpha$ -selective retaining GTs.<sup>4</sup> Very recently, the functionally essential Notch-modifying xylosyltransferase has also been suggested to follow

this S<sub>N</sub>i-pathway.<sup>11</sup> Together these studies suggest that the unusual, front-face mechanism may, in fact, play an important and potentially widespread role in nature, when considering the importance and ubiquity of glycosyltransferases. Thus far, no β-selective retaining reaction has been observed. One apparent crucial feature of the α-selective mechanism suggested in these studies (**Scheme 1**) is the role of an asymmetric and shielding environment (the active site) as a ‘reaction compartment’ with sufficient space to not only accommodate the nucleophile and the leaving group on the same face but to do so in a protective manner that allows sufficient lifetime for oxocarbenium ion-like intermediates. In essence, the active site provides a ‘protective box’ that allows the acceptor nucleophile to separate the ion-pair that is generated from the donor electrophile.

Together these suggest common features (suitable shielding by active site moieties to exclude solvent; no competing protein nucleophile; reduced requirement for protein general acid/base; and suitable leaving group pK<sub>a</sub>) that, in principle, could be engineered rather than simply observed. Here we demonstrate that the front-face reaction is operative not only in retaining GTs but can also be created in engineered GHs through the exploitation of such features. Selection of a suitable, robust GH scaffold creates an enzyme with highly specific transglycosylation activity capable of stereospecific creation of β-glycosidic linkages from activated β-donors such as *p*-nitrophenyl glycosides, and incapable of hydrolyzing the unactivated glycosidic linkages in the product. Mechanistic investigations (including kinetic, biochemical, mutagenic, structural and computational studies) suggest that this novel, unnatural ‘synthase’ utilizes front-face nucleophilic substitution, similar to that proposed for retaining GTs. To the best of our knowledge, this is the first description of a frontal face mechanism of a β-retaining enzyme.

## Results and Discussion (~3200 words)

**Design and Creation of a Nucleophile-free GH.** The robust and representative GH family 1 scaffold was chosen as a protein platform for design. The retaining  $\beta$ -glycosidase from *Sulfolobus solfataricus* (SS $\beta$ G) has shown stability to mutation,<sup>20, 21</sup> solvents<sup>22</sup> and even under typically denaturing conditions.<sup>23, 24</sup> Prior nucleophile-free mutants bearing smaller residues than the natural Glu387 (e.g., Gly387<sup>25</sup>) have been shown to act as classical, *inverting* glycosynthases<sup>26-28</sup> with suitable ( $\alpha$ -glycosyl fluoride) substrates.<sup>25</sup> In contrast, our initial modeling suggested that to ensure sufficient protection and putative stabilizing interactions and yet small enough to be accommodated, only certain bulkier residues (e.g. Tyr, Phe) would prove suitable. Tyr387 was therefore chosen and site-directed mutagenesis of SS $\beta$ G-WT, yielded stable, folded, soluble protein SS $\beta$ G-E387Y, C-terminally-His-tagged to allow exhaustive nickel affinity chromatography (**Supplementary Figure 1**) giving good protein yields of ~28 mg per L of growth. N-terminal sequencing, LC-mass spectrometry (ESI-MS, found 57,450; expected 57,447 Da) (**Supplementary Table 1**) and circular dichroism (CD) analysis (**Supplementary Figure 2**) confirmed identity and unaffected secondary structure, respectively.

**The Glu387Tyr Nucleophile-Mutant Displays Altered Catalytic Activity.** By design, *para*-nitrophenoxide ( $\text{pK}_{\text{aH}} \sim 7$ )<sup>29</sup> with a similar  $\text{pK}_{\text{a}}$  to those of UDP ( $\text{pK}_{\text{aH1}} \sim 7$ ,  $\text{pK}_{\text{aH2}} \sim 9$ ) was chosen as a suitable leaving group for our putative ‘activated’ substrates. Determination of the kinetic parameters (**Table 1**) of SS $\beta$ G-E387Y towards *p*-nitrophenyl  $\beta$ -D-glycosides and comparison with SS $\beta$ G-WT revealed reduced but clear activity towards pNP $\beta$ Glc and pNP $\beta$ Gal substrates. Consistent with the loss of SS $\beta$ G-WT’s nucleophilic Glu387 residue, the decrease in activity was manifested exclusively in  $k_{\text{cat}}$ . Notably, substrate **selectivity** (as judged by

$k_{\text{cat}}/K_M$ ) was reversed from Gal:Glc = 1:1.6 in SS $\beta$ G-WT to 3:1 in Ss $\beta$ G-E387Y, a ratio that more closely reflects inherent, chemical reactivity of Gal vs Glc.<sup>30</sup> Interestingly, tyrosyl residues are observed in similar positions to Tyr387 in glycosidase enzymes that exploit substrate-assisted catalysis, such as the hexosaminidases.<sup>31</sup> These are thought to stabilize the formation of corresponding oxazolinium ion intermediates. However, Ss $\beta$ G-E387Y displayed no hexosaminidase activity either towards pNP $\beta$ GlcNAc or even corresponding activated oxazoline substrates (2-methyl-(1,2-dideoxy- $\alpha$ -D-glucofuranose)[2,1-d]- $\Delta^2$ -oxazoline) (**Supplementary Figure S3**). Consistent with the designed requirement for a suitable activated leaving group, Ss $\beta$ G-E387Y failed to hydrolyze either methyl  $\beta$ -D-galactopyranoside (MeGal) or *p*-nitrophenyl 6-O-( $\beta$ -D-galactopyranosyl)- $\beta$ -D-galactopyranoside (pNPGal1,6Gal).

Incubation with mechanism-based inhibitor<sup>32</sup> 2,4-dinitrophenyl 2-deoxy-2-fluoro- $\beta$ -D-galactopyranoside (**see Supplementary Methods and Discussion**) with no significant effect discounted the possibility of activity arising from Ss $\beta$ G-WT or other (e.g. endogenous expression host *E. coli*) glycosidases that use nucleophilic catalysis. It also intriguingly suggested that this altered catalytic activity of Ss $\beta$ G-E387Y was no longer nucleophile-dependent (*vide infra*). When Ss $\beta$ G-E387Y was thermally denatured (16-20h at 45°C) all activity was lost, implying that native protein conformation is required for its catalytic activity.

**SS $\beta$ G-E387Y is a 'Synthase'**. Given this striking selectivity for activated substrates, with negligible activity towards the hydrolysis of unactivated glycosides (and hence potential products), SS $\beta$ G-E387Y suggested itself as a potentially useful catalyst for glycosidic bond formation from activated *p*NP substrates. A range of representative monosaccharides as nucleophilic acceptors were surveyed under different conditions (**Table 2**).

Non-aromatic sugar acceptors were not processed to any significant extent by Ss $\beta$ G-E387Y, resulting in reactions that instead primarily gave Gal $\beta$ GalpNP disaccharidic products (**Table 2**, entries i-viii), suggesting a strong preference for utilizing GalpNP as an acceptor. This observed preference for aromatic sugar acceptors is consistent with aromatic stacking interactions in the + 1 or + 2 acceptor pockets used by the GH naturally for binding oligosaccharide substrates.<sup>33, 34</sup> Indeed, aromatic Gal $\beta$ , Glc $\beta$  and Man $\alpha$  glycosides all proved to be suitable nucleophile substrates (**Table 2**, entries ix-xii). Unlike several other synthases, under these conditions trisaccharides and higher or branched oligosaccharides (from uncontrolled 'self condensation') were not synthesized in measurable amounts; these are normally isolated in reactions catalyzed by classical glycosynthases<sup>35-37</sup> including, notably, a variant derived from Ss $\beta$ G.<sup>25</sup> Only under more extreme conditions were small amounts of trisaccharides observed (see below and **Supplementary Methods**). In all reactions, either exclusive 1,6- or 1,6-/1,3-linked regioselectivity was observed;<sup>38</sup> in contrast to the behavior of other Ss $\beta$ G-related catalysts,<sup>25, 39</sup> no 1,4-linked disaccharides were isolated. Notably, all transglycosylation reactions displayed exclusive, retentive  $\beta$ -stereoselectivity.

