

Wen-Yih Jeng,^{a,b,c} Nai-Chen Wang,^{a,b} Cheng-Tse Lin,^{a,b} Wei-Jung Chang,^{a,b} Chia-I Liu^{a,b} and Andrew H.-J. Wang^{a,b*}

^aInstitute of Biological Chemistry, Academia Sinica, Taipei 115, Taiwan, ^bCore Facility for Protein Production and X-ray Structural Analysis, Academia Sinica, Taipei 115, Taiwan, and ^cUniversity Center for Bioscience and Biotechnology, National Cheng Kung University, Tainan 701, Taiwan

Correspondence e-mail:
ahjwang@gate.sinica.edu.tw

High-resolution structures of *Neotermes koshunensis* β -glucosidase mutants provide insights into the catalytic mechanism and the synthesis of glucoconjugates

NkBgl, a β -glucosidase from *Neotermes koshunensis*, is a β -retaining glycosyl hydrolase family 1 enzyme that cleaves β -glucosidic linkages in disaccharide or glucose-substituted molecules. β -Glucosidases have been widely used in several applications. For example, mutagenesis of the attacking nucleophile in β -glucosidase has been conducted to convert it into a glycosynthase for the synthesis of oligosaccharides. Here, several high-resolution structures of wild-type or mutated *NkBgl* in complex with different ligand molecules are reported. In the wild-type *NkBgl* structures it was found that glucose-like glucosidase inhibitors bind to the glycone-binding pocket, allowing the buffer molecule HEPES to remain in the aglycone-binding pocket. In the crystal structures of *NkBgl* E193A, E193S and E193D mutants Glu193 not only acts as the catalytic acid/base but also plays an important role in controlling substrate entry and product release. Furthermore, in crystal structures of the *NkBgl* E193D mutant it was found that new glucoconjugates were generated by the conjugation of glucose (hydrolyzed product) and HEPES/EPPS/opipramol (buffer components). Based on the wild-type and E193D-mutant structures of *NkBgl*, the glucosidic bond of cellobiose or salicin was hydrolyzed and a new bond was subsequently formed between glucose and HEPES/EPPS/opipramol to generate new glucopyranosidic products through the transglycosylation reaction in the *NkBgl* E193D mutant. This finding highlights an innovative way to further improve β -glucosidases for the enzymatic synthesis of oligosaccharides.

Received 18 January 2012
Accepted 26 March 2012

PDB References: *NkBgl* complexes, 3vif; 3vig; 3vih; 3vii; 3vij; 3vik; 3vil; 3vim; 3vin; 3vio; 3vip.

1. Introduction

β -Glucosidases, a heterogeneous group of exo-type glycosyl hydrolases, have been widely used in numerous applications. β -Glucosidases normally catalyze the hydrolysis of glucosidic linkages in disaccharide or glucose-substituted molecules, but transglycosylation under conditions that favour a back reaction can be applied to promote the synthesis of oligosaccharides (Wang & Huang, 2009). Oligosaccharides play numerous roles in biological systems such as molecular recognition in intercellular communication and signal transduction and have shown considerable potential as therapeutics (Dwek, 1996; Varki, 1993; Zopf & Roth, 1996). Converting β -glucosidases into glycosynthases or glycoligases by protein engineering is an attractive approach to the enzymatic synthesis of oligosaccharides owing to the advantages of the controllable stereospecificity and regiospecificity of the reaction (Crout & Vic, 1998; Hanson *et al.*, 2004). The first engineered glycosynthase was constructed from a mutated *Agrobacterium* sp. β -glucosidase by converting the catalytic carboxylate nucleophile residue into a non-nucleophilic residue (Mackenzie *et al.*,

Table 1Data-collection and refinement statistics for *NkBgl* structures.

Values in parentheses are for the highest resolution shell.

Crystal	WT (gluconolactone + HEPES)	WT (1-deoxynojirimycin + HEPES)	WT (glycerol)	WT (bis-Tris)	E193A (glucose)	E193S (cellobiose)
PDB code	3vif	3vig	3vih	3vii	3vij	3vik
Data collection						
Radiation source	PF BL5A	PF BL5A	PF BL5A	NSRRC BL13B1	SPRING-8 BL44XU	SPRING-8 BL12B2
Wavelength (Å)	0.80000	0.80000	0.80000	0.80000	0.90000	0.90000
Space group	C2	C2	C2	C2	C2	C2
Unit-cell parameters						
<i>a</i> (Å)	92.95	93.11	93.16	93.02	92.95	92.97
<i>b</i> (Å)	68.48	68.56	68.53	68.47	68.40	68.71
<i>c</i> (Å)	74.85	75.14	75.47	74.77	75.05	75.83
β (°)	95.33	95.54	95.58	95.32	95.48	95.73
Resolution (Å)	30–1.00 (1.04–1.00)	30–0.99 (1.03–0.99)	30–1.03 (1.07–1.03)	30–0.97 (1.00–0.97)	30–1.03 (1.07–1.03)	30–1.10 (1.14–1.10)
No. of reflections	246183 (24089)	259744 (25857)	224739 (23108)	271148 (27240)	229380 (22898)	190640 (19015)
Completeness (%)	98.0 (96.1)	99.7 (99.5)	96.8 (99.8)	98.8 (99.6)	99.6 (99.9)	99.9 (100)
Multiplicity	5.5 (5.5)	5.4 (5.1)	5.1 (4.9)	4.7 (4.5)	4.9 (4.5)	4.9 (4.8)
R_{merge} (%)	5.7 (88.8)	6.3 (90.8)	8.0 (93.1)	5.2 (59.8)	5.1 (42.6)	6.4 (50.8)
$\langle I/\sigma(I) \rangle$	30.7 (1.9)	31.7 (2.0)	23.3 (2.3)	29.6 (2.1)	30.8 (3.4)	23.4 (2.4)
Overall Wilson <i>B</i> factor (Å ²)	10.1	10.4	11.3	10.0	9.9	13.8
Refinement						
Resolution (Å)	28.20–1.00 (1.05–1.00)	28.30–0.99 (1.04–0.99)	28.00–1.03 (1.09–1.03)	27.20–0.97 (1.02–0.97)	26.40–1.03 (1.08–1.03)	22.95–1.10 (1.16–1.10)
Reflections (work)	233672 (33139)	246458 (35710)	213213 (31817)	257311 (37629)	217731 (31476)	180910 (25878)
Reflections (free)	12371 (1754)	13058 (1837)	11242 (1680)	13672 (1975)	11481 (1725)	9575 (1394)
R_{work} (%)	12.2 (22.9)	13.0 (25.2)	14.8 (21.1)	12.6 (21.7)	12.0 (18.8)	14.7 (24.1)
R_{free} (%)	14.2 (24.5)	14.6 (25.4)	16.8 (22.5)	14.5 (22.4)	13.5 (20.3)	16.8 (24.6)
Geometry deviations						
Bond lengths (Å)	0.007	0.007	0.007	0.007	0.006	0.007
Bond angles (°)	1.4	1.4	1.4	1.4	1.4	1.4
Mean <i>B</i> values (Å ²)/No. of atoms						
Protein atoms	9.7/3911	10.1/3892	11.1/3872	10.0/3869	9.1/3892	12.8/3914
Water molecules	26.5/708	26.6/679	27.1/753	26.9/724	25.3/686	28.9/624
Other atoms	16.3/36	15.2/34	15.6/15	16.8/21	11.7/21	16.7/29
Ramachandran plot† (%)						
Favoured	97.5	97.4	97.7	97.4	97.4	97.7
Allowed	2.5	2.6	2.3	2.6	2.6	2.3

1998). In the past decade, glycosynthases or glycoligases have been engineered to perform glycoside synthesis from activated glycosyl fluoride donors of the opposite anomer and acceptors via the transglycosylation reaction (Perugino *et al.*, 2004; Shaikh & Withers, 2008; Kim *et al.*, 2010).

