BIOLOGICAL CRYSTALLOGRAPHY

ISSN 1399-0047

Received 30 October 2014
Accepted 21 April 2015

Edited by P. Langan, Oak Ridge National Laboratory, USA

‡ Present address: LISM, UMR7255 CNRS, 31 Chemin Joseph Aiguier, 13402 Marseille CEDEX 20, France and AFMB–UMR7257 CNRS–Aix-Marseille Université–Case 932, 163 Avenue de Luminy, 13288 Marseille CEDEX 09, France.

Keywords: CHIT1; GH18 chitinase; crystal structures; protonation states; hydrolysis; catalytic mechanism.

PDB references: human chitotriosidase catalytic domain, 4wjx; 4wka; complex with chitobiose (2.5 m*M*), 4wkf; complex with chitobiose (1 m*M*), 4wkh; complex with chitobiose (300 m*M*), 4wk9

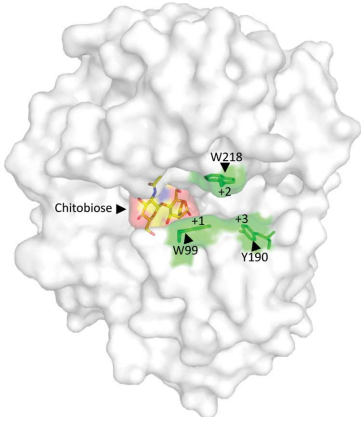
Supporting information: this article has supporting information at journals.iucr.org/d

New insights into the enzymatic mechanism of human chitotriosidase (CHIT1) catalytic domain by atomic resolution X-ray diffraction and hybrid QM/MM

Firas Fadel,[a]*‡ Yuguang Zhao,[b] Raul Cachau,[c] Alexandra Cousido-Siah,[a] Francesc X. Ruiz,[a] Karl Harlos,[b] Eduardo Howard,[a,d] Andre Mitschler[a] and Alberto Podjarny[a]*

[a]Department of Integrative Structural Biology, Institut de Génétique et de Biologie Moléculaire et Cellulaire, CNRS/INSERM/UdS, 1 Rue Laurent Fries, 67404 Illkirch CEDEX, France, [b]Division of Structural Biology, Wellcome Trust Centre for Human Genetics, Oxford University, Roosevelt Drive, Headington, Oxford, England, [c]Leidos Biomedical Research Inc. Advanced Biomedical Computer Center, Information Systems Program, Frederick National Laboratory for Cancer Research, Frederick, Maryland, USA, and [d]IFLYSIB, Conicet, UNLP, Calle 59 No. 789, La Plata, Argentina. *Correspondence e-mail: firasfadel1@gmail.com, podjarny@igbmc.fr

Chitotriosidase (CHIT1) is a human chitinase belonging to the highly conserved glycosyl hydrolase family 18 (GH18). GH18 enzymes hydrolyze chitin, an *N*-acetylglucosamine polymer synthesized by lower organisms for structural purposes. Recently, CHIT1 has attracted attention owing to its upregulation in immune-system disorders and as a marker of Gaucher disease. The 39 kDa catalytic domain shows a conserved cluster of three acidic residues, Glu140, Asp138 and Asp136, involved in the hydrolysis reaction. Under an excess concentration of substrate, CHIT1 and other homologues perform an additional activity, transglycosylation. To understand the catalytic mechanism of GH18 chitinases and the dual enzymatic activity, the structure and mechanism of CHIT1 were analyzed in detail. The resolution of the crystals of the catalytic domain was improved from 1.65 Å (PDB entry 1waw) to 0.95–1.10 Å for the apo and pseudo-apo forms and the complex with chitobiose, allowing the determination of the protonation states within the active site. This information was extended by hybrid quantum mechanics/molecular mechanics (QM/MM) calculations. The results suggest a new mechanism involving changes in the conformation and protonation state of the catalytic triad, as well as a new role for Tyr27, providing new insights into the hydrolysis and transglycosylation activities.



