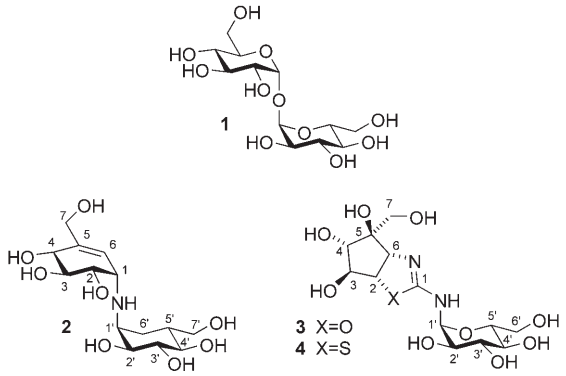
DOI: 10.1002/anie.200604825

Enzyme Structures

Molecular Basis for Trehalase Inhibition Revealed by the Structure of Trehalase in Complex with Potent Inhibitors**

*Robert P. Gibson, Tracey M. Gloster, Shirley Roberts, R. Anthony J. Warren, Isabel Storch de Gracia, Ángela García, Jose L. Chiara, and Gideon J. Davies**

Studies of the unusual nonreducing disaccharide trehalose (α-D-glucopyranosyl α-D-glucopyranoside; **1**; Scheme 1) and its metabolism date back to the isolation of trehalose from the "trehala" cocoons of weevils of the genus *Larinus*.[1] A variety of physiological roles for the disaccharide have been demonstrated in organisms ranging from bacteria and fungi to invertebrates and higher plants. Trehalose is a common source of energy, and its role as a universal stress protectant is also becoming increasingly clear. Cellular damage resulting from temperature extremes, osmotic/oxidative stress, toxic chemical exposure, desiccation, hydrostatic-pressure changes, or nutrient starvation has an effect on the production and utilization of trehalose.[2] Other studies have revealed that trehalose forms the backbone of the structural cord factor from *Mycobacterium tuberculosis*, and recent work has demonstrated that trehalose plays a vital role as a signaling molecule in plant development.[3]



Scheme 1. Trehalose (**1**), and the trehalase inhibitors validoxylamine A (**2**), trehazolin (**3**), and 1-thiatrehazolin (**4**). The numbering of **2**–**4** was chosen to aid comparison with trehalose.

Trehalose is particularly important for insects, as it is hydrolyzed into glucose, which is vital for insect flight. The inhibition of the breakdown of trehalose is, thus, an interesting target for novel insecticides. The hydrolysis of trehalose is catalyzed by glycoside hydrolases, trehalases, which fall into family GH37 of the Carbohydrate-Active Enzyme (CAZy) classification.[4] Although approximately 130 open reading frames are classified in this family, no three-dimensional structure was available for any of these enzymes prior to this work. GH37 trehalases (Enzyme Commission (EC) number 3.2.1.28) hydrolyze one of the two glycosidic bonds in trehalose with inversion of the anomeric configuration.[5] Whilst many trehalase inhibitors are active as insecticides[6,7] (for example, **2** and **3**; Scheme 1), the details of their mode of action were previously unknown. Here we present the first three-dimensional structure of a trehalase, the periplasmic Tre37A enzyme from *Escherichia coli*, which was solved by single-wavelength anomalous dispersion (SAD) methods with diffraction data to 1.5-Å resolution. The enzyme structure was determined in complex with the potent inhibitors validoxylamine A (**2**) and 1-thiatrehazolin (**4**; Scheme 1). Compounds **2**–**4** have been shown to have kinetic inhibition (K_i) and calorimetric dissociation (K_d) constants in the low nanomolar range, with the interactions being characterized by a large favorable enthalpy change (ΔH).