Having demonstrated initial synthetic potential, the synthetic application was explored in a model reaction of donor pNPGal with acceptor Ph $\beta$ Glc (**Supplementary Table S4**). Strikingly, variation of conditions allowed the improvement of the synthesis(S):hydrolysis(H) ratio to up to >99. Under these conditions, the enzyme is both selective and essentially, exclusively synthetic, yielding PheGlc1,6Gal as the predominant product in >70% isolable yield with only the formation of smaller amounts of trisaccharides as side products (**Table 2**, entry xiii). In control experiments under essentially identical conditions, Ss $\beta$ G-WT simply hydrolyzed the donor sugar and gave none of the desired synthetic product. No transglycosylation activity was observed using  $\alpha$ -D-galactopyranosyl fluoride donor and representative acceptors: Ss $\beta$ G-

E387Y does not process donor substrates with  $\alpha$ -anomeric configuration, thereby confirming that Ss $\beta$ G-E387Y does not act as a classical glycosynthase. Notably, in comparison to reactions catalyzed by glycosidases, which typically give transglycosylation yields from 20-40%,<sup>40</sup> the general yields of transglycosylation products synthesized with Ss $\beta$ G-E387Y (several > 80%) were high and only rivaled by some of the more potent glycosynthases.<sup>41</sup>

Although it should be noted that estimated transglycosylation rates ( $k_{\text{cat}}/K_M \sim 0.0052 - 0.025 \text{ min}^{-1}\text{mM}^{-1}$ ) are  $\sim 2,000$ -fold lower compared to classical glycosynthases (see below for further details).

**Mechanistic Analysis: Ss $\beta$ G-E387Y does not require a nucleophile or a general acid/base.** This useful transglycosylation / 'synthase' activity again highlighted the differing mechanism of Ss $\beta$ G-E387Y and suggested comparison with natural, trans-glycosidases. The trans-sialidase from *Trypanosoma cruzi* of GH family 33 utilizes a tyrosine residue as a nucleophile,<sup>42</sup> and although modeling and design (*vide supra*) had suggested incompatible geometries for Tyr387 in Ss $\beta$ G-E387Y to play this role, we attempted to clarify this aspect of its mechanism. First, to test Tyr387 as a catalytic nucleophile, trapping experiments were designed that were intended to yield a covalent intermediate from mechanism-based fluorosugar inactivators.<sup>32</sup> Thus, Ss $\beta$ G-E387Y was incubated with DNP-2FGlc<sup>32</sup> (1000 equivalents, 45°C, pH 6.5 50 mM sodium phosphate buffer) and analyzed by LC-MS (**Figure 1 and Supplementary Figure S3**). Over 6h, no change in Ss $\beta$ G-E387Y's hydrolytic activity was observed. Concomitant monitoring of DNP release (absorbance at 405 nm) revealed no acceleration over uncatalyzed chemical DNP-2FGlc hydrolysis. *Agrobacterium faecalis*  $\beta$ -glucosidase can form a stable  $\alpha$ -D-glucopyranosyl tyrosine product at non-relevant Y298 upon mutation of the active site nucleophile,<sup>43</sup> peptide 'mapping' did not show trapping of Tyr387.

Neither proteolytic (trypsin, pepsin, thermolysin, clostripain)-MSMS and/or CNBr-cleavage-MSMS (including neutral loss analysis of the 2FGlc moiety) indicated peptides with attached 2FGlc moieties (**Supplementary Figures S5,S6,S7**), even though the coverage of this 'mapping' successfully included peptides containing Y387 (and E206) as putative trapping sites. In control experiments, under essentially similar conditions, Ss $\beta$ G-WT was successfully labeled (**Supplementary Figure S8-S10**). Together these results suggested that Tyr387 (or even Glu206) was not acting as a catalytically nucleophilic residue in Ss $\beta$ G-E387Y (and that observed mass changes in the total protein MS were distributed non-specifically at low abundance over multiple non-specific locations that could not be detected by proteolytic-cleavage-MSMS analyses).

Next, to further probe the mechanism of Ss $\beta$ G-E387Y, and prompted by this apparent lack of any functioning nucleophilic catalytic residue, a range of representative mutants of Ss $\beta$ G were constructed (**Table 3**). Their identities (primary and secondary structure) were confirmed by ESI-MS (**Supplementary Table S1**) and CD analysis (**Supplementary Figure S2**).

None of these mutations caused a dramatic loss of function; indeed, the similar activities of Ss $\beta$ G-E387Y, -E387F, -E206A:E387Y, and -Y322F:E387Y suggested that none of these residues were necessary for the observed catalytic mechanism, i.e. none play a required role as a nucleophile or a general acid/base in their catalytic mechanisms. It is particularly notable that, consistent with the designed mechanism (*vide supra*) the additional mutation of the acid/base residue (Glu206) along with that of nucleophile (Glu387) to give Ss $\beta$ G-E206A:E387Y had no detrimental effect on activity; in the catalytic mechanism a general acid/base catalyst is also apparently not required, consistent with design (**Scheme 1**). This is also consistent with the observation that the basic limb of the pH profile of Ss $\beta$ G-E387Y was

also shifted ~0.6 pKa units to a value similar to that for *para*-nitrophenol (**Supplementary Figure S11**).

Finally, transglycosylation kinetics were determined for Ss $\beta$ G-E387Y with a range of substrates (**Supplementary Table S2 and Supplementary Figure 12**). Notably, both activity (as judged by  $k_{\text{cat}}/K_{\text{M}}$ ) and regioselectivity (1,6 vs 1,3, **Supplementary Figure 12b**) varied with leaving group; tentative linear free energy analysis (**Supplementary Figure 13**) reveals a small  $\beta$  value (-0.049), consistent with computational analysis suggesting a step-wise mechanism with a higher barrier for the collapse of oxocarbenium-ion intermediate than that for leaving group departure (*vide infra*).

**Structural Determinants of Catalysis in Ss $\beta$ G-E387Y.** To further probe the mechanism of Ss $\beta$ G-E387Y, the *apo* x-ray crystal structure of Ss $\beta$ G-E387Y was successfully determined (**Figure 2a and Supplementary Figure S13, Supplementary Methods and Supplementary Table S3**) and compared to the previously reported Ss $\beta$ G-WT structure.<sup>24</sup> Despite the mutation, the structures can be superimposed with very little divergence; the r.m.s. deviation is 0.26 Å as calculated using 486 C $\alpha$  positions. Essentially in the active site, only 2 amino acids have shifted significantly as a result of the mutation *i.e.* Tyr322 and His342 (**Figure 2a**). Attempts to generate *holo* structures in complex with either substrate or inhibitor were unsuccessful. Therefore, the structures of appropriate ternary complexes were modeled informed by both the *apo* Ss $\beta$ G-E387Y structure and structural alignments with Ss $\beta$ G-WT<sup>44</sup> complexed with D-galactohydroximolactam (pdb: 1uwt) (**Supplementary Fig. S14**). The Ss $\beta$ G-E387Y active site is very similar to that of Ss $\beta$ G-WT (**Figure 2a**), consistent with the similar  $K_{\text{M}}$  values obtained for pNP $\beta$ Gal and pNP $\beta$ Glc substrates for Ss $\beta$ G-WT and Ss $\beta$ G-E387Y Ss $\beta$ G (**Table 1**).

A combination of classical molecular dynamics and metadynamics techniques were used to model a ternary Michaelis complex of Ss $\beta$ G-E387Y with two molecules of pNP $\beta$ Gal, as putative acceptor and donor substrates corresponding to one of the observed synthase activities (*vide supra*). In a first step, the two molecules were manually placed at the entrance of the enzyme catalytic groove (see **Supplementary Methods**). After 200 nanoseconds of molecular dynamics (MD) simulation, one of the molecules partially entered the catalytic site, sitting at  $\sim 8$  Å from the catalytic residues, whereas the other remained at the entrance (**Supplementary Figure S16a**). Further MD simulation did not lead to significant change, indicating that complete entrance of the two molecules is associated with a certain free energy barrier. Therefore, the ligand binding process was activated using an enhanced-sampling technique (metadynamics).<sup>45</sup> Two collective variables were chosen to drive the binding of the two pNP $\beta$ Gal molecules to the active site of Ss $\beta$ G-E387Y. The first (CV<sub>1</sub>, **Supplementary Figure S17 and Supplementary Discussion**) measures the degree of penetration of the first pNP $\beta$ Gal molecule (as the *donor*) into the active site; the second (CV<sub>2</sub>) accounts for the formation of a O1 $\cdots$ H' interaction, i.e. it measures the distance between the *donor* and *acceptor* molecules.