NkBgl, a β -retaining glycoside hydrolase family 1 β -glucosidase (EC 3.2.1.21), was isolated and identified from the salivary glands of the termite *Neotermes koshunensis* by Tokuda *et al.* (2002). *NkBgl* is composed of 498 amino acids containing a 20-residue signal peptide at the N-terminus which is removed during the secretion process. In our previous work, functional recombinant *NkBgl* was expressed in *Escherichia coli* and crystal structures of its complexes with Tris (inhibitor-like) and *pNPG* (substrate) were reported (Jeng *et al.*, 2011).

In the present study, several amino acids, including the two catalytic glutamates, of *NkBgl* were mutated in order to elucidate the detailed mechanism of hydrolysis and transglycosylation by *NkBgl*. Several structures of *NkBgl* mutants mutated at the catalytic acid/base Glu193 in complex with different ligand molecules were determined at high resolution (~ 1.0 Å). Based on observations of the crystal structures of the *NkBgl* E193A, E193S and E193D mutants, we speculate that Glu193 not only acts as the catalytic acid/base but also

plays an important role in controlling substrate entry and product release. We found that new glucoconjugated products were lodged in the active-site pocket in the crystal structures of the *NkBgl* E193D mutant, implying that this mutation enhances transglycosylation activity. The *NkBgl* E193D mutant appears to be suitable for development as a useful tool for the enzymatic synthesis of numerous therapeutic oligosaccharides via the transglycosylation reaction.

2. Materials and methods

2.1. Protein expression and purification

The cloning, purification and crystallization of *NkBgl* have previously been reported (Jeng *et al.*, 2011). Briefly, the gene for *NkBgl* was cloned into pET21a vector and expressed in *E. coli* BL21 (DE3) cells. All mutants were generated using the PCR-based QuikChange site-directed mutagenesis kit (Stratagene). For easier purification, a His₆ extension was added to the C-terminus of *NkBgl*. The protein was purified using Ni²⁺-nitrilotriacetic acid affinity chromatography. After confirmation by SDS-PAGE, the purified protein was concentrated and placed in storage buffer (50 mM HEPES,

Table 1 (continued)

Crystal	E193S (salicin)	E193D (HEPES-1-glucoside, cellobiose)	E193D (HEPES-1-glucoside, salicin)	E193D (EPPS-1-glucoside, salicin)	E193D (opipramol-1-glucoside, salicin)
PDB code	3vil	3vim	3vin	3vio	3vip
Data collection					
Radiation source	NSRRC BL13C1	NSRRC BL13B1	SPRING-8 BL12B2	SPRING-8 BL44XU	NSRRC BL13C1
Wavelength (Å)	0.97622	0.80000	0.90000	0.90000	0.97622
Space group	C2	C2	C2	C2	C2
Unit-cell parameters					
<i>a</i> (Å)	92.91	93.12	93.15	92.95	92.79
<i>b</i> (Å)	68.54	68.56	68.57	68.51	68.50
<i>c</i> (Å)	75.66	75.46	75.64	75.33	75.08
β (°)	95.70	95.72	95.61	95.55	95.38
Resolution (Å)	30–1.15 (1.19–1.15)	30–1.03 (1.07–1.03)	30–1.13 (1.17–1.13)	30–1.12 (1.16–1.12)	30–1.28 (1.33–1.28)
No. of reflections	166145 (16597)	223573 (21839)	173272 (17010)	179748 (17958)	119961 (11961)
Completeness (%)	99.4 (99.7)	96.4 (94.5)	98.0 (96.9)	99.3 (99.5)	99.7 (99.8)
Multiplicity	4.7 (4.5)	4.8 (4.6)	4.5 (4.2)	4.6 (4.4)	3.8 (3.8)
<i>R</i> _{merge} (%)	4.6 (45.8)	4.5 (53.6)	4.5 (52.8)	6.6 (53.1)	3.6 (19.7)
<i>I</i> / σ (<i>I</i>)	34.5 (3.7)	31.0 (2.5)	28.7 (2.4)	22.9 (3.1)	39.3 (8.3)
Overall Wilson <i>B</i> factor (Å ²)	13.8	10.0	13.2	11.7	15.3
Refinement					
Resolution (Å)	25.20–1.15 (1.21–1.15)	19.20–1.03 (1.09–1.03)	22.50–1.13 (1.19–1.13)	28.20–1.12 (1.18–1.12)	22.90–1.28 (1.35–1.28)
Reflections (work)	157615 (22871)	212221 (30025)	164457 (23417)	170533 (24069)	113838 (16313)
Reflections (free)	8352 (1242)	11218 (1605)	8684 (1160)	8978 (1270)	6025 (858)
<i>R</i> _{work} (%)	12.8 (21.6)	12.8 (17.0)	13.5 (23.2)	12.9 (18.5)	12.0 (13.6)
<i>R</i> _{free} (%)	15.4 (23.5)	14.3 (18.9)	15.6 (25.6)	15.1 (21.0)	14.4 (17.4)
Geometry deviations					
Bond lengths (Å)	0.007	0.006	0.007	0.007	0.007
Bond angles (°)	1.4	1.4	1.4	1.4	1.4
Mean <i>B</i> values (Å ²)/No. of atoms					
Protein atoms	12.8/3884	9.4/3915	11.5/3891	10.7/3926	10.6/3899
Water molecules	31.0/669	25.1/651	28.9/677	27.3/627	27.7/640
Other atoms	24.5/36	13.0/34	20.2/35	20.9/35	24.6/62
Ramachandran plot† (%)					
Favoured	97.4	97.9	97.7	97.7	97.9
Allowed	2.6	2.1	2.3	2.3	2.1

† Categories were defined by *RAMPAGE*.

100 mM NaCl pH 8.0) using 30 kDa cutoff membranes with Jumbosep (Pall) and Amicon Ultra-15 (Millipore) centrifugal devices.

2.2. Crystallization

The protein was prepared at 25 mg ml^{−1} in 50 mM HEPES buffer pH 8.0 containing 100 mM NaCl and 5 mM DTT for crystallization. Crystals were grown from a drop composed of 1.5 µl protein solution and 1.5 µl reservoir solution consisting of 18–21% (w/v) PEG 3350 and 0.1–0.25 M MgCl₂ in 0.1 M bis-Tris (Fluka) buffer pH 6.5 equilibrated at 298 K against 250 µl reservoir solution using the sitting-drop vapour-diffusion method. Crystals of the *NkBgl*–gluconolactone and *NkBgl*–1-deoxynojirimycin complexes were obtained in cocrystallization experiments with *NkBgl* protein pre-mixed with 5 mM gluconolactone and 1 mM 1-deoxynojirimycin, respectively. Crystals of *NkBgl* E193S in complex with salicin or cellobiose were obtained in cocrystallization experiments with *NkBgl* E193S protein pre-mixed with 10 mM salicin or cellobiose. Crystals of *NkBgl* E193D in complex with HEPES-1-glucoside were obtained in cocrystallization experiments with *NkBgl* E193D protein pre-mixed with 10 mM salicin or cellobiose. Crystals of the *NkBgl*–glycerol, *NkBgl*–bis-Tris and *NkBgl* E193A–glucose complexes were obtained under the same conditions as those described above without adding any

other chemicals. Crystals of *NkBgl* E193D in complex with EPPS-1-glucoside and opipramol-1-glucoside were obtained in cocrystallization experiments with the *NkBgl* E193D protein pre-mixed with 10 mM salicin and 4 mM EPPS or opipramol and incubated at 303 K for 16 h before crystallization. To avoid interference of HEPES in the cocrystallization experiments of *NkBgl* E193D in complex with EPPS-1-glucoside and opipramol-1-glucoside, the *NkBgl* E193D protein was prepared at 25 mg ml^{−1} in 20 mM imidazole buffer pH 7.0 containing 100 mM NaCl and 5 mM DTT. Both wild-type *NkBgl* and the *NkBgl* E193S mutant were cocrystallized under the same conditions as *NkBgl* E193D to identify whether these proteins could synthesize the same new glucopyranosidic products.