Tre37A was prepared by periplasmic expression using its own signal peptide. This protein was used for all subsequent kinetic, calorimetric, and structural studies. Enzymatic activity was measured using an assay in which glucose was detected using glucose oxidase/peroxidase linking enzymes, following trehalose hydrolysis. At pH 5.5 and 37 °C, Tre37A has a Michaelis constant (K_M) of 0.41 mM and a turnover number (k_{cat}) of 199 s^{-1} (Figure 1). Compounds **2**–**4** have all previously been reported to be tight-binding slow-onset inhibitors of trehalase,[8,9] notably with the pig-kidney enzyme. In our laboratory, trehazolin (**3**) decayed after storage at −20 °C, highlighting the advantages of 1-thiatrehazolin (**4**), which remains stable for (at least) several years. Inhibition studies (performed following a 20-min pre-incubation between the enzyme and inhibitor to mitigate slow-onset inhibition) show that **2** and **4** are indeed very potent inhibitors of Tre37A, with K_i values of 10 and 9 nM, respectively. Isothermal titration calorimetry (carried out at pH 7 and 25 °C) also revealed that both **2** and **4** are tight-binding inhibitors of Tre37A, with K_d values of 7 and 15 nM, respectively (Figure 2). Both binding events are characterized by a large favorable enthalpy of binding (ΔH values of −18.9 and −12.7 kcal mol^{-1} for **2** and **4**,

[*] Dr. R. P. Gibson, Dr. T. M. Gloster, S. Roberts, Prof. R. A. J. Warren, Prof. G. J. Davies
Department of Chemistry
York Structural Biology Laboratory
York, YO10 5YW (UK)
Fax: (+44) 1904-328266
E-mail: davies@ysbl.york.ac.uk
I. Storch de Gracia, Á. García, Dr. J. L. Chiara
Instituto de Química Orgánica General, CSIC
Juan de la Cierva, 3
28006 Madrid (Spain)

[**] R.P.G. and T.M.G. contributed equally to this work. This work was funded by the Biotechnology and Biological Sciences Research Council (BBSRC). G.J.D. holds a Royal Society Wolfson Merit Award. The authors thank Anthony O'Sullivan (Syngenta) for the provision of compound **2** and for useful discussions.

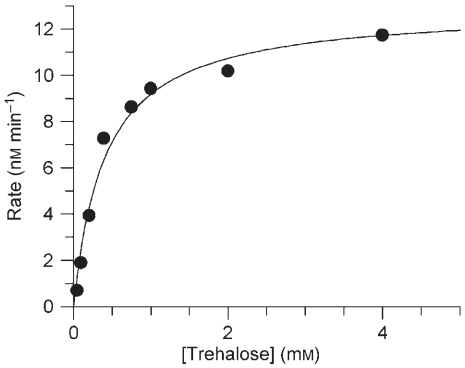


Figure 1. Michaelis–Menten kinetics for the hydrolysis of trehalose by Tre37A; the reaction rate is plotted against trehalose concentration.

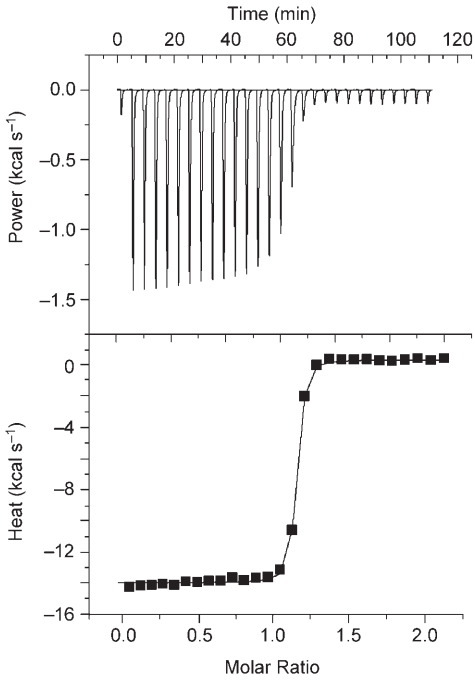


Figure 2. Isothermal titration calorimetry (ITC) data for the binding of **4** to Tre37A. Top: Raw titration data of the power supplied to the system to maintain a constant temperature plotted against time. Bottom: The bimolecular fit of the normalized heats of interaction plotted against the molar ratio of **4** to Tre37A.

respectively), but an unfavorable entropy change ($T\Delta S$ values of -7.8 and -2.1 kcal mol^{-1} for **2** and **4**, respectively).

The three-dimensional structure of Tre37A was solved using SAD methods with a selenomethionine (SeMet) derivative of the enzyme, which was initially crystallized in the presence of **2**. X-ray diffraction data were collected to 2.2-Å resolution, and the Tre37A model could be traced from Pro37 to Pro547 (final R_{cryst} of 0.16 and R_{free} of 0.20), with some of the surface residues being disordered. Subsequently, the protein coordinates from this structure were used as the starting model for the refinement of the structure of a Tre37A complex with **4**, with data to 1.5-Å resolution (final R_{cryst} of 0.15 and R_{free} of 0.19).