The free energy landscape (FEL) of ligand binding obtained from the classical metadynamics simulation (**Supplementary Figure S16c**) shows an energy minimum (the global one) in which the two pNP $\beta$ Gal molecules are inside the enzyme active site (the ternary complex, shown in **Figure 2c**). Analysis of the water content around the active site shows that a number of water molecules are displaced during binding ( $13 \pm 4$  from a region of  $\leq 5$  Å from Y387 and Y322). Among the remaining water molecules, there are two that are located within 5 Å of the donor anomeric carbon. Although these water molecules are not well oriented for catalysis, they could account for the observed residual hydrolysis. Close examination of the

orientation of the two molecules in the active site reveals that the hydroxymethyl group of the acceptor molecule is located on the same face of the donor sugar as the *p*-nitrophenyl group (i.e. the leaving group) of the donor molecule. This is an optimum topology for a front-face mechanism, which could ultimately lead to a transglycosylation product with net retention of configuration. The terminal hydrogen atom of the acceptor hydroxymethyl group points towards the glycosidic oxygen of the donor molecule, favoring the formation of a 1,6-glycosidic linkage, consistent with the observed regiochemical preferences of SsβG-E387Y. This hydrogen bonding interaction may provide a guide for the nucleophile to the same face as the leaving group, akin to interactions observed in retaining “S<sub>N</sub>i-like” GTs.<sup>14, 15</sup> Furthermore, this is consistent with the intended, designed role of the leaving group glycosidic oxygen as a general base that deprotonates the incoming protic OH-6-hydroxyl (**Scheme 1**). It is also consistent with the non-detrimental effect on activity of the removal of the general acid/base residue (Glu206) in SsβG-E387Y:E206A; in SsβG-E387Y with *p*NPGal the phenolic base appears sufficient to deprotonate the incoming hydroxyl nucleophile.

There are crucial substrate-protein interactions (**Figure 2b**) that contribute to the stability of the above “front-face arrangement”. First of all, Tyr387 forms stabilizing *donor* sugar···π interactions<sup>46</sup> (sugar hydrogen atoms point towards the center of the Y387 phenol ring, with distances < 3 Å, **Figure 2c**), consistent with the overlay of the starting *apo* SsβG-E387Y x-ray crystal structure with the SsβG-WT•inhibitor complex (**Supplementary Figure S14**). Second, Tyr322 has swung to form π···π stacking interactions with the *acceptor* *p*NPGal moiety (the distance between carbon atoms of both six-membered rings amounts to ~ 3.5 Å). This, in turn, appears to position the OH-6-hydroxyl group in an optimum orientation to attack the anomeric carbon of the sugar donor. These π···π stacking interactions explain why *p*NPβGal and other aromatic glycosides are preferred substrates for the synthase activity of SsβG-E387Y (*vide*

*supra*). Essentially identical analysis of a possible O3-regioselective pathway also generated an appropriate Michaelis complex (**Supplementary Figure 18**). The binding modes corresponding to the 1,6- or 1,3-reaction are quite different, especially for the acceptor molecule. However, notably, in *both* cases (1,6 and 1,3), the *donor* sugar is stabilized by CH $\cdots\pi$  interactions engendered by Y387. In the corresponding 1,3- pathway the major difference is that in the *acceptor* the aglycon is oriented away from Y322 enabling sugar-CH $\cdots\pi$  interactions between acceptor and donor (c.f. *acceptor* aglycon  $\pi\cdots\pi$  interactions with Y322 for the 1,6-, see above). Thus, in both cases  $\pi\cdots\pi$  and sugar $\cdots\pi$  interactions stabilize the substrates in optimum orientation for catalysis. Together these structural analyses (x-ray structure and metadynamics simulations of ligand binding) suggest clearly that the donor anomeric carbon is spatially accessible to the acceptor OH-6 or OH-3 hydroxyl groups from the ‘front face’.

***QM/MM Analysis of the Mechanism and Reaction Landscape of Ss $\beta$ G-E387Y.*** QM/MM simulations, using the metadynamics approach, were performed to elucidate precise details of this unusual glycosyl transfer reaction at atomic detail and to obtain the free energy landscape from which, in turn, reaction coordinates can be defined. From the ternary complex determined above (**Figure 2b,c**) three collective variables, corresponding to the main bonds undergoing breaking or formation, were used (**Supplementary Fig. S15 and Supplementary Discussion**). As a test of one of the critical design elements in this “S<sub>N</sub>I-synthase”, it is important to note that none of the CVs used ‘self-select’ any specific reaction pathway. The free energy landscape for the transglycosylation reaction, reconstructed from the QM/MM metadynamics simulation (**Figure 3a**) shows three main minima and two transition states (TS). The free energy difference between the reactants state and the highest TS amounted to ~ 25

kcal.mol<sup>-1</sup>, similar to the value obtained for the OtsA glycosyltransferase with essentially similar computational methodology.<sup>14</sup>

The structure of the reactants complex (**R** in **Figure 3c**) is very similar to the one from classical (i.e. force-field based) metadynamics simulation (**Figure 2c**), except that the donor galactosyl ring is distorted into a <sup>1</sup>S<sub>3</sub> conformation in the QM/MM structure as opposed to a relaxed <sup>4</sup>C<sub>1</sub>. This is not surprising in view of the known limitations of force-fields to describe the precise conformation of the sugar ring in glycoside hydrolases.<sup>47, 48</sup> The more detailed QM/MM metadynamics simulations instead support a distorted conformation for the saccharide ring at the -1 donor enzyme subsite, essentially similar to that expected for a β-glucoside hydrolase mechanism.<sup>49, 50</sup> Of particular interest is the hydrogen bond between the hydroxymethyl group of the acceptor molecule and the leaving group (*p*NP) of the donor molecule in the reactants complex. This type of interaction, previously observed on the basis of QM/MM calculations for GTs<sup>14, 15, 17, 18</sup> (the hydrogen bond forms either at the reactants complex or in the early stages of the reaction), is a common feature of enzymes operating via a front-face mechanism and part of the design invoked for SsβG-E387Y (**Figure 1** and *vide supra*).

The reaction pathway (**Figure 3**) starts with the elongation of the C1-O1 bond of the donor molecule (the C-O distance increases more than 1 Å when going from **R** to **1**, **Figure 3b** and **Supplementary Table S5**). This bond is completely broken at intermediate **2** (C1-O1 = 3.4 Å). At this point of the reaction, the distance between the donor and the acceptor (C1⋯O6') is still long (~3 Å), indicating the formation of an oxocarbenium–phenoxide ion pair. Further evidence for the change in electronic configuration at the anomeric carbon atom is the shift to a trigonal geometry, which is also associated with changes in the conformation of the pyranose ring along the reaction (see **Figure 3d** and discussion below). This change is accompanied by a

decrease in the C1-O5 bond length (from 1.41 Å to 1.27 Å, Table 5) and an increase of the charge of the anomeric center (by 0.30 e<sup>-</sup> when going from **R** to **2**).

The oxocarbenium ion-pair corresponds to a minimum along the reaction pathway. It is stabilized by the O6'-H···O1 hydrogen bond (**2** in **Figure 3c**), which has also a role in orienting the acceptor sugar for the subsequent nucleophilic attack. Afterwards, a slight displacement of the hydroxymethyl moiety coupled to a proton transfer, from the hydroxymethyl to the *p*NP leaving group, forms the new glycosidic bond (**3** → **P** in **Figure 3a**). Notably, the observation of a slightly higher barrier ~3 kcal/mol for collapse of the oxocarbenium ion is not only consistent with prior observations in GTs<sup>14, 17</sup> but also with the low β<sub>lg</sub> determined experimentally (see above). As a further characterization of this species, we extracted two snapshots of the metadynamics simulation that correspond to minimum **2** and performed geometry optimizations and subsequent QM/MM MD simulations (see **Supplementary Methods**). The ion-pair species was stable under optimization and MD simulation with a lifetime > 15 ps. This again indicates that the ion-pair species is a minimum of the free energy landscape. Interestingly, *in silico* mutation of Y387 to F387 generates an oxocarbenium-ion species that is still a stable minimum, with a slightly longer distance between the aryl ring and the sugar donor anomeric carbon compared with the E387Y variant. This is consistent with the experimental findings that the E387F variant still exhibits clear activity (**Table 3**). An alternative mechanism in which the oxocarbenium ion collapses with the E206 acid base residue,<sup>51</sup> was also considered and tested (**Supplementary Methods** and **Supplementary Figure S19**). However, this mechanism was discarded in view of the high-energy barrier obtained and the low stability of such an intermediate. Therefore, the simulation shows that cleavage of the donor Gal-β-*p*NP bond and formation of the Galβ1,6Gal bond are entirely asynchronous and follow a front-side stepwise mechanism.