2.3. Diffraction data collection, structure determination and refinement

All crystals were flash-cooled to 100 K in a stream of cold nitrogen prior to data collection. Prior to flash-cooling, the crystals were soaked briefly in a cryoprotectant solution. The cryoprotectant solution consisted of 23% (w/v) PEG 3350, 0.12 M NaCl, 0.25 M MgCl₂ and either 8% (w/v) glycerol or 10% (w/v) ethylene glycol. X-ray diffraction data were collected on SPXF beamlines BL13B1 and BL13C1 at the National Synchrotron Radiation Research Center (NSRRC),

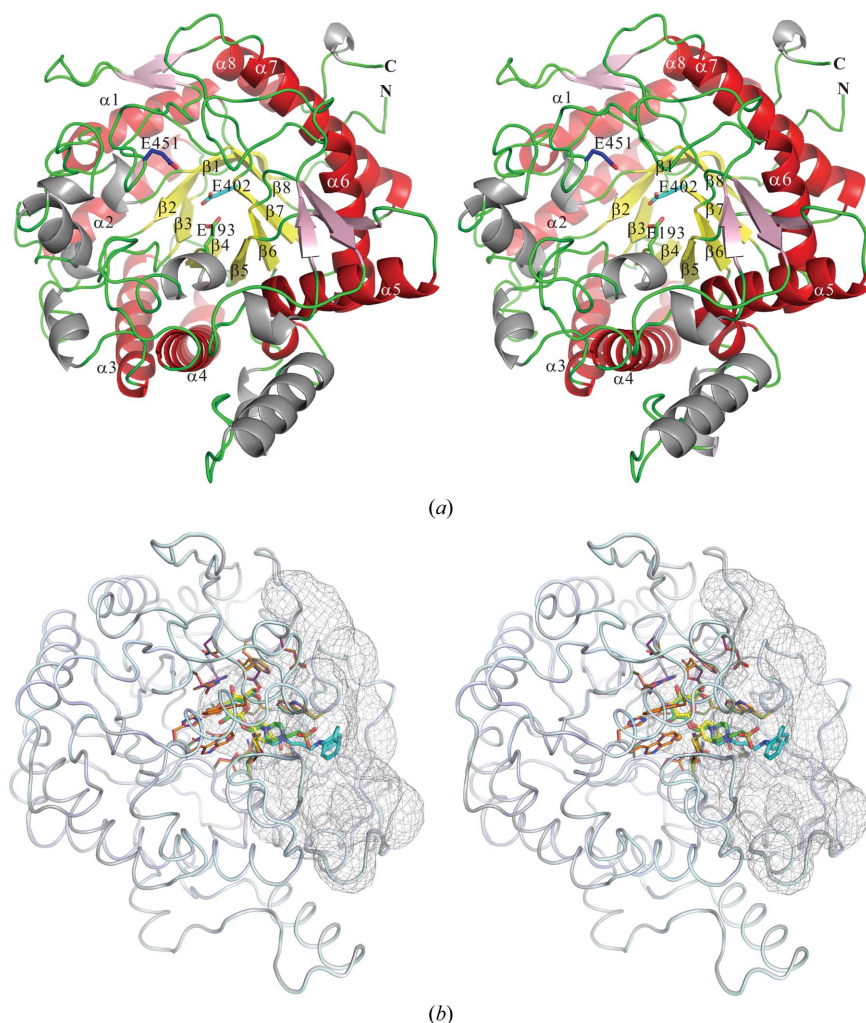


Figure 1

Structural information on *NkBgl*. (a) The overall structure of *NkBgl* adopts a typical $(\alpha/\beta)_8$ TIM-barrel fold. A stereoview of the *NkBgl* structure is shown as a ribbon representation. The helices and strands which constitute the $(\alpha/\beta)_8$ TIM-barrel fold are labelled $\alpha 1$ – $\alpha 8$ (red) and $\beta 1$ – $\beta 8$ (yellow), respectively. The loops, extra helices and extra strands are shown in green, grey and light pink, respectively. The side chains of the three important glutamates in the active site of *NkBgl* are shown as stick models. (b) The bell-shaped pocket of the *NkBgl* active site is illustrated as a grey mesh model. The shape and volume of the active-site pocket of *NkBgl* (PDB entry 3ahz; Jeng *et al.*, 2011) were generated and calculated using cleft analysis in PDBsum (Laskowski *et al.*, 2005).

Taiwan, on beamline BL5A at the Photon Factory (PF), Japan and on beamlines BL12B2 and BL44XU at SPring-8, Japan. The diffraction images were processed using the *HKL-2000* program package (Otwinowski & Minor, 1997). Structures were determined by molecular replacement using the deposited structure of *NkBgl* (PDB entry 3ahz; Jeng *et al.*, 2011) as a search model. All structure determinations were carried out using alternate cycles of model building in *Coot* (Emsley & Cowtan, 2004) and refinement in *REFMAC5* (Winn *et al.*, 2003; Murshudov *et al.*, 2011) with TLS-group tensors and anisotropic *B* factors. The stereochemistry and structure of the final models were analyzed by *RAMPAGE* (Lovell *et al.*, 2003) and *SFCHECK* (Vaguine *et al.*, 1999) from the *CCP4* program suite (Collaborative Computational Project, Number 4, 1994; Winn *et al.*, 2011). Data-collection and refinement

statistics are summarized in Table 1. The atomic coordinates and structure factors were deposited in the Protein Data Bank; the PDB deposition codes are provided in Table 1. Structural figures were produced using *PyMOL* (<http://www.pymol.org>).

2.4. Kinetic assays

Resorufin- β -D-glucopyranoside, a fluorescent substrate, was used for kinetic analysis of wild-type and mutated *NkBgl*. All reactions were performed in 100 mM phosphate–citrate buffer pH 5.5 at 298 K. Reactions were carried out using a final concentration of 0.1–200 μ M resorufin- β -D-glucopyranoside in a total volume of 1 ml. The reaction was initiated by adding 10 μ l wild-type *NkBgl* to give a final concentration of 0.1–4 nM enzyme, whereas the *NkBgl* mutant proteins were at a concentration of 0.01–1 μ M. The release of resorufin was monitored continuously at excitation and emission wavelengths of 535 and 595 nm, respectively, using a Fluorolog-3 spectrofluorometer (Horiba). K_m and k_{cat} values were calculated by fitting the data to the Michaelis–Menten equation using *GraphPad Prism* 5.0.

3. Results

3.1. The overall structure of *NkBgl*

The crystal structures of wild-type or mutated *NkBgl* in complex with various ligands were determined at 0.97–1.28 Å resolution (Table 1). The overall structure of *NkBgl* is similar to the structures of glycosyl hydrolase family 1 enzymes, exhibiting a classical $(\alpha/\beta)_8$ TIM-barrel fold (Fig. 1a). The bell-shaped pocket of the *NkBgl* active site is approximately 20 Å deep and 3000 Å³

in volume and was located on connecting loops at the C-terminal end of the β -sheets of the TIM barrel (Figs. 1 and 2). Glu193 is the catalytic acid/base of *NkBgl* and Glu402 is the nucleophile. The high resolution of the 11 structures presented in this study allowed us to unequivocally visualize the alternative conformations exhibited by the side chains of 38 amino-acid residues of *NkBgl*. The structure of the *NkBgl* E193D–EPPS-1-glucoside complex had 22 residues adopting alternative conformations, the largest number among the 11 structures. Details of the alternative conformations in each structure are summarized in Supplementary Table S1¹. In addition, the Ala208–Pro209, Glu350–Pro351 and Trp444–

¹ Supplementary material has been deposited in the IUCr electronic archive (Reference: MH5058). Services for accessing this material are described at the back of the journal.