The structure of Tre37A, the first of a trehalase, consists of an $(\alpha/\alpha)_6$ barrel (Figure 3a) similar to that found for other α-toroidal glycosidases, notably GH94 chitobiose phosphorylases,[10] GH15 glucoamylases,[11] and GH65 maltose phosphorylases.[12] All of these enzymes display the $(\alpha/\alpha)_6$ toroid fold; they are classified in the glucoamylase family by the secondary-structure matching (SSM) server,[13] and all perform catalysis with inversion of the anomeric configuration, as is the case for the GH37 trehalases.

The complexes of Tre37A with **2** and **4** (Figure 3b,c) reveal the interactions in the -1 and $+1$ subsites. In both complexes, the positions of the protein side chains are invariant, but intriguingly, the disaccharide mimics are completely buried within the structure. Significant conformational change would be required for substrate entry and departure into the occluded active centre. Such conformational change may contribute both to the slow onset of inhibition observed in kinetic studies by others[9] and the unfavorable entropy of interaction. In the leaving-group $+1$

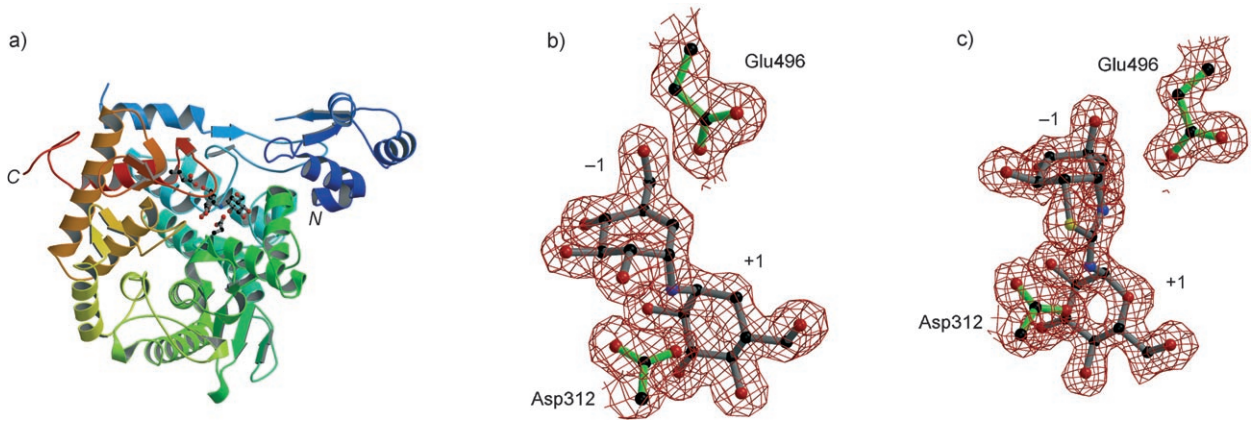


Figure 3. a) Ribbon representation of Tre37A in complex with validoxylamine A (**2**), color-ramped from the N terminus (blue) to the C terminus (red); **2** and the residues proposed to act as the acid (Asp312) and base (Glu496) in the hydrolysis mechanism are shown in ball-and-stick representation. Ball-and-stick representations of b) **2** and c) 1-thiatrehazolin (**4**) bound to Tre37A (residues Glu496 and Asp312 are shown; $+1$ and -1 subsites are labeled); the observed electron density for the maximum-likelihood weighted $2F_{obs}-F_{calcd}$ map is contoured at 1σ. C black, N blue, O red, S yellow. The figure was drawn using BOBSCRIPT.[14]

subsite, the pseudosugar ring of **2** and the glucose moiety of **4** (which both have a $^4C_1$ chair conformation) lie in a nearly identical position (Figure 4), which promotes favorable hydrogen bonding with residues Arg152, Asn196, Arg205, Arg277, Glu279, and Glu511 (Figure 5). As was also found for