The donor conformational itinerary observed in Ss $\beta$ G-E387Y during transglycosylation (**Figure 3c**) was:  ${}^1S_3$  (reactants) –  ${}^4H_3/E_3$  (reaction intermediate) –  ${}^4C_1$  (products). This pathway is the same predicted experimentally<sup>49, 50, 52</sup> and theoretically<sup>53</sup> for retaining  $\beta$ -D-gluco-active glycoside hydrolases such as Ss $\beta$ G-WT. Remarkably, therefore, despite the very different mechanism, the engineered ‘S<sub>N</sub>i-synthase’ Ss $\beta$ G-E387Y synthesizes glycosidic bonds by exploiting essentially the same conformational itinerary (and associated distortional strategies to guide catalysis) used by the WT enzyme for hydrolysis. This suggests that, independent of the type of reaction catalyzed by the enzyme, the active site serves as a ‘box’ for the donor to accommodate a given reduced set of pyranose ring conformers.

## Conclusions (~500 words)

Until now, frontal face or S<sub>N</sub>i-like mechanisms have only been implied in retaining  $\alpha$ -glycosyltransferases; the engineered system we present here constitutes an example of a retaining glycosyltransferase-like enzyme with  $\beta$ -glycosidic bond selectivity. Structural and computational analyses support a critical role for the installed Tyr387 through sugar- $\pi$  and  $\pi$ - $\pi$  interactions in recruiting to the Michaelis complex (**Figure 2c**) and in stabilizing the reaction pathway through the formation of a hydrogen bond between the acceptor OH and the donor glycosidic oxygen. Given that the dehydroxylating Tyr $\rightarrow$ Phe mutation in Ss $\beta$ G-E387F does not affect activity, it suggests that any such stabilization might not be (entirely) *via* interactions with the hydroxyl group and/or is not dramatically altered by the change in  $\pi$ -density that this would also cause; this slight effect is supported by computation. Mutagenesis of an analogous tyrosine to phenylalanine in human cytosolic  $\beta$ -glucosidase, caused only a 2-5 fold decrease in  $k_{cat}$ , with minimal effect on  $K_M$ ; this too suggested that a polarisable  $\pi$ -aromatic ring system might have the capacity for transition state stabilization.<sup>54</sup> Free energy landscape analyses

show some shortening of the sugar-phenol distances  $\sim 0.5$  Å at the point of ion pair formation, consistent with  $\pi$ -cation stabilization, albeit at a distance  $\sim 5$ -6 Å. Consistent with this reasoning, the aromatic residues (Tyr or Phe) at position 387 were found to be essential for activity: removal of the aromatic group by mutagenesis to Ala in Ss $\beta$ G-E387A resulted in a protein with no activity (**Table 4**).

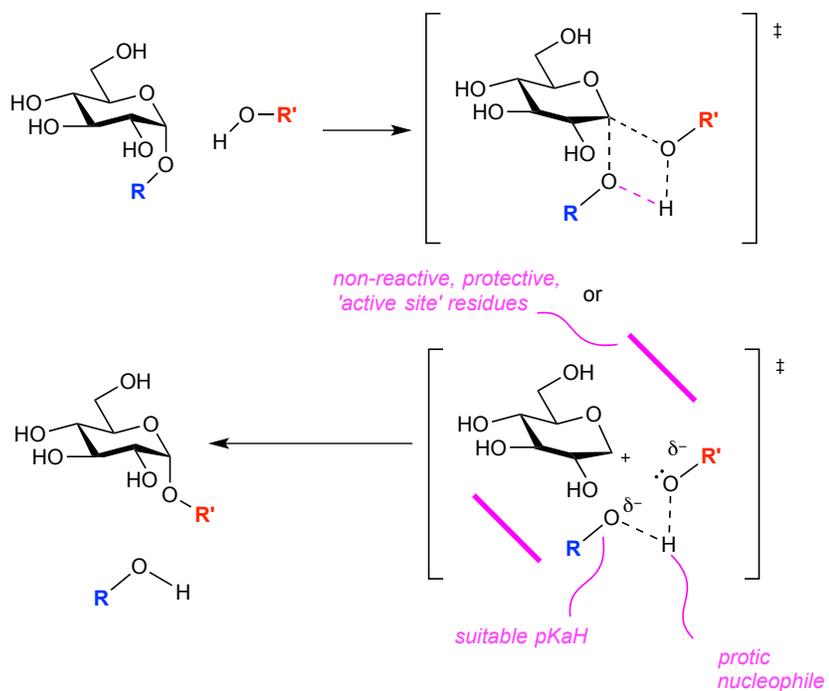
The front-face mechanism therefore appears to proceed *via* an oxocarbenium ion-pair intermediate that, due to the greater steric bulk of the active site upon tyrosine introduction, is largely prevented from reacting with water to give the hydrolysis product. Instead, an acceptor bound in the +1 subsite, preferentially stabilized by the relocated Tyr322 residue, attacks the carbocation. The enzyme scaffold provides a shaped 'protein box' (primarily for the donor) devoid of any catalytic residue but that nonetheless provides stabilization and specifies that reactants can only form  $\beta$ -products. This reactivity and selectivity is provided (at least in part) by the box's favoring of particular conformers along the corresponding itinerary (**Figure 3c**). Such a 'box' is highly reminiscent of the catalytic activity proposed for serine protease mutants that, although lacking their entire catalytic triad, nonetheless show rate accelerations of  $\sim 10^3$ -fold over background.<sup>55</sup> Notably the 'box' provided by catalytic antibodies that act as glycosidases<sup>56</sup> that also lack participating residues are similarly highly hydrophobic and, indeed, less efficient (rate accelerations of  $\sim 10^3$ -fold over background;  $k_{cat}$  0.007 min<sup>-1</sup>,  $K_M$  0.53 mM) than the designed 'S<sub>N</sub>i synthase' that we have created here (rate accelerations of  $\sim 10^5$ -fold over background;  $k_{cat}$  0.48 min<sup>-1</sup>,  $K_M$  0.17 mM). It should be noted that our 'S<sub>N</sub>i synthase' is, in turn, a similar magnitude less active than prior 'S<sub>N</sub>2 synthases'. Further future activity optimization might be considered, through forced evolution strategies, for example. Given the previously suggested<sup>57</sup> 'conceptual kinship' of some glycosyl units and terpenes it is interesting to note that our initial inspection of known structures of terpene cyclase structures

suggests prominently placed aromatic sidechains, akin to the Y387 that we have discovered here.<sup>58</sup> Altogether, these results suggest that the, once seemingly improbable and rare, same-face nucleophilic substitution is a viable mechanistic possibility in many ways in nature and can be considered a viable accessible mechanism in the design of catalysts for substitution.<sup>59</sup>

**Author Contributions.** JIF designed and performed calculations. SMH, SSL, MK performed the biochemical experiments. SMH, KM, AF-S determined x-ray structures. All authors analyzed results. CR, SSL, BGD wrote the manuscript. All authors except AF-S read and commented on the manuscript.

**Acknowledgements.** We thank the EPSRC and High Force Research (SMH), the BBSRC (SSL), MINECO (grant CTQ2014-55174 to CR) and AGAUR (grant and 2014SGR-987 to CR) for funding. BGD was a Royal Society Wolfson Research Merit Award recipient during the course of this work. We acknowledge the computer support provided by the Barcelona Supercomputing Center (BSC-CNS). This paper is dedicated to the memory of Tony Fordham-Skelton, a friend, mentor and comrade who is still very much missed.

**Supporting Information Available:** Supporting Methods, Figures, Tables and Discussion. This material is available free of charge via the Internet.

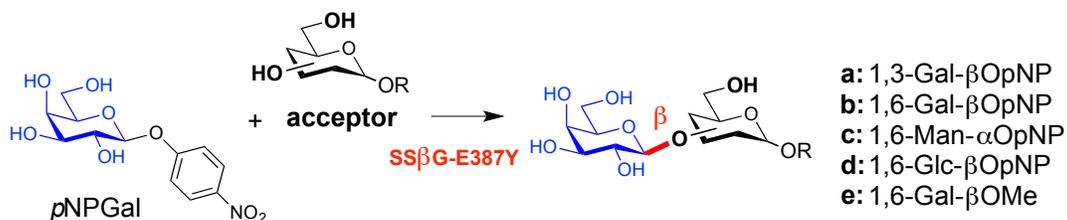


**Scheme 1.** Front-face reaction mechanism of  $\alpha$ -selective retaining glycosyltransferases.

**Table 1. Hydrolytic kinetic parameters for SsβG-E387Y.** Sodium phosphate buffer (50 mM, pH 6.5) at 45°C. Assays were initiated by adding enzyme (10 μL) to substrate (190 μL) and *p*-nitrophenolate (*p*NP) release was monitored at 405 nm on a 96-well plate reader. na = no detectable activity