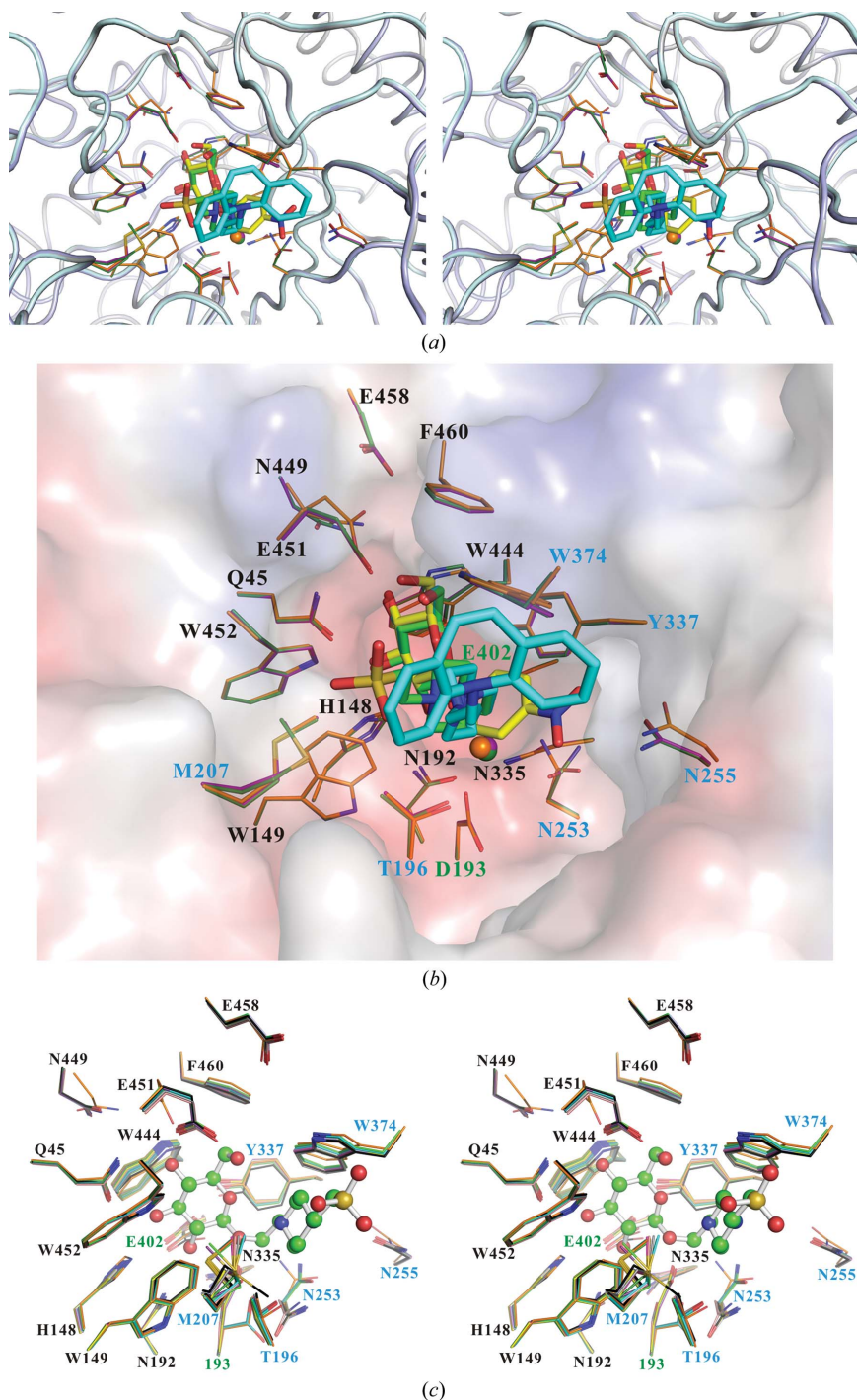


Figure 2

A close-up view of the active-site pocket of *NkBgl*. (a, b) Close-up views of the active-site pockets of *NkBgl* E193D in complex with HEPES-1-glucoside (from cellobiose), opipramol-1-glucoside and *pNPG* are shown as (a) cartoon loop and (b) electrostatic surface representations. The superimposed HEPES-1-glucoside, opipramol-1-glucoside and *pNPG* molecules are shown as stick models. (c) A stereoview of the superimposed surrounding residues of the active sites of 13 *NkBgl* structures. The new glucopyranosidic product, HEPES-1-glucoside, is shown as a ball-and-stick model. The side chains of *NkBgl* are shown as line models. Water molecules are shown as sphere models. Positive charges on the electrostatic surfaces are shown in blue and negative charges in red. C atoms of HEPES-1-glucoside, opipramol-1-glucoside and *pNPG* are shown in green, cyan and yellow, respectively. C atoms of side chains and water molecules in the structures of HEPES-1-glucoside, opipramol-1-glucoside and *pNPG* are shown in green, purple and orange, respectively. Protein O atoms are shown in red, N atoms in blue and S atoms in gold. The *NkBgl* catalytic residues are labelled in green, the glycone-binding site residues in black and the aglycone-binding site residues in cyan.

Ser445 peptide bonds were observed to adopt a *cis* configuration for various reasons.

3.2. Comparison of *NkBgl* structures

The superposition of 13 crystal structures, including wild-type and mutated *NkBgl*s in complex with various ligands in the active site and two structures that have previously been reported (Jeng *et al.*, 2011), revealed the highly conserved architecture of these proteins (Supplementary Fig. S1). To assess the structural differences among these proteins, the coordinates of the PDB files were uploaded onto the *PDBFold* website and the r.m.s.d.s between these models were calculated to be in the range 0.034–0.168 Å over 472 C α atoms. Details of the r.m.s.d. values of these structures are summarized in Supplementary Table S2.

The residues around the active site of the 13 superposed crystal structures of wild-type or mutated *NkBgl* are shown in Fig. 2(c). When the catalytic acid/base of *NkBgl* was mutated from Glu193 to the shorter Asp193, the space for substrate binding was enlarged and resulted in a rotation of the Asn253 side chain towards the substrate-binding pocket (Fig. 2c). In addition, the orientations of the Asn255, Asn449 and Glu451 side chains in the *NkBgl* E193D–*pNPG* structure differed more than usual from the other 12 structures that were solved. This variability may have resulted from the accommodation of the relatively large and rigid *pNPG* in *NkBgl* E193D (Figs. 2b and 2c). The catalytic cavities of the 13 crystal structures maintained the same shape and were of similar size. To understand the conformational variability of the active site among the *NkBgl* structures, the distances between the C α atom of residue 193 and the C α atoms of Tyr337, Trp374, Glu402, Trp444 and Glu451, as well as the C ϵ^3 atom of Trp374, were measured (Supplementary Table S3).

3.3. Glucosidase inhibitors promote the retention of HEPES in the aglycone-binding pocket of *NkBgl*

Gluconolactone (GNL) and 1-deoxy-nojirimycin (DNJ) are glucose-like inhibitors of β -glucosidase. In the structures of wild-type *NkBgl* cocrystallized with GNL or DNJ, not only were the inhibitors found to be bound in each glycone-binding pocket, but an unexpected HEPES molecule was also found to be bound in each aglycone-

binding pocket (Figs. 3*a* and 3*b*). In the *NkBgl*–GNL structure the hydroxyl group of HEPES forms direct hydrogen bonds to the carboxyl group of Glu193 and the O1 atom of GNL (Fig. 3*a*). In the *NkBgl*–DNJ structure the hydroxyl group of HEPES forms direct and water-mediated hydrogen bonds to the carboxyl group of Glu193 and the N5 atom of DNJ, respectively (Fig. 3*b*). Moreover, in both the *NkBgl*–GNL and *NkBgl*–DNJ structures hydrophobic interactions between the piperazine ring of HEPES and the indole ring of Trp374 were observed (Figs. 3*a* and 3*b*). No HEPES was observed in the structures of wild-type *NkBgl* in complex with glycerol or bis-Tris (Supplementary Figs. S2*a* and S2*b*). Based on these observations, we hypothesize that the glucose-like inhibitors that bind to the glycone-binding pocket may contribute to the retention of HEPES in the aglycone-binding pocket of *NkBgl* through additional hydrogen bonds between the inhibitors and HEPES.