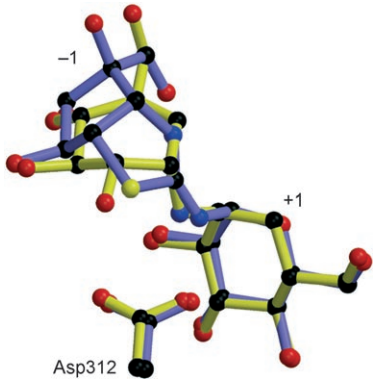


Figure 4. Ball-and-stick representation of the superposition of validoxylamine A (**2**; yellow) and 1-thiatrehazolin (**4**; blue) bound to Tre37A (residue Asp312 is shown; +1 and −1 subsites are labeled). C black, N blue, O red, S yellow.

other enzymes such as glucoamylases, there is no classical hydrophobic platform in this subsite. This feature, in tandem with the hydrogen-bonding template, provides the driving force for the accommodation of the α,α-1,1-linked disaccharide, as opposed to α-1,4-linked substrates. The extensive hydrogen bonding and the lack of hydrophobic stacking in the +1 subsite provide a persuasive structural rationale for the reduced affinity of simplified hydrophobic derivatives of **2**–**4**.[15,16]

Catalysis by trehalases is known to occur with inversion of the anomeric configuration, as is also the case for its closest structural homologues from families GH15 and GH65. Glycoside hydrolysis with inversion of the anomeric configuration follows a canonical mechanism: proton assistance to leaving-group departure is complemented by Brønsted base assistance to nucleophilic attack by a water molecule, in a single-displacement mechanism. The structure does not support a mechanism involving ring opening, which was once considered.[5] The complexes of Tre37A with **2** and **4**, and the similarity of the active centre to those of glucoamylases, strongly implicate Asp312 (spatially approximate to Glu179 of *Aspergillus* glucoamylase[11]) and Glu496 (directly equivalent to Glu400 of *Aspergillus* glucoamylase[11]) as the catalytic acid and base, respectively, in the inversion mechanism. Asp312 is poised 2.8 Å from the "glycosidic" nitrogen atom of both **2** and **4** (Figures 3 and 4). Glu496 sits above the β face of the pseudosugar ring in **2**, and is poised to activate a water molecule (located 4 Å from C1 of **2**) for nucleophilic attack. In the complex with **4**, no electron density that could be assigned to a potential water molecule is detected, as the hydroxy group at C5 of **4** occupies a similar position to that occupied by the putative attacking water molecule in the complex with **2**. The hydroxy group of Tyr512 may impede the attacking water molecule, which implies that small conformational rearrangements of the determined structures are necessary for catalysis.

In the −1 subsite, there are differences in the positions of the valienamine and tetrahydrocyclopentathiazole moieties of **2** and **4**, and the interactions they make with Tre37A. In the "catalytic" −1 subsite, the valienamine ring of **2** has a $^2H_3$ half-chair conformation, as dictated by the double bond between C5 and C6, whilst the cyclopentane ring of **4** adopts a ^2E envelope conformation. As expected, **2** binds essentially as seen previously for the inhibition of α-1,4-glucan-hydrolyzing enzymes by acarbose.[11,17] Surprisingly, despite substantial differences in their chemical structures, **2** and **4** make similar interactions with the protein.

The hydroxy groups at C3 and C4 of **2** superimpose well with their equivalents in **4** (Figure 4); in both cases, these hydroxy groups make hydrogen bonds with Gly310 (main chain), Trp159, Gln207, and Asp160. The hydroxy group at C2 of **2** interacts with Trp447 and Asp312, and although **4** possesses no equivalent exocyclic substituent, the sulfur atom of the thiazoline ring makes comparable interactions. In a similar compensatory way, the hydroxy group of the hydroxymethyl group at C5 of **2** hydrogen bonds with Asp160, as does the axial hydroxy group at C5 of **4** (Figures 4 and 5). In both **2** and **4**, the nitrogen atom equivalent to the glycosidic oxygen atom of trehalose interacts with the catalytic acid Asp312. Presumably, the favorable geometry of this interaction, enhanced by the potentially transition-state-mimicking conformations of the two moieties in the −1 subsite, contributes significantly to the high affinity. Furthermore, the nitrogen atom of the thiazoline ring of **4** lies in the position that the endocyclic O5 atom would occupy in the natural substrate, which is likely to contribute to the favorable enthalpy of interaction of **4**.