<b>SsβG</b>	<b>Substrate</b>	<b><math>k_{cat} / s^{-1}</math></b>	<b><math>K_M / mM</math></b>	<b><math>k_{cat}/K_M / M^{-1}s^{-1}</math></b>
WT	<i>p</i> NPGal	7.20 ± 1.06	0.57 ± 0.07	11140
WT	<i>p</i> NPGlc	4.35 ± 0.53	0.15 ± .0.01	17777
E387Y	<i>p</i> NPGal	0.008 ± 0.0011	0.17 ± 0.02	44.4
E387Y	<i>p</i> NPGlc	0.002 ± 0.0004	0.17 ± 0.02	14.9
E387Y	<i>p</i> NPGlcNAc	na	na	na
E387Y	<i>p</i> NPGal1,6Gal	na	na	na
E387Y	GlcNAc-Δ <sup>2</sup> -oxazoline	na	na	na
E387Y	MeGal	na	na	na

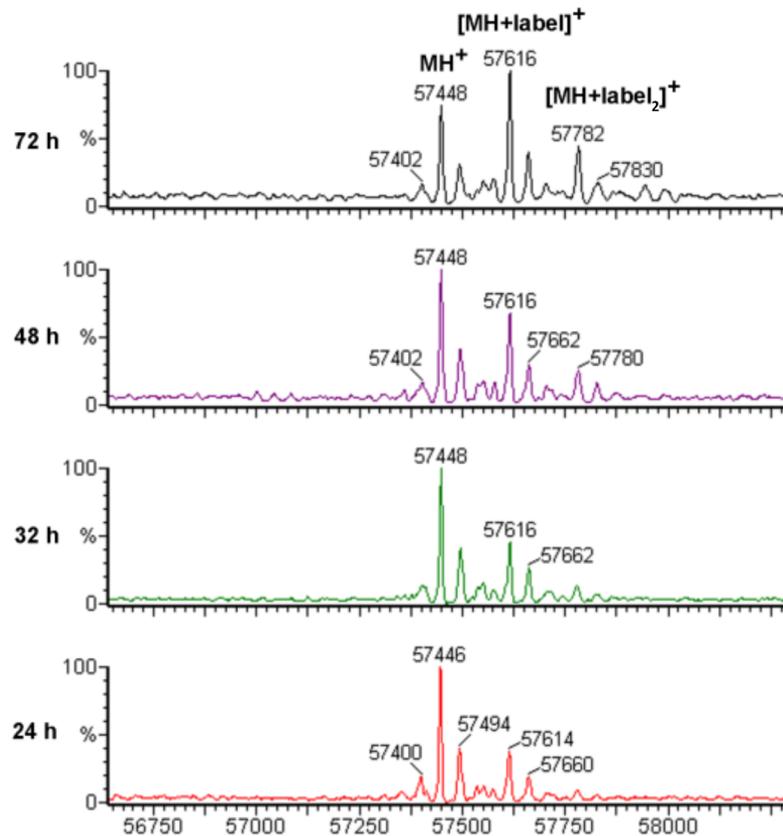
**Table 2. Ss $\beta$ G-E387Y catalyzes transglycosylation.** Disaccharides synthesised from *p*NPGal as a glycosyl donor.



	Acceptor	Temp / °C	Yield / % <sup>[a]</sup>								S / H	Conversion <sup>[d]</sup> / %
			a	b	c	d	e	H <sup>[b]</sup>	S <sup>[c]</sup>	Total		
i	Me $\beta$ Gal	45	18	24	-	-	2	37	44	81	1.2	92
ii	Me $\beta$ Gal	80	51	36	-	-	1	<1	88	88	>88	78
iii	cellobiose	45	14	15	-	-	-	44	29	73	0.7	100
iv	cellobiose	80	22	27	-	-	-	6	49	55	8.2	79
v	lactose	45	21	29	-	-	-	33	50	75	2.0	80
vi	lactose	80	30	54	-	-	-	16	84	100	5.3	91
vii	Me $\beta$ Man	45	16	38	-	-	-	46	54	100	1.2	100
viii	Me $\beta$ Man	80	39	46	-	-	-	15	85	100	5.7	92
ix	Ph $\beta$ Glc	45	9	46	-	26	-	17	81	98	4.8	97
x	Ph $\beta$ Glc	80	0	28	-	12	-	37	-	-	-	100
xi	Ph $\alpha$ Man	45	0	3	12	-	-	85	15	100	0.2	100
xii	Ph $\alpha$ Man	80	1	10	25	-	-	64	36	100	0.6	100
xiii	Ph $\beta$ Glc <sup>[e]</sup>	45	5	0	-	72	-	<1	>99	100	>100	_ <sup>[e]</sup>

<sup>[a]</sup> Yields were determined by NMR analysis of the per-acetylated reaction mixture, separated by flash chromatography and based on the recovery of starting material. Reaction times were determined by period of catalytic activity i.e. until no further progression ~15h or longer. <sup>[b]</sup> Total yield of hydrolysis products. <sup>[c]</sup> Total yield of glycosides/synthesis products. <sup>[d]</sup> based on

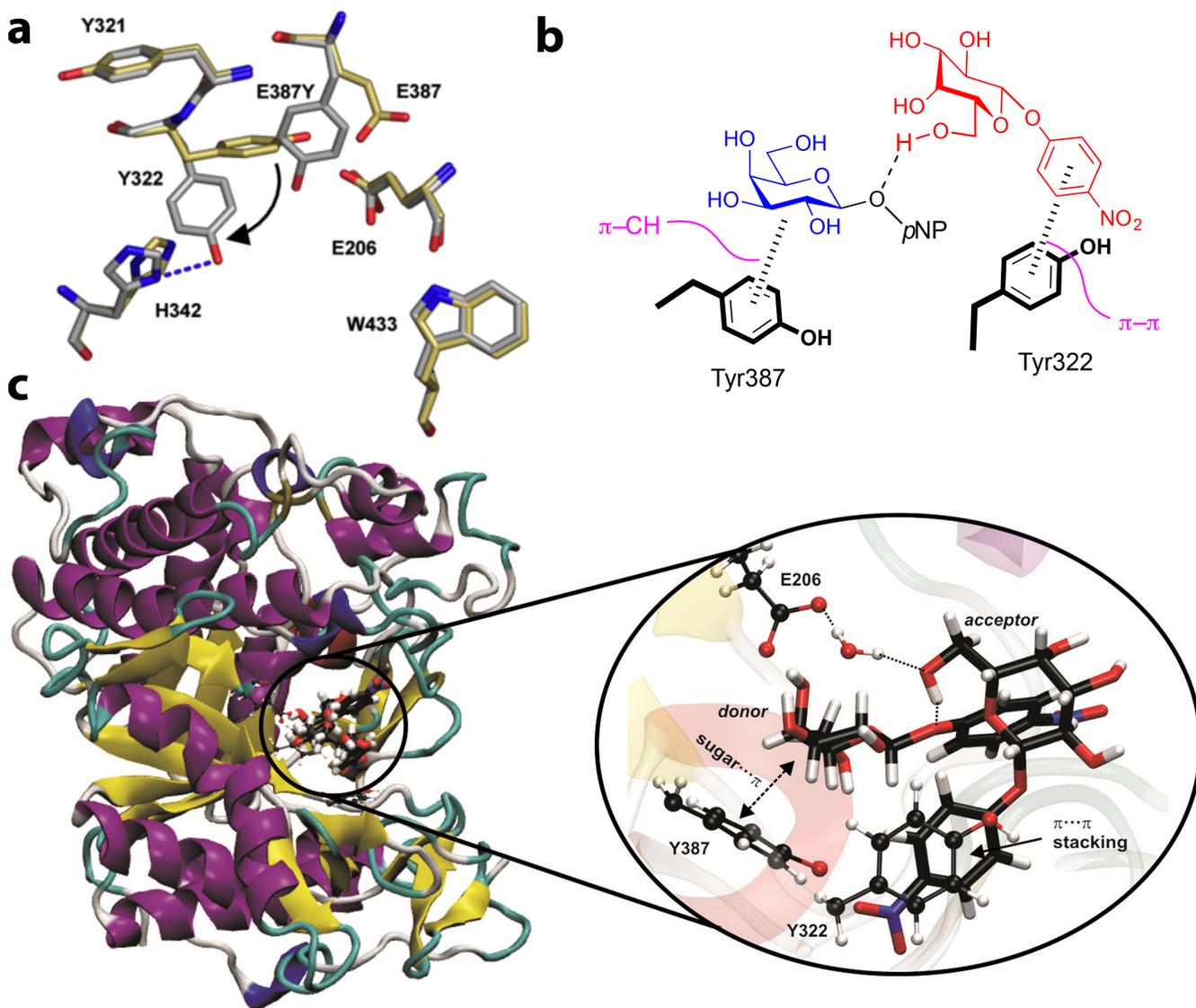
the consumption of starting material <sup>[e]</sup> After optimization for yield, including additional production of trisaccharide as mass balance – see **Supplementary Table S4**.