© 2015 International Union of Crystallography

1. Introduction

Chitinases belong to the class of glycosyl hydrolases (GHs) that degrade chitin, an abundant natural polysaccharide, by cleaving the β-(1,4) linkages of its *N*-acetylglucosamine (NAG) chain (Gooday, 1990). Chitinases can be subdivided into two families, family 18 (GH18) and family 19 (GH19), that differ in structure and mechanism (Henrissat & Davies, 1997). GH18 chitinases have been identified in many organisms varying from lower organisms to humans. Interestingly, chitinases fulfill diverse functional roles in different species. While they ensure carbon and nitrogen sources by degrading chitin in bacteria, they are involved in growth and morphogenesis in fungi and insects. Further, chitinases have been shown to perform a protective role against chitin-containing pathogens in plants and mammals.

Files: d/lp5005/lp5005.3d d/lp5005/lp5005.sgml **LP5005** FA IU-1512/21(26)5 1512/20(26)5 () LP5005 PROOFS D:FA:2015:71:5:0:0–0

In the past decade, several crystal structures of GH18 chitinases have been solved. According to these structures, the catalytic domains of this enzyme family consist of an $(\alpha/\beta)_8$ TIM-barrel fold, with a high degree of conservation in their active-site cleft composed of aromatic residues that contribute to substrate binding (van Aalten *et al.*, 2000; Perrakis *et al.*, 1994; Fusetti *et al.*, 2002; Olland *et al.*, 2009; Yang *et al.*, 2010; Terwisscha van Scheltinga *et al.*, 1996). The cleavage of the chitin polymer takes place between subsites −1 and +1. The catalytic triad is positioned at subsite −1, which is located at the bottom of the substrate-binding cleft. The highly conserved motif (D*X*D*X*E) that characterizes the GH18 chitinases (van Aalten *et al.*, 2001) includes the catalytic triad, which consists of two aspartates and a glutamate. The glutamate of this motif has been identified as the key catalytic acid/base residue, which is presumed to be protonated on the outer O atom of its side chain. In general, the majority of GH18 apo-form crystal structures have shown the middle aspartate to form a short hydrogen bond to the first aspartate (conformation *A*). However, it has been reported that in the presence of the substrate the middle aspartate turns towards the catalytic glutamate and forms a hydrogen bond to this glutamate (conformation *B*) (van Aalten *et al.*, 2000). In addition to its participation in catalysis, the middle aspartate has been reported to assist in stabilization of the conformation of the substrate subsequent to its binding (Fusetti *et al.*, 2002; Olland *et al.*, 2009; van Aalten *et al.*, 2000; Songsiriritthigul *et al.*, 2008).

Regarding the enzymatic mechanism, GH18 chitinases are considered to be retaining enzymes, which implies the retention of the initial anomeric carbon configuration in the hydrolysis product (White & Rose, 1997; Davies & Henrissat, 1995). Previous studies have proposed that the binding of the substrate generates a distortion of the −1 NAG subunit into a boat conformation preceding the formation of the non-covalent oxazolinium-ion intermediate. The acetamido group of the −1 distorted sugar performs the nucleophilic attack, enabling the formation of this intermediate ion, *i.e.* the so-called substrate-assisted mechanism (Brameld & Goddard, 1998; Tews *et al.*, 1997). Simultaneously with the nucleophilic attack, the catalytic glutamate protonates the glycosidic O atom located between the −1 and +1 NAGs, which leads to the cleavage of chitin (van Aalten *et al.*, 2001; Suginta & Sritho, 2012).

In addition to the hydrolysis reaction, many chitinases, including human chitinases, show a distinctive property consisting of the capacity to shift the activity from hydrolysis to transglycosylation (TG) in the presence of excess substrate concentrations. In TG, the enzymes catalyze the formation of glycosidic bonds between donor and acceptor sugar units, which leads to repolymerization of the substrate (Zakariassen *et al.*, 2011). However, the precise mechanism of this reaction is not yet clearly understood.

By combining X-ray data, biochemical experiments and computational calculations, extensive efforts have been devoted to elucidating the structure–function relationship, including the features of substrate binding, in GH18 chitinases. Despite these efforts, the detailed structural mechanistic basis is not fully understood and many aspects remain controversial. Indeed, the protonation pattern of the catalytic site and the proton pathway during the enzymatic reaction have not yet been elucidated. To determine the protonation states of the catalytic site, it is necessary to obtain crystals of better quality to achieve a sufficiently high resolution (1 Å or better).