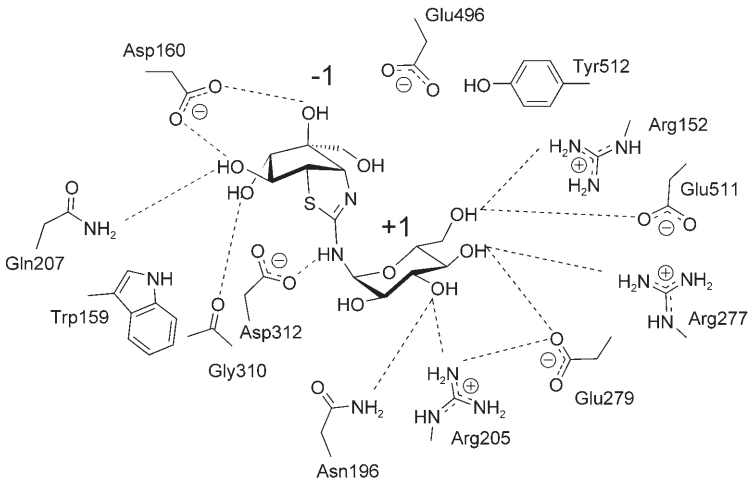


Figure 5. Interactions between 1-thiatrehazolin (**4**) and Tre37A. Hydrogen bonds are shown by dashed lines.

A number of other glycosidase inhibitors feature thiazoline moieties. Allosamidin, for example, possesses a tetrahydrocyclopentoxazole ring system similar to that of **4** and is a potent chitinase

inhibitor.[18] Similarly, thiazoline derivatives of *N*-acetylglucosamine are tight-binding hexosaminidase inhibitors.[19] In this case, however, the thiazoline moiety mimics the covalent oxazolinium ion intermediate, or a closely related transition state,[20] formed during the substrate-assisted double-displacement reaction of a retaining glycosidase. In contrast, **3** and **4** inhibit an inverting glycosidase, and their favorable tight binding reflects an adventitious mimicry of the hydrogen bonding of the transition state during a single-displacement reaction mechanism. This situation is similar to general glycosidase inhibition using five-membered rings (recently described for mannostatins[21]). The exact nature of the transition state conformation for glycoside hydrolysis by trehalases remains undefined, with both $^{4}H_{3}$ half-chair and $^{2,5}B$ boat conformations of the six-membered rings being possible (by virtue of work on analogous systems[22,23]). A feature of compounds **2**–**4** is that, whilst none closely mimic either the $^{4}H_{3}$ or the $^{2,5}B$ conformation, all are reasonable mimics of both conformations. To some extent, the conformational flexibility of the hydroxymethyl group at C5 of **2** and of both the axial and the equatorial substituents at C5 in **3** and **4**, allows the compounds to optimize hydrogen bonding, despite the apparent differences in ring conformations. Such strategies to overcome the chemical challenges of transition-state mimicry, which build on the benefits of aglycon interactions (as illustrated by compounds **2**–**4**), should provide inspiration for future generations of glycosidase inhibitors.

Experimental Section

The gene encoding Tre37A was amplified by the polymerase chain reaction (PCR) and ligated into the *Nde*I and *Xho*I sites of a pET22b expression vector (Novagen) for periplasmic expression with the native Tre37A signal peptide. *E. coli* BL21 (DE3) cells were transformed with the plasmid containing the Tre37A gene and grown at 37 °C in Luria–Bertani (LB) broth containing ampicillin (50 μg mL^{-1}), until an optical density at 600 nm (OD_{600}) of approximately 0.7 AU was reached. Over-expression was initiated by the addition of isopropyl thio-β-D-galactoside to a final concentration of 1 mM, and cells continued to grow at 30 °C for 12 h. Cells were harvested by centrifugation, and Tre37A was obtained by osmotically shocking the cells. Tre37A was purified by ion-exchange chromatography (HiTrap MonoQ, Amersham Biosciences), and by gel filtration (Superdex 200 16/60), using standard procedures. A selenomethionine (SeMet) derivative of Tre37A was prepared using the auto-induction methodology of Studier[24] and purified in the same way.