**Figure 1. Incubation of SsβG-E387Y with Covalent Inhibitor DNP-2FGlc.** Reaction with 2,4-dinitrophenyl 2-deoxy-2-fluoro-β-D-glucopyranoside was monitored over time by ESI-MS. Slow reaction and emergence of additional peaks ( $2 \times +165 \pm 3$  Da etc) after extended incubation and with an apparent statistical distribution suggest non-specific chemical modification; incubation with 2FGlc did not cause direct glycation (**Supplementary Figure S3**). (either directly or likely following uncatalyzed chemical DNP-2FGlc hydrolysis and glycation). This non-specific, non-‘activity-based’ cause is also consistent with the thermal denaturation of SsβG-E387Y at 45°C >16h (*vide supra*).

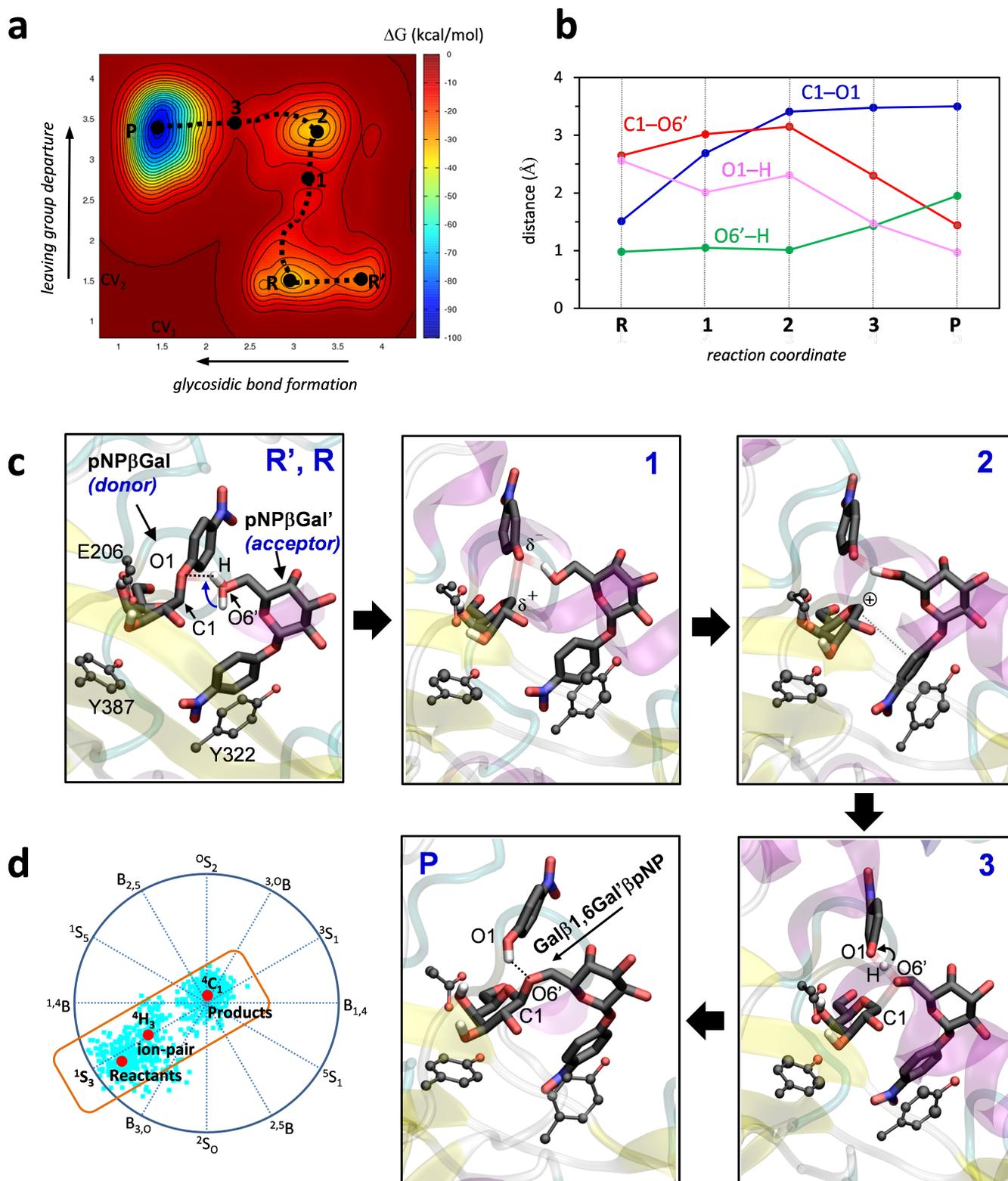
**Table 3. Hydrolytic kinetic parameters for Ss $\beta$ G-mutants.** Hydrolysis of pNPGal at 45°C in 50 mM phosphate buffer, pH 6.5. na = no observable activity.

<b>Ss<math>\beta</math>G variant</b>	<b><math>k_{\text{cat}} / \text{s}^{-1}</math></b>	<b><math>K_{\text{M}} / \text{mM}</math></b>	<b><math>k_{\text{cat}} / K_{\text{M}} / \text{M}^{-1}\text{s}^{-1}</math></b>
E387Y	0.008 $\pm$ 0.001	0.17 $\pm$ 0.02	45
E387F	0.005 $\pm$ 0.0004	0.07 $\pm$ 0.02	79
E387A	na	na	na
E206A:E387Y	0.011 $\pm$ 0.001	0.10 $\pm$ 0.03	106
Y322F:E387Y	0.007 $\pm$ 0.0003	0.08 $\pm$ 0.01	79



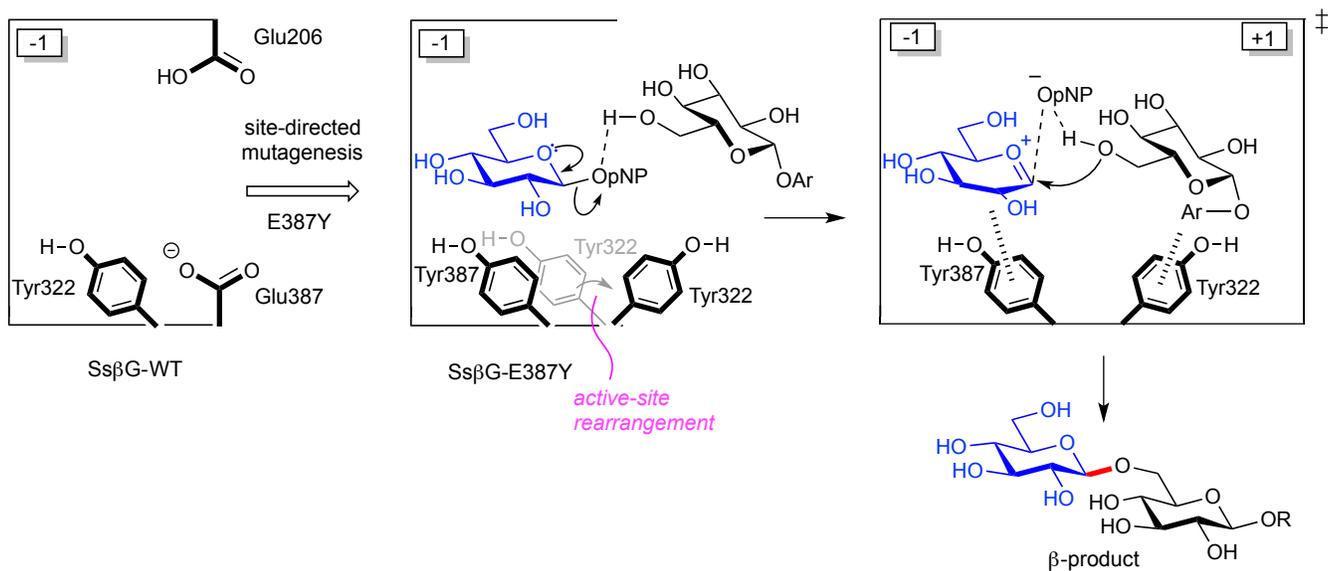
**Figure 2. Structural Analysis of SS $\beta$ G-E387Y.** (a) The X-ray structure of *apo* SS $\beta$ G-E387Y (determined in this work: pdb 5i3d, silver) superimposed on SS $\beta$ G-WT (pdb: 1gow, gold) shows the highly localized rearrangement (indicated by curled black arrow) of residues Y322 and H342 to accommodate the changed residue at 388 (E387Y). The hydroxyl of Y322 is within  $\sim 3.1$  Å of the N $\delta$ 1 of H342, suggesting that a hydrogen bond stabilizes this amino acid side chain migration (blue dashes). Essentially negligible alterations are observed in the rest of the structure. (b) Schematic interaction diagram of proposed substrate-protein interactions based on (a) and (c): Y387 forms stabilizing *donor* sugar $\cdots\pi$  interactions<sup>46</sup> (sugar hydrogen atoms point towards the center of the Y387 phenol ring, with distances  $< 3$  Å, see (c)); the