3.4. The control of substrate entry and product release by catalytic acid/base mutation of *NkBgl*

The single or double mutations E193D, E193S, E193A, E402Q, E402A, E451Q, E193D/E402Q, E193D/E451Q and E193A/E402A of *NkBgl* were prepared for crystallization. When the Glu402 residue of *NkBgl* was mutated to Gln or Ala, or when Glu451 was mutated to Gln, the mutant proteins failed to bind substrate in the crystals on cocrystallization or crystal soaking. In addition, all of these mutant proteins demonstrated extremely low activity toward the substrates (Table 2). These results suggest that mutation of Glu402 and Glu451 may make *NkBgl* proteins lose their substrate-binding ability. We speculate that both the Glu402 and the Glu451 residues are important for recognition of the glucose moiety of the substrates and also contribute to glucose binding.

The crystal structure of the *NkBgl* E193S mutant in complex with cellobiose shows that the cellobiose molecule is located at an incorrect position that is too distant from the attacking nucleophile of *NkBgl* for hydrolysis to take place (Fig. 3c). Additional hydrogen bonds are formed between the amino group of Asn253 and the O6' atom of the reducing-end glucose moiety of cellobiose and also between the amino group of Asn255 and the O1' atom of the reducing-end glucose (Fig. 3c). Meanwhile, two water

molecules displace the original binding positions of the O4 and O5 atoms of the nonreducing glucose moiety of cellobiose (Fig. 3c).

In the crystal structure of the *NkBgl* E193S mutant in complex with salicin, a nearly intact salicin located at the correct position for nucleophilic attack by *NkBgl* was

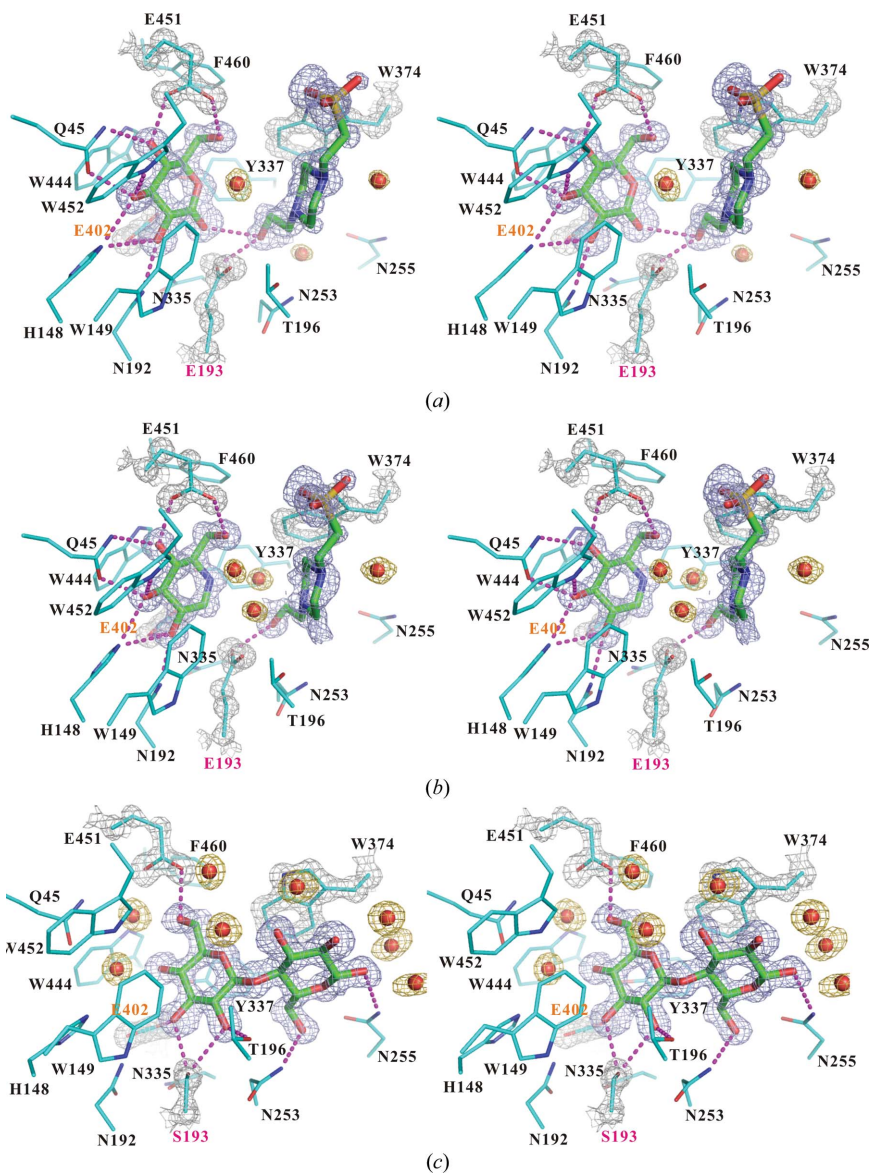


Figure 3

A stereoview of ligand molecules and the surrounding residues in the active sites of *NkBgl* structures. Wild-type *NkBgl* is shown in complex with (a) gluconolactone and HEPES and (b) 1-deoxynojirimycin and HEPES. (c) The *NkBgl* E193S mutant in complex with cellobiose. $|F_o| - |F_c|$ difference Fourier maps for the ligand molecules contoured at 3.0σ are shown in light blue. $2|F_o| - |F_c|$ maps for the side chains of selected residues contoured at 2.0σ are shown in grey, whereas water molecules contoured at 1.5σ are shown in yellow/orange. The side chains of residues around the active site of *NkBgl* are shown as line models. Water molecules are shown as sphere models. The ligand molecules are shown as stick models. The Glu402 residue labelled in orange is the catalytic nucleophile of *NkBgl*. The acid/base residues of wild-type or mutated *NkBgl*s are labelled in magenta. Hydrogen-bond interactions between residues and ligand molecules are shown as magenta dotted lines. C atoms of proteins are shown in cyan and those of ligand molecules in green. O atoms are shown in red, N atoms in blue and S atoms in gold. Next to the carboxyl group of Asp193, a water molecule, which is presumably involved in both the hydrolysis and transglycosylation reactions, is indicated by a magenta asterisk.

observed in the active site (Fig. 4a and Supplementary Fig. S2c). Two elongated extra areas of electron density were observed near the C1 and O1 atoms of salicin (Fig. 4a). The distances from the Ser193 O^γ atom to the salicin C1 and O1 atoms were 4.17 and 3.62 Å, respectively. The distance from the Thr196 O^{γ1} atom to the salicin O1 atom was 4.15 Å. From these distances, it appeared that Ser193 and Thr196 were unable to directly attack the salicin C1 or O1 atoms. However, a 2-hydroxybenzyl alcohol molecule, one of the hydrolyzed products of salicin, was observed in the crystal-packing space surrounded by the side chains of Glu347, Tyr349 and Tyr354 of the *NkBgl* molecule at (*x*, *y*, *z*) as well as Glu288 and Glu289 of the symmetry-related *NkBgl* molecule at (*−x* − 1/2, *y* + 1/2, *−z*) (Fig. 4b). This 2-hydroxybenzyl alcohol may be an impurity from a chemical reagent or a product of the hydro-

lysis of salicin by *NkBgl* E193S with the assistance of small weak acids (Takahashi & Wraight, 1991).

In the crystal structures of the *NkBgl* E193A mutant, a glucose molecule, which is the end product of β-glucosidase, continued to occupy the glycone-binding site of the mutant protein during expression in *E. coli* (Supplementary Fig. S2d). In both cocrystallization and soaking experiments, cellobiose or salicin (substrates of β-glucosidase) at 50 mM could not compete with the glucose molecule and displace it from the glycone-binding site of the *NkBgl* E193A mutant protein. The tight glucose-binding phenomenon in the E193A mutant was not observed in the E193D and E193S mutants of *NkBgl*.