Validoxylamine A (**2**) was prepared by the acid removal of glucose from validamycin A, which was isolated from an agricultural fungicide (Takeda Chemical Company; according to the method in reference [30]). 1-Thiatrehazolin (**4**) was prepared as described in reference [9].

Tre37A (15 mg mL^{-1}) in the presence of **2** (1 mM) was crystallized from poly(ethylene glycol) (PEG) 3350 (25% w/v) and bis(2-hydroxyethyl)aminotris(hydroxymethyl)methane (Bis-Tris)/HCl buffer (0.1 M, pH 6.5). These crystals were used as seeds for the SeMet protein and for the native protein with **4** (1 mM). X-ray diffraction data for SeMet Tre37A in complex with **2** and native Tre37A in complex with **4** were collected at 100 K at the European Synchrotron Radiation Facility (ESRF) on beamlines ID14-4 and ID29, respectively. A single-wavelength anomalous dispersion (SAD) experiment, optimized for the f'' component of the anomalous scattering, was used to collect the data for SeMet Tre37A. The data were processed using MOSFLM[25] from the CCP4 suite.[26] Heavy-atom phasing was performed with autoSHARP,[27] subsequent model building with COOT,[28] and refinement with REFMAC.[29] The atomic coordinates have been deposited in the Protein Data Bank (PDB) with accession codes 2JF4 (SeMet Tre37A with **2**) and 2JG0 (Tre37A with **4**); additional experimental details are given in the PDB headers.

The kinetics of the hydrolysis of trehalose by Tre37A was investigated using a stopped assay, in which glucose was detected using glucose oxidase/peroxidase linking enzymes (Megazyme). Assays were performed at 37 °C in sodium maleate buffer (75 mM, pH 5.5) containing bovine serum albumin (BSA; 0.38 mg mL^{-1}). Measurements were made at trehalose concentrations of 0.05–4 mM; Tre37A was present at a final concentration of 5.4 nM. 200-μL aliquots of the reaction mixture were taken at time intervals, and after boiling the aliquots for 2 min, 1 mL of the glucose oxidase/peroxidase solution was added. Assays were otherwise performed as described in the manufacturer's instructions. The reaction rates were determined and the data were fitted to the Michaelis–Menten equation using GRAFIT (Erithacus Software Ltd.). The inhibition constant (K_i) values were determined in the same way, in the presence of **2** or **4** (10 nM) and at a final Tre37A concentration of 1 nM (to ensure a 10-fold excess of the inhibitor). Tre37A was pre-incubated with each of the inhibitors for 20 min prior to the start of the assay to prevent any complications with slow-onset inhibition. The data were similarly fitted using GRAFIT to obtain an apparent Michaelis constant (K_M^{app}), and K_i values were determined using the equation $K_M^{app} = K_M(1+[I]/K_i)$, where [I] is the concentration of the inhibitor.

Isothermal titration calorimetry (ITC) was performed using a VC calorimeter (Microcal) at 25 °C. Tre37A (31–37 μM) was dialyzed into sodium phosphate buffer (50 mM, pH 7.0), and inhibitors **2** and **4** (0.5 mM) were diluted in the same buffer. Titrations were performed by injecting 10-μL aliquots of the ligand into Tre37A. The data were corrected for the heat of dilution by subtracting the heat at a high molar ratio of inhibitor to enzyme. The stoichiometry (n), enthalpy change (ΔH), and equilibrium association constant (K_a) were determined by fitting a bimolecular model to the data using Microcal Origin. The Gibbs free energy (ΔG) and $T\Delta S$ were calculated using the equations $\Delta G = -RT\ln K_a = \Delta H - T\Delta S$.

Received: November 28, 2006
Published online: April 23, 2007

Keywords: inhibitors · protein structures · trehalases · trehazolins · validoxylamines