localized Y322 rearrangement creates  $\pi\cdots\pi$  stacking interactions with the *acceptor* pNPGal moiety. This, in turn, positions the acceptor OH-6 in an orientation to attack the anomeric carbon of the sugar donor. **(c)** Structure of Ss $\beta$ G-E387Y in complex with two pNP $\beta$ Gal molecules. This Michaelis complex was obtained from classical metadynamics simulations (see **Supplementary Methods**) based upon the determined *apo* x-ray structure (determined in this work: pdb 5i3d, silver) shown in (a). The inlay shows an expanded view of the active site.



**Figure 3. Free Energy Landscape and Atomic rearrangement along the  $S_Ni$  reaction pathway. (a) Free energy landscape (FEL) reconstructed from the metadynamics simulation**

of the transglycosylation reaction (projection on two collective variables  $CV_1$  and  $CV_2$ ). Contour lines are at 5 kcal/mol. The second transition state (labelled as **3** on the reaction pathway) is above in energy with respect to the first one (labelled as **1**) by 3 kcal/mol. **(b)** Evolution of the most relevant distances involving the donor and the acceptor along the reaction coordinate (see atom numbering in **Figure 2**). Each distance corresponds to an average from all configurations falling into a small region around the corresponding point of the FEL. Data also given in **Supplementary Table S5**. **(c)** Hydrogen atoms have been omitted for clarity, except the one being transferred from the sugar acceptor to the *p*NP leaving group of the donor molecule and the hydroxyl hydrogen atoms of the Gal donor that interact with E206. Bonds being broken/formed are represented by a transparent bond (configurations 1 and 3), whereas dotted lines indicate hydrogen bonding interactions. **(d)** Conformational itinerary of the glucosyl ring along the reaction coordinate. All ring conformations of the metadynamics simulation are mapped onto a projection of the Cremer-Pople sphere from the North pole. Cyan points represent the ring conformations visited by the glucose glycon. The red dots indicate the ring conformation at the reactants, products and intermediate of the reaction (topology **2**).



**Scheme 2.** Proposed frontal face nucleophilic substitution mechanism of SsβG-E387Y.

## References

1. Sinnott, M.L. & Jencks, W.P. Solvolysis of D-Glucopyranosyl Derivatives in Mixtures of Ethanol and 2,2,2-Trifluoroethanol. *Journal of the American Chemical Society* **102**, 2026-2032 (1980).
2. Persson, K. et al. Crystal structure of the retaining galactosyltransferase LgtC from *Neisseria meningitidis* in complex with donor and acceptor analogs. *Nature Structural Biology* **8**, 166-175 (2001).
3. Gibson, R.P., Turkenburg, J.P., Charnock, S.J., Lloyd, R. & Davies, G.J. Insights into Trehalose Synthesis Provided by the Structure of the Retaining Glucosyltransferase OtsA. *Chemistry & Biology* **9**, 1337-1346 (2002).
4. Lairson, L.L., Henrissat, B., Davies, G.J. & Withers, S.G. Glycosyltransferases: Structures, Functions and Mechanisms. *Annual Review of Biochemistry* **77**, 521-555 (2008).
5. Koshland, D.E. Stereochemistry and the Mechanism of Enzymatic Reactions. *Biological Reviews* **28**, 416-436 (1953).
6. Lewis, E.S. & Boozer, C.E. The Kinetics and Stereochemistry of the Decomposition of Secondary Alkyl Chlorosulfites<sup>1</sup>. *Journal of the American Chemical Society* **74**, 308-311 (1952).
7. Chan, J., Tang, A. & Bennet, A.J. A Stepwise Solvent-Promoted S<sub>N</sub>i Reaction of α-D-Glucopyranosyl Fluoride: Mechanistic Implications for Retaining Glycosyltransferases. *Journal of the American Chemical Society* **134**, 1212-1220 (2012).
8. Vetting, M.W., Frantom, P.A. & Blanchard, J.S. Structural and Enzymatic Analysis of MshA from *Corynebacterium glutamicum*. *Journal of Biological Chemistry* **283**, 15834-15844 (2008).
9. Batt, S.M. et al. Acceptor substrate discrimination in phosphatidyl-myo-inositol mannoside synthesis: structural and mutational analysis of mannosyltransferase *Corynebacterium glutamicum* PimB'. *Journal of Biological Chemistry* **285**, 37741-37752 (2010).
10. Chaikuad, A. et al. Conformational plasticity of glycogenin and its maltosaccharide substrate during glycogen biogenesis. *Proceedings of National Academy of Sciences USA* **108**, 21028-21033 (2011).
11. Yu, H. et al. Notch-modifying xylosyltransferase structures support an S<sub>N</sub>i-like retaining mechanism. *Nature Chemical Biology* **11**, 847-854 (2015).
12. Errey, J.C. et al. Mechanistic Insight into Enzymatic Glycosyl Transfer with Retention of Configuration through Analysis of Glycomimetic Inhibitors. *Angewandte Chemie International Edition* **49**, 1234-1237 (2010).
13. Lee, S.S. et al. Mechanistic evidence for a front-side, S<sub>N</sub>i-type reaction in a retaining glycosyltransferase. *Nature Chemical Biology* **7**, 631-638 (2011).
14. Ardevol, A. & Rovira, C. The Molecular Mechanism of Enzymatic Glycosyl Transfer with Retention of Configuration: Evidence for a Short-Lived Oxocarbenium-Like Species. *Angewandte Chemie International Edition* **50**, 10897-10901 (2011).
15. Gomez, H., Polyak, I., Thiel, W., Lluch, J.M. & Masgrau, L. Retaining Glycosyltransferase Mechanism Studied by QM/MM Methods: Lipopolysaccharyl-α-1,4-galactosyltransferase C Transfers α-Galactose via an Oxocarbenium Ion-like Transition State. *Journal of the American Chemical Society* **134**, 4743-4752 (2012).
16. Bobovska, A., Tvaroska, I. & Kona, J. A theoretical study on the catalytic mechanism of the retaining [small alpha]-1,2-mannosyltransferase Kre2p/Mnt1p: the impact of