Based on the structures of the *NkBgl* E193S mutant in complex with salicin or cellobiose (Fig. 3c and Supplementary Fig. S2c), the formation of additional hydrogen bonds between

cellobiose and the side chains of Asn253 and Asn255 prevents the cellobiose molecule from moving further into the active site of *NkBgl*. Enzymatic activity assays of wild-type *NkBgl* toward salicin or cellobiose substrates were in concert with the observations in the crystal structures (Supplementary Fig. S3). The kinetic studies showed the following order of affinity of *NkBgl* towards the substrate resorufin-β-D-glucopyranoside: E193D > wild type > E193A ≈ E193S (Table 2). Furthermore, based on our observations on the crystal structures of *NkBgl* E193A, E193S and E193D mutants, we propose that the Glu193 residue not only acts as the catalytic acid/base but also assists in substrate entry and product release from *NkBgl*.

3.5. The *NkBgl* E193D mutant favours the transglycosylation reaction

In the crystal structures of *NkBgl* E193D cocrystallized with substrates such as cellobiose or salicin, a new gluco-pyranosidic product was found in the active site. This product was likely to be formed by coupling the C1 atom of glucose with the hydroxyl group of HEPES *via* a covalent bond (Fig. 3d and Supplementary Fig. S2e). Other similar products were also found in the active site of *NkBgl* when the HEPES was replaced by EPPS or opipramol in the cocrystallization of *NkBgl* E193D and salicin (Figs. 3e and 3f). In these structures, the glucose moieties of the glucoconjugated products were bound to the glycone-binding site, as are the glucose-like inhibitors or glucose in the structures of *NkBgl*-GNL, *NkBgl*-DNJ and *NkBgl* E193A-glucose. In the

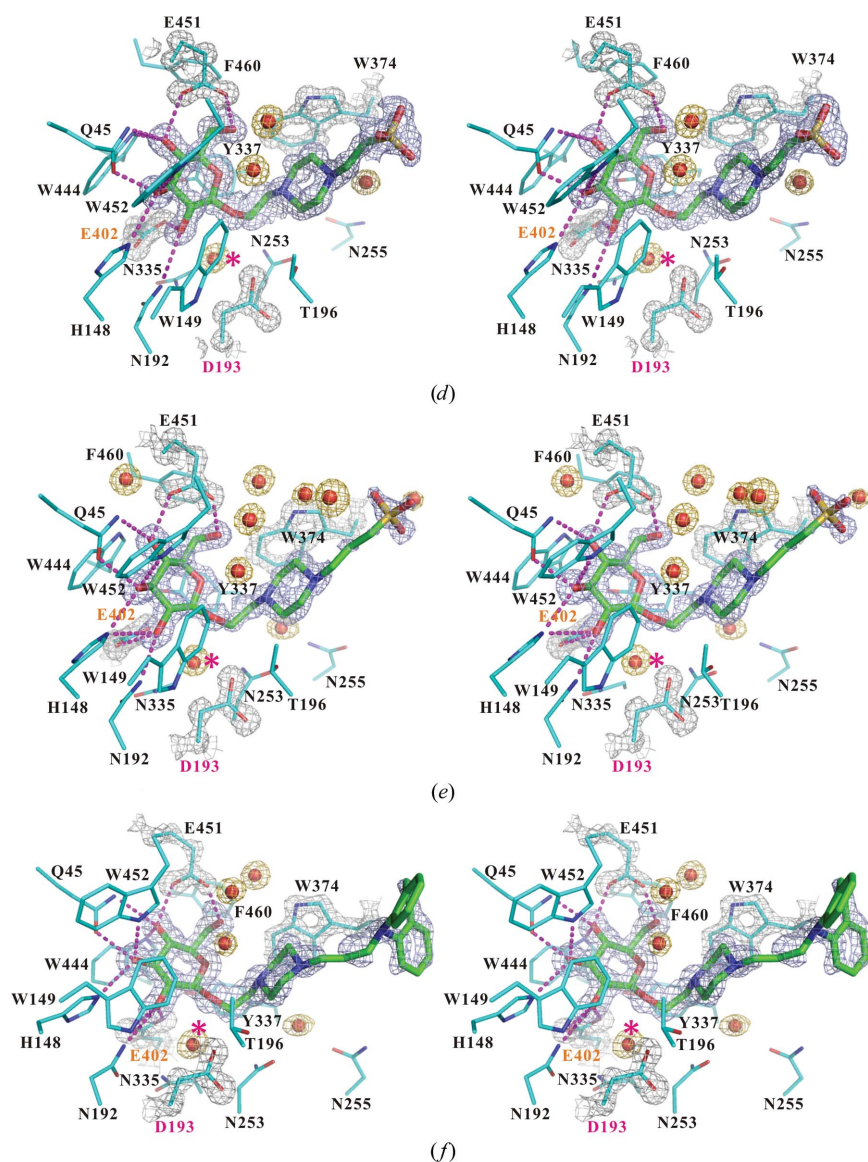


Figure 3 (continued)

(d–f) The *NkBgl* E193D mutant in complex with HEPES-1-glucoside (d), EPPS-1-glucoside (e) or opipramol-1-glucoside (f). Ligand molecules in (d) create a new adduct generated from cellobiose and HEPES, whereas (e) and (f) show new adducts generated from salicin and EPPS and opipramol, respectively.

Table 2

Kinetic parameters of wild-type and mutated *NkBgl*s towards resorufin- β -D-glucopyranoside.

Values are shown with standard deviations calculated from three replicates using a single prepared protein sample.

Protein	k_{cat} (min^{-1})	K_{m} (μM)	$k_{\text{cat}}/K_{\text{m}}$ ($\text{min}^{-1} \text{mM}^{-1}$)
WT	540.54 ± 18.05	11.64 ± 1.45	47170 ± 7028
E193D	7.69 ± 0.14	0.94 ± 0.08	8240 ± 828
E193S	0.355 ± 0.013	29.02 ± 3.25	12.39 ± 1.69
E193A	0.485 ± 0.015	29.68 ± 2.34	16.44 ± 1.61
E402Q	0.078 ± 0.003	21.48 ± 2.68	3.69 ± 0.56
E402A	0.034 ± 0.002	28.49 ± 3.96	1.22 ± 0.21
E451Q	3.53 ± 0.10	28.29 ± 2.35	125.65 ± 12.74

aglycone-binding site, where the HEPES/EPPS/opipramol moieties of the glucoconjugated products were bound, the piperazine rings of these products interacted hydrophobically with the indole ring of Trp374, as does HEPES in the *NkBgl*-GNL and *NkBgl*-DNJ structures (Figs. 3*a*, 3*b*, 3*d*, 3*e*, 3*f* and Supplementary Fig. S2*e*). To identify whether the transglycosylation reaction occurs in the presence of other primary alcohols, HEPES was replaced by 1-butanol, 1-hexanol and 1-octanol in the cocrystallization of the *NkBgl* E193D protein with salicin/cellobiose. However, no new products were

observed in the crystal structures. We conclude that hydrophobic interactions between the indole ring of Trp374 and the piperazine ring of HEPES/EPPS/opipramol appear to be crucial in stabilizing the binding of HEPES/EPPS/opipramol molecules in the aglycone pocket of *NkBgl* to form new compounds.