[1] G. R. Wyatt, G. F. Kalf, *J. Gen. Physiol.* **1957**, *40*, 833.
[2] J. H. Crowe, L. M. Crowe, D. Chapman, *Science* **1984**, *223*, 701.
[3] N. Avonce, B. Leyman, J. Thevelein, G. Iturriaga, *Biochem. Soc. Trans.* **2005**, *33*, 276.
[4] P. M. Coutinho, B. Henrissat in *Recent Advances in Carbohydrate Bioengineering* (Eds.: H. J. Gilbert, G. J. Davies, B. Henrissat, B. Svensson), Royal Society of Chemistry, Cambridge, **1999**, pp. 3.
[5] J. Defaye, H. Driguez, B. Henrissat, *Carbohydr. Res.* **1983**, *124*, 265.
[6] S. N. Thompson, *Adv. Insect Physiol.* **2003**, *31*, 205.
[7] M. C. P. Silva, W. R. Terra, C. Ferreira, *Comp. Biochem. Physiol. Part B* **2006**, *143*, 367.
[8] H. M. Salleh, J. F. Honek, *FEBS Lett.* **1990**, *262*, 359.
[9] J. L. Chiara, I. S. de Gracia, Á. García, Á. Bastida, S. Bobo, M. D. Martín-Ortega, *ChemBioChem* **2005**, *6*, 186.
[10] M. Hidaka, Y. Honda, M. Kitaoka, S. Nirasawa, K. Hayashi, T. Wakagi, H. Shoun, S. Fushinobu, *Structure* **2004**, *12*, 937.
[11] A. E. Aleshin, L. M. Firsov, R. B. Honzatko, *J. Biol. Chem.* **1994**, *269*, 15631.

[12] M. P. Egloff, J. Uppenberg, L. Haalck, H. van Tilbeurgh, *Structure* **2001**, *9*, 689.
[13] E. Krissinel, K. Henrick, *Acta Crystallogr. Sect. D* **2004**, *60*, 2256.
[14] R. M. Esnouf, *J. Mol. Graphics Modell.* **1997**, *15*, 132.
[15] Y. Kobayashi, *Carbohydr. Res.* **1999**, *315*, 3.
[16] R. Lysek, C. Schutz, S. Favre, A. C. O'Sullivan, C. Pillonel, T. Krulle, P. M. J. Jung, I. Clotet-Codina, J. A. Este, P. Vogel, *Bioorg. Med. Chem.* **2006**, *14*, 6255.
[17] A. M. Brzozowski, D. M. Lawson, J. P. Turkenburg, H. Bisgaard-Frantzen, A. Svendsen, T. V. Borchert, Z. Dauter, K. S. Wilson, G. J. Davies, *Biochemistry* **2000**, *39*, 9099.
[18] F. V. Rao, D. R. Houston, R. G. Boot, J. M. Aerts, S. Sakuda, D. M. van Aalten, *J. Biol. Chem.* **2003**, *278*, 20110.
[19] S. Knapp, D. J. Vocadlo, Z. N. Gao, B. Kirk, J. P. Lou, S. G. Withers, *J. Am. Chem. Soc.* **1996**, *118*, 6804.
[20] G. E. Whitworth, M. S. Macauley, K. A. Stubbs, R. J. Dennis, E. J. Taylor, G. J. Davies, I. R. Greig, D. J. Vocadlo, *J. Am. Chem. Soc.* **2007**, *129*, 635.
[21] S. P. Kawatkar, D. A. Kuntz, R. J. Woods, D. R. Rose, G. J. Boons, *J. Am. Chem. Soc.* **2006**, *128*, 8310.
[22] Y. Tanaka, W. Tao, J. S. Blanchard, E. J. Hehre, *J. Biol. Chem.* **1994**, *269*, 32306
[23] G. J. Davies, V. M.-A. Ducros, A. Varrot, D. L. Zechel, *Biochem. Soc. Trans.* **2003**, *31*, 523.
[24] F. W. Studier, *Protein Expression Purif.* **2005**, *41*, 207.
[25] A. G. W. Leslie, *Joint CCP4 and ESF–EAMCB Newsletter on Protein Crystallography* **1992**, 26.
[26] Collaborative Computational Project Number 4; *Acta Crystallogr. Sect. D* **1994**, *50*, 760.
[27] E. de la Fortelle, G. Bricogne, *Methods Enzymol.* **1997**, *276*, 472.
[28] P. Emsley, K. Cowtan, *Acta Crystallogr. Sect. D* **2004**, *60*, 2126.
[29] G. N. Murshudov, A. A. Vagin, E. J. Dodson, *Acta Crystallogr. Sect. D* **1997**, *53*, 240.
[30] Y. Kameda, N. Asano, T. Yamaguchi, K. Matsui, S. Horii, H. Fukase, *J. Antibiot.* **1986**, *39*, 1491.