In this study, we investigated the catalytic domain of the human chitotriosidase (CHIT1) as a model to probe the mode of action of the GH18 chitinases. This enzyme is one of two active chitinases that have been identified in humans and have been reported to be involved in the innate immune response as well as being a biomarker of Gaucher disease (Hollak *et al.*, 1994). Further to its hydrolysis activity, it shows a high transglycosylation rate, which is a widespread phenomenon in GH18 chitinases (Aguilera *et al.*, 2003). In 2002, the first X-ray crystal structure of CHIT1 (PDB entry 1guv) was determined at 2.35 Å resolution by Fusetti and coworkers (Fusetti *et al.*, 2002); subsequently, Rao and coworkers obtained a structure of CHIT1 in complex with the inhibitor argifin at 1.65 Å resolution (PDB entry 1waw; Rao *et al.*, 2005). These structures showed that this enzyme shares the TIM-barrel three-dimensional folding of GH18 chitinases, with the conserved catalytic motif (D*X*D*X*E) located in the active site (Fusetti *et al.*, 2002). However, more detailed information is still required for a full explanation of proton-translocation processes. Thus, in order to extend our knowledge regarding the catalytic properties, we obtained new X-ray data to resolutions of between 0.95 and 1.10 Å for CHIT1 in apo and pseudo-apo forms and in complex with chitobiose, an *N*-acetylglucosamine (NAG) dimer. Our crystal structures reveal the detailed internal organization of the active-site residues as well as their interactions with chitobiose and allow us to investigate the protonation state of the catalytic triad. Quantum-mechanics calculations further supplemented our crystallographic findings, confirming the observed protonation states of the catalytic triad and providing novel insights into the proton pathway during the hydrolytic reaction.

2. Materials and methods

2.1. Cloning, expression and purification

The full-length human chitotriosidase CHIT1 gene was used as a template to generate the construct corresponding to the catalytic domain (residues 1–386; CHIT1) by the polymerase chain reaction (PCR) using the following primers (Sigma): 5′-AATTCAAGCTTGCCACCATGGTGCGGTCTGTGG-3′ (N-terminal derived sense primer) and two antisense primers to generate the 3′ end encoding Ser386 with an additional thrombin site and a His tag, 5′-GTGATGGTGATGGTG-GTGAGAACCGCGTGGCACCAGACTCAGTTCCTGCC-GTAGC-3′ and 5′-ATTATCGCGATACTAGTCTCGAGT-CATTAGTGATGGTGATGGTGGTG-3′. The PCR product was cloned into the pHL expression vector (Aricescu *et al.*, 2006). CHIT1 was transiently expressed in adherent HEK293T cells grown in roller bottles as described by Zhao *et*

Files: d/lp5005/lp5005.3d d/lp5005/lp5005.sgml LP5005 FA IU-1512/21(26)5 1512/20(26)5 () LP5005 PROOFS D:FA:2015:71:5:0:0–0

Table 1
Data-collection and refinement statistics for CHIT1.

Values in parentheses are for the highest resolution shell.