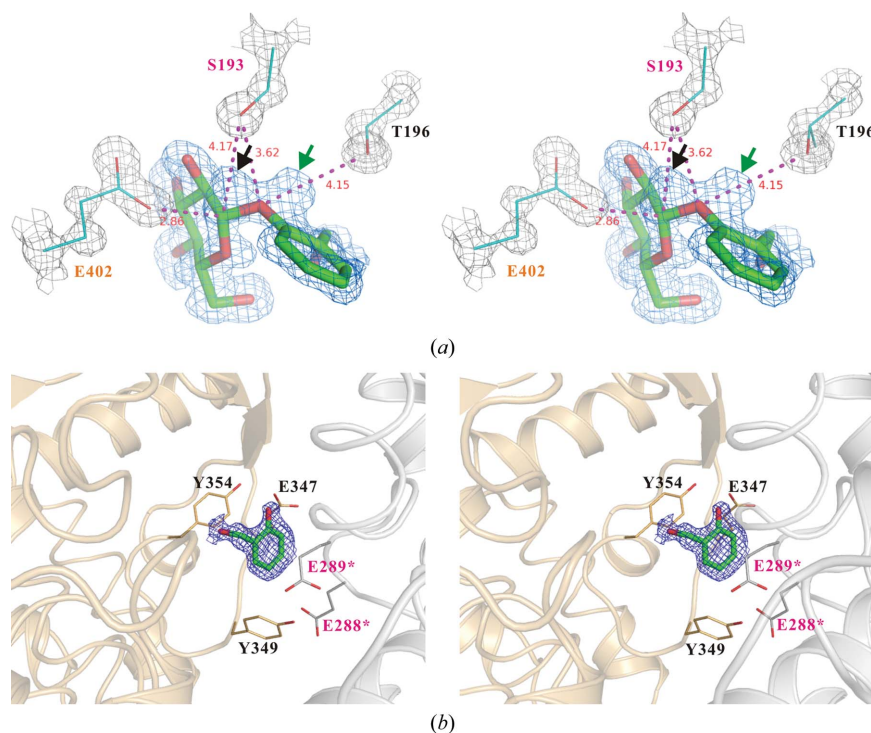
These glucoconjugated adducts produced *via* the transglycosylation reaction catalyzed by the *NkBgl* E193D mutant were not observed in the structures of wild-type *NkBgl* and the *NkBgl* E193S mutant. Hence, the transglycosylation reaction that contributes to the enzymatic synthesis of oligosaccharides is more favourable in the *NkBgl* E193D mutant than in wild-type *NkBgl* and the *NkBgl* E193S mutant. In addition to the structural evidence, a new product, opipramol-1-glucoside, was isolated by HPLC and validated by mass spectrometry (Supplementary Fig. S4).

4. Discussion

In accordance with the structures of *NkBgl* E193D in complex with substrates/products, a retaining β -glucosidase mechanism of wild-type *NkBgl* activity is presented in Fig. 5(*a*). At the active site, Glu402 acts as a nucleophile to attack the anomeric carbon directly, resulting in the formation of a glycosyl-enzyme intermediate and the release of the hydrolyzed product bound at the aglycone-binding site. Subsequently, the deprotonated acidic carboxylate of Glu193 launches a second nucleophilic attack *via* a water molecule towards the anomeric carbon of the resulting glycosyl enzyme to produce glucose as the hydrolysis product.

In *NkBgl* E193D, the mutation of the catalytic acid/base increases the distance between the anomeric carbon of the substrate and the carboxyl group of aspartate relative to that of glutamate. A water molecule located close to the carboxyl group of Asp193 and the anomeric carbon of the substrate/product was found in the structures of *NkBgl* E193D in complex with substrates/products (Figs. 3*d*, 3*e*, 3*f* and Supplementary Fig. S2*e*). In contrast, no water molecules were found at similar positions in other structures. Hence, we speculate that the β -carboxyl group of Asp193 attacks the anomeric carbon of cellobiose/salicin through the nearby water molecule and causes the removal of an *R* group from the substrate in the hydrolysis step (Fig. 5*b*).

It is known that alcohols can enhance the hydrolytic activity of β -glucosidases owing to the contribution of the transglycosylation reaction in reducing feedback inhibition (Umezurike, 1988; Gopalan *et al.*, 1989). The water molecule replaced by an alcohol

**Figure 4**

Electron-density maps around the active site and the crystal-packing space of the *NkBgl* E193S mutant in complex with salicin. (*a*) A nearly intact salicin was observed in the active site after nucleophilic attack. (*b*) A 2-hydroxybenzyl alcohol molecule, a hydrolysis product of salicin, was observed in the crystal-packing space surrounded by the side chains of Glu347, Tyr349 and Tyr354 of the *NkBgl* molecule at (x, y, z) as well as Glu288 and Glu289 of the symmetry-related *NkBgl* molecule at ($-x - 1/2, y + 1/2, -z$). $|F_o| - |F_c|$ difference Fourier maps for salicin and 2-hydroxybenzyl alcohol contoured at 3.0σ are shown in light blue and blue, respectively. $2|F_o| - |F_c|$ maps for the side chains of selective residues contoured at 2.0σ are shown in grey. The side chains of *NkBgl* are shown as line models. Salicin and 2-hydroxybenzyl alcohol are shown as stick models.

during hydrolysis of the glycosyl-enzyme intermediate has been proposed to take part in the mechanism of the alcohol-

stimulated transglycosylation (Watt *et al.*, 1998). We propose that the primary hydroxyl groups of HEPES/EPPS/opipramol