- different metal ions on catalysis. *Organic & Biomolecular Chemistry* **12**, 4201-4210 (2014).
17. Lira-Navarrete, E. et al. Substrate-Guided Front-Face Reaction Revealed by Combined Structural Snapshots and Metadynamics for the Polypeptide N-Acetylgalactosaminyltransferase 2. *Angewandte Chemie International Edition* **53**, 8206-8210 (2014).
  18. Gomez, H. et al. A computational and experimental study of O-glycosylation. Catalysis by human UDP-GalNAc polypeptide:GalNAc transferase-T2. *Organic & Biomolecular Chemistry* **12**, 2645-2655 (2014).
  19. Albesa-Jove, D. et al. A Native Ternary Complex Trapped in a Crystal Reveals the Catalytic Mechanism of a Retaining Glycosyltransferase. *Angewandte Chemie International Edition* **54**, 9898-9902 (2015).
  20. Corbett, K., Fordham-Skelton, A.P., Gatehouse, J.A. & Davis, B.G. Tailoring the substrate specificity of the  $\beta$ -glycosidase from the thermophilic archeon *Sulfolobus solfataricus*. *FEBS Letters* **509**, 355-360 (2001).
  21. Hancock, S.M., Corbett, K., Fordham-Skelton, A.P., Gatehouse, J.A. & Davis, B.G. Developing promiscuous glycosidases for glycoside synthesis: residues W433 and E432 in *Sulfolobus solfataricus*  $\beta$ -glycosidase are important glucoside and galactoside specificity determinants. *ChemBioChem* **6**, 866-875 (2004).
  22. Trincone, A., Improta, R. & Gambacorta, G. Enzymatic synthesis of polyol and masked polyol glucosides using  $\beta$ -glycosidase of *Sulfolobus solfataricus*. *Biocatalysis and Biotransformations* **12**, 77-88 (1995).
  23. Trincone, A. et al. Enzyme catalysed synthesis of alkyl  $\beta$ -D-glycosides with crude homogenate of *Sulfolobus solfataricus*. *Biotechnology Letters* **13**, 235-240 (1991).
  24. Aguilar, C.F. et al. Crystal structure of the  $\beta$ -glycosidase from the hyperthermophilic archeon *Sulfolobus solfataricus*: resilience as a key factor in thermostability. *Journal of Molecular Biology* **271**, 789-802 (1997).
  25. Trincone, A., Perugino, G., Rossi, M. & Moracci, M. A novel thermophilic glycosynthase that effects branching glycosylation. *Bioorganic and Medicinal Chemistry Letters* **10**, 365-368 (2000).
  26. Mackenzie, L.F., Wang, Q., Warren, R.A.J. & Withers, S.G. Glycosynthases: mutant glycosidases for oligosaccharide synthesis. *Journal of the American Chemical Society* **120**, 5583-5584 (1998).
  27. Moracci, M., Trincone, A. & Rossi, M. Glycosynthases: new enzymes for oligosaccharide synthesis. *Journal of Molecular Catalysis B: Enzymatic* **11**, 155-163 (2001).
  28. Jakeman, D.L. & Withers, S.G. Glycosynthases: new tools for oligosaccharide synthesis. *Trends in Glycosciences and Glycotechnology* **14**, 13-25 (2002).
  29. Perrin, D.D., Dempsey, B. & Serjeant, E.P. *pKa Prediction for Organic Acids and Bases*. (Chapman and Hall, London; 1981).
  30. Zhang, Z. et al. Programmable One-Pot Oligosaccharide Synthesis. *Journal of the American Chemical Society* **121**, 734-753 (1999).
  31. Williams, S.J., Mark, B.L., Vocadlo, D.J., James, M.N.G. & Withers, S.G. Aspartate 313 in the *Streptomyces plicatus* Hexosaminidase Plays a Critical Role in Substrate-assisted Catalysis by Orienting the 2-Acetamido Group and Stabilizing the Transition State. *Journal of Biological Chemistry* **277**, 40055-40065 (2002).

32. Withers, S.G., Street, I.P., Bird, P. & Dolphin, D.H. 2-deoxy-2-fluoroglucosidase: a novel class of mechanism based glucosidase inhibitors. *Journal of the American Chemical Society* **109**, 7530-7531 (1987).
33. Lopez, R. & Fernandez-Mayoralas, A. Enzymatic galactosidation of modified monosaccharides: study of the enzyme selectivity for the acceptor and its application to the synthesis of disaccharides. *Journal of Organic Chemistry* **59**, 737-745 (1994).
34. Yamamoto, K. & Davis, B.G. Creation of an  $\alpha$ -Mannosynthase from a Broad Glycosidase Scaffold. *Angewandte Chemie International Edition* **51**, 7449-7453 (2012).
35. Kim, Y.-W., Lee, S.S., Warren, R.A.J. & Withers, S.G. Directed evolution of a glycosynthase from *Agrobacterium* sp. increases its catalytic activity dramatically and expands its substrate repertoire. *Journal of Biological Chemistry* **279**, 42787-42793 (2004).
36. Mayer, C. et al. Directed evolution of new glycosynthases from *Agrobacterium* beta-glucosidase: a general screen to detect enzymes for oligosaccharide synthesis. *Chemistry & Biology* **8**, 437-443 (2001).
37. Kim, Y.W., Chen, H. & Withers, S.G. Enzymatic transglycosylation of xylose using a glycosynthase. *Carbohydrate Research* **340**, 2735-2741 (2005).
38. Petzelbauer, I., Reiter, A., Splechtna, B., Kosma, P. & Nidetzky, B. Transgalactosylation by thermostable beta-glycosidases from *Pyrococcus furiosus* and *Sulfolobus solfataricus*. *European Journal of Biochemistry* **267**, 5055-5066 (2000).
39. Reuter, S., Rusborg Nygaard, A. & Zimmermann, W. beta-Galactooligosaccharide synthesis with beta-galactosidases from *Sulfolobus solfataricus*, *Aspergillus oryzae*, and *Escherichia coli*. *Enzyme and Microbial Technology* **25**, 509-516 (1999).
40. Crout, D.H.G. & Vic, G. Glycosidases and glycosyl transferases in glycoside and oligosaccharide synthesis. *Current Opinion in Chemical Biology* **2**, 98-111 (1998).
41. Shim, J.-H., Chen, H., Rich, J.R., Goddard-Borger, E.D. & Withers, S.G. Directed evolution of a  $\beta$ -glucosidase from *Agrobacterium* sp. enhances its glycosynthase catalytic activity toward C3-OH modified donor sugar. *PEDS* **25**, 465-472 (2012).
42. Watts, A.G. et al. *Trypanosoma cruzi* trans-sialidase operates through a covalent sialyl-enzyme intermediate: tyrosine is the catalytic nucleophile. *Journal of the American Chemical Society* **125**, 7532-7533 (2003).
43. Lawson, S.L., Warren, R.A.J. & Withers, S.G. Mechanistic consequences of replacing the active site nucleophile Glu-358 in *Agrobacterium* sp. beta-glucosidase with a cysteine residue. *Biochemical Journal* **330**, 203-209 (1998).
44. Gloster, T.M. et al. Structural studies of the beta-glycosidase from *Sulfolobus solfataricus* in complex with covalently and non-covalently bound inhibitors. *Biochemistry* **43**, 6101-6109 (2004).
45. Laio, A. & Parrinello, M. Escaping free-energy minima. *Proceedings of the National Academy of Sciences* **99**, 12562-12566 (2002).
46. Asensio, J.L., Ardá, A., Cañada, F.J. & Jiménez-Barbero, J. Carbohydrate-Aromatic Interactions. *Accounts of Chemical Research* **46**, 946-954 (2013).
47. Biarnés, X., Nieto, J., Planas, A. & Rovira, C. Substrate Distortion in the Michaelis Complex of *Bacillus* 1,3-1,4- $\beta$ -Glucanase: Insight From First Principles Molecular Dynamics Simulations. *Journal of Biological Chemistry* **281**, 1432-1441 (2006).
48. Ardèvol, A. & Rovira, C. Reaction Mechanisms in Carbohydrate-Active Enzymes: Glycoside Hydrolases and Glycosyltransferases. Insights from ab Initio Quantum Mechanics/Molecular Mechanics Dynamic Simulations. *Journal of the American Chemical Society* **137**, 7528-7547 (2015).

49. Davies, G.J., Planas, A. & Rovira, C. Conformational Analyses of the Reaction Coordinate of Glycosidases. *Accounts of Chemical Research* **45**, 308-316 (2012).
50. Speciale, G., Thompson, A.J., Davies, G.J. & Williams, S.J. Dissecting conformational contributions to glycosidase catalysis and inhibition. *Current Opinion in Structural Biology* **28**, 1-13 (2014).
51. Bottoni, A., Miscione, G.P. & DeVivo, M. A theoretical DFT investigation of the lysozyme mechanism: computational evidence for a covalent intermediate pathway. *Proteins* **59**, 118-130 (2005).
52. Vocadlo, D.J. & Davies, G.J. Mechanistic insights into glycosidase chemistry. *Current Opinion in Chemical Biology* **12**, 539-555 (2008).
53. Biarnés, X., Ardèvol, A., Iglesias-Fernández, J., Planas, A. & Rovira, C. Catalytic Itinerary in 1,3-1,4- $\beta$ -Glucanase Unraveled by QM/MM Metadynamics. Charge Is Not Yet Fully Developed at the Oxocarbenium Ion-like Transition State. *Journal of the American Chemical Society* **133**, 20301-20309 (2011).
54. Berrin, J.-G. et al. Substrate (aglycone) specificity of human cytosolic  $\alpha$ -glucosidase. *Biochemical Journal* **373**, 41-48 (2003).
55. Carter, P. & Wells, J.A. Dissecting the catalytic triad of a serine protease. *Nature* **332**, 564-568 (1988).
56. Janda, K.D. et al. Chemical selection for catalysis in combinatorial antibody libraries. *Science* **275**, 945-948 (1997).
57. Schreiber, S.L. Rethinking relationships between natural products. *Nature Chemical Biology* **3**, 352-352 (2007).
58. We would like to thank a Reviewer for this interesting suggestion.
59. An, J., Denton, R.M., Lambert, T.H. & Nacsa, E.D. The development of catalytic nucleophilic substitution reactions: challenges, progress and future directions. *Organic & Biomolecular Chemistry* **12**, 2993-3003 (2014).