	Pseudo-apo form	Apo form	Chitobiose, 0.3 m*M*	Chitobiose, 1 m*M*	Chitobiose, 2.5 m*M*
PDB code	4wka	4wjx	4wk9	4wkh	4wkf
Synchrotron, beamline	SLS, X06DA (PXIII)	SLS, X06DA (PXIII)	SLS, X06DA (PXIII)	SLS, X06DA (PXIII)	SLS, X06DA (PXIII)
Wavelength (Å)	0.8	0.8	0.8	0.8	0.8
Resolution range (Å)	50–0.95 (0.98–0.95)	50–1.00 (1.04–1.00)	50–1.10 (1.14–1.10)	50–1.05 (1.09–1.05)	50–1.10 (1.14–1.10)
Space group	$P2_12_12$	$P2_12_12$	$P2_12_12$	$P2_12_12$	$P2_12_12$
Unit-cell parameters (Å)					
a	85.33	85.69	85.67	85.50	85.502
b	103.70	105.75	106.18	105.52	103.434
c	41.69	41.52	41.43	41.475	841.58
Total reflections	2811923 (19720)	1121774 (10116)	1805445	1091078	1041198
Unique reflections	223038 (18926)	188444 (9732)	152249	175071	144158 (13705)
Multiplicity	12.6 (9.9)	6.0 (2.4)	11.9 (8.6)	6.2 (5.0)	7.2 (5.7)
Completeness (%)	95.99 (82.51)	92.51 (48.18)	99.51 (95.97)	99.78 (98.46)	96.32 (92.70)
Mean $I/\sigma(I)$	36.08 (2.57)	27.50 (1.77)	40.0 (3.57)	31.92 (2.24)	20.38 (2.61)
Wilson *B* factor (Å^2)	8.82	10.13	9.82	10.59	7.66
R_{merge}	0.059 (0.766)	0.041 (0.439)	0.057 (0.552)	0.046 (0.681)	0.085 (0.602)
R factor	0.1142 (0.194)	0.1355 (0.231)	0.1396 (0.167)	0.1434 (0.220)	0.1433 (0.189)
R_{free}	0.1222 (0.189)	0.1476 (0.247)	0.1545 (0.185)	0.1534 (0.2293)	0.1622 (0.2083)
No. of atoms					
Total	7587	4107	3660	3778	6516
Macromolecules	3518	3526	3096	3171	3062
Ligands	20	40	29	87	58
Water molecules	586	526	535	512	367
No. of protein residues	369	369	370	369	368
R.m.s.d., bonds (Å)	0.008	0.005	0.006	0.013	0.011
R.m.s.d., angles (°)	1.38	1.15	1.19	1.35	1.36
Ramachandran favoured (%)	98	98	98	99	99
Ramachandran outliers (%)	0	0	0	0	0
Clashscore	5.65	3.87	3.18	3.45	3.06
Average *B* factors (Å^2)					
Overall	12.70	13.80	13.40	14.00	10.20
Macromolecules	10.80	12.10	11.50	12.20	9.20
Ligands	24.80	29.00	13.90	14.50	7.50
Solvent	22.90	23.90	24.50	24.80	18.70

al. (2011). After dialysis against 25 m*M* phosphate-buffered saline (PBS) pH 8.0 at 4°C, the secreted protein was purified from the medium using an immobilized metal-affinity chromatography (IMAC) batch procedure. CHIT1 was further purified by size-exclusion chromatography on a Superdex 200 16/60 (GE Healthcare) column in 10 m*M* HEPES, 150 m*M* NaCl pH 7.5. The protein purity was assessed by SDS–PAGE followed by Coomassie Brilliant Blue staining (Laemmli, 1970). The enzyme concentration was determined from the absorption at 280 nm using an UV NanoDrop 1000 spectrophotometer (Thermo Scientific). The molar extinction coefficient was calculated as 73 590 M^{-1} cm^{-1} using the *ProtParam* tool on the ExPasy server (Gasteiger *et al.*, 2005).

2.2. Enzyme-activity measurements

The enzymatic activity was determined using commercial synthetic fluorogenic substrates (Hollak *et al.*, 1994). Briefly, 0.25 n*M* CHIT1 was incubated at 37°C with various concentrations (0–200 µ*M*) of the substrate 4-methylumbelliferyl *β*-*N*,*N*′,*N*″-triacetylchitotrioside (4-MU-NAG$_3$; Sigma). The assays were performed in 100 m*M* phosphate–citrate buffer pH 5.6 containing 1 mg ml^{-1} bovine serum albumin. After 30 min, the assay was stopped by the addition of 100 µl 0.3 *M* glycine–NaOH pH 10.3. The product 4-methylumbelliferone was quantified using a microplate fluorometer (excitation at 360 nm/emission at 440 nm).

2.3. Crystallization and data collection

For crystallization of the apo form of CHIT1, the protein was concentrated to 9 mg ml^{-1} in 10 m*M* HEPES, 150 m*M* NaCl pH 7.5 buffer and crystals were grown by the hanging-drop vapour-diffusion method at 17°C by mixing 1.5 µl protein solution with an equal volume of reservoir solution containing a low concentration of micro-seeds. The crystals reached maximum dimensions of