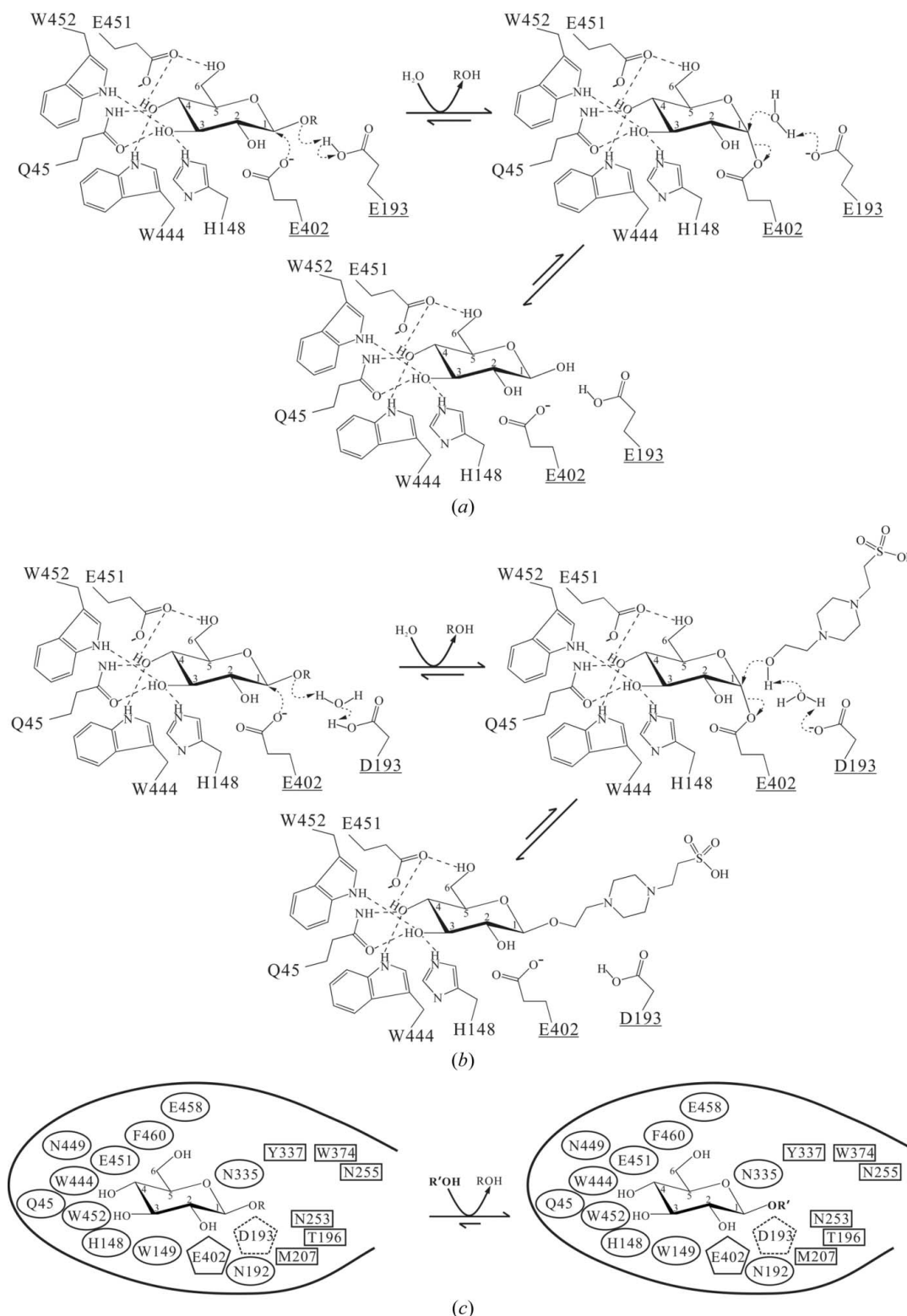


Figure 5

A schematic illustration of the proposed reaction mechanisms. (a) Hydrolysis of the glucosidic bond is catalyzed by wild-type *NkBgI*. (b) *NkBgI* E193D catalyzes the hydrolysis and formation of the glucosidic bond. Only residues linked by direct hydrogen bonds and catalytic residues are shown; the catalytic residues are highlighted. Hydrogen bonds are indicated by dashed lines. (c) *NkBgI* E193D favours the transglycosylation reaction. Residues around the glycone-binding pocket and the aglycone-binding pocket are shown in ovals and in rectangles, respectively.

in *NkBgl* E193D take the place of water as a nucleophilic molecule to perform the water-mediated attack on the anomeric carbon in the hydrolysis of the intermediate, generating new glucopyranosidic products in the transglycosylation step (Fig. 5*b*). In our study, HEPES/EPPS/opipramol, which are considered to be more nucleophilic than the water molecule in the formation of the new adduct, served as primary alcohols.

Because the active site of *NkBgl* has a long aglycone-binding pocket and a bell-shaped exit, it should be easy for it to release the part of the hydrolyzed product that is bound to the aglycone site and it is likely to be beneficial to import compounds as the acceptor of transglycosylation in such structural environments (Figs. 1 and 2). In our study, *NkBgl* E193D was capable of synthesizing soluble glucopyranosidic products using acceptors of various lengths from the short HEPES to the long and large opipramol. We speculate that the bulk-solvent-exposed and bell-shaped exit of the *NkBgl* active site has the advantage of having a large capacity to hold acceptors of various lengths to synthesize large glucopyranosidic products through the transglycosylation reaction (Fig. 1*b*).

Stabilizing the acceptor in the aglycone-binding pocket is important for carrying out the transglycosylation reaction and is helpful in forming new glucoconjugated adducts (Tran *et al.*, 2010). In our study, the indole ring of Trp374 of *NkBgl* played a role in stabilizing the piperazine ring of HEPES/EPPS/opipramol. To utilize the glucoconjugation feature of *NkBgl* E193D in biotechnological and medical applications, it will be necessary to engineer residues around the aglycone-binding pocket of *NkBgl* E193D to accommodate various acceptors to form different glucopyranosidic products (Fig. 5*c*). Alternatively, other β -glucosidases may be chosen as the starting enzymes in which the respective catalytic acid/base glutamate is changed to aspartate.

In conclusion, in the present study we describe atomic resolution crystal structures of wild-type *NkBgl* in complex with glucose-like glucosidase inhibitors such as gluconolactone and 1-deoxynojirimycin. From these structures, we found that inhibitors that bind to the glycone site may lead HEPES to occupy the aglycone site of *NkBgl*. We also present atomic resolution crystal structures of catalytic acid/base mutants of *NkBgl*, including E193D, E193S and E193A, to provide insight into the enzyme reactions of substrate import and product expulsion. Furthermore, we demonstrate for the first time that the *NkBgl* E193D mutant can transfer the HEPES/EPPS/opipramol moiety onto the glucose moiety of cellobiose/salicin to synthesize glucoconjugated adducts. From a structural viewpoint, the *NkBgl* E193D mutant enzyme first cleaves the glucosidic bond of cellobiose/salicin using its β -glucosidase hydrolase activity and a new bond is subsequently formed between glucose and HEPES/EPPS/opipramol to generate new products through the transglycosylation reaction. These results suggest that the E193D mutant of *NkBgl*

may serve as a useful tool in the synthesis of numerous therapeutic oligosaccharides.

We thank NSRRC, Taiwan and SPring-8 and the Photon Factory, Japan for beam time allocations. We are grateful to Ms Hui-Ling Shr of the Core Facility for Protein Production and X-Ray Structural Analysis (Taipei, Taiwan) for crystallization screening. This work was supported by Academia Sinica and the National Science Council of Taiwan (grant Nos. NSC 97-3112-B-001-017, NSC 97-3114-P-001-001, NSC 98-3112-B-001-024 and NSC 99-3113-B-001-001 to AH-JW).

References

- Bradford, M. M. (1976). *Anal. Biochem.* **72**, 248–254.
- Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* **D50**, 760–763.
- Crout, D. H. & Vic, G. (1998). *Curr. Opin. Chem. Biol.* **2**, 98–111.
- Dwek, R. A. (1996). *Chem. Rev.* **96**, 683–720.
- Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* **D60**, 2126–2132.
- Gopalan, V., Glew, R. H., Libell, D. P. & DePetro, J. J. (1989). *J. Biol. Chem.* **264**, 15418–15422.
- Hanson, S., Best, M., Bryan, M. C. & Wong, C.-H. (2004). *Trends Biochem. Sci.* **29**, 656–663.
- Jeng, W.-Y., Wang, N.-C., Lin, M.-H., Lin, C.-T., Liaw, Y.-C., Chang, W.-J., Liu, C.-I., Liang, P.-H. & Wang, A. H.-J. (2011). *J. Struct. Biol.* **173**, 46–56.
- Kim, Y.-W., Zhang, R., Chen, H. & Withers, S. G. (2010). *Chem. Commun.* **46**, 8725–8727.
- Laskowski, R. A., Chistyakov, V. V. & Thornton, J. M. (2005). *Nucleic Acids Res.* **33**, D266–D268.
- Lovell, S. C., Davis, I. W., Arendall, W. B. III, de Bakker, P. I., Word, J. M., Prisant, M. G., Richardson, J. S. & Richardson, D. C. (2003). *Proteins*, **50**, 437–450.
- Mackenzie, L. F., Wang, Q., Warren, R. A. J. & Withers, S. G. (1998). *J. Am. Chem. Soc.* **120**, 5583–5584.
- Miller, G. L. (1959). *Anal. Chem.* **31**, 426–428.
- Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* **D67**, 355–367.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Perugini, G., Trincone, A., Rossi, M. & Moracci, M. (2004). *Trends Biotechnol.* **22**, 31–37.
- Shaikh, F. A. & Withers, S. G. (2008). *Biochem. Cell Biol.* **86**, 169–177.
- Takahashi, E. & Wraight, C. A. (1991). *FEBS Lett.* **283**, 140–144.
- Tokuda, G., Saito, H. & Watanabe, H. (2002). *Insect Biochem. Mol. Biol.* **32**, 1681–1689.
- Tran, V., Hoffmann, L., Rabiller, C., Tellier, C. & Dion, M. (2010). *Protein Eng. Des. Sel.* **23**, 43–49.
- Umezurike, G. M. (1988). *Biochem. J.* **254**, 73–76.
- Vaguine, A. A., Richelle, J. & Wodak, S. J. (1999). *Acta Cryst.* **D55**, 191–205.
- Varki, A. (1993). *Glycobiology*, **3**, 97–130.
- Wang, L.-X. & Huang, W. (2009). *Curr. Opin. Chem. Biol.* **13**, 592–600.
- Watt, D. K., Ono, H. & Hayashi, K. (1998). *Biochim. Biophys. Acta*, **1385**, 78–88.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.
- Winn, M. D., Murshudov, G. N. & Papiz, M. Z. (2003). *Methods Enzymol.* **374**, 300–321.
- Zopf, D. & Roth, S. (1996). *Lancet*, **347**, 1017–1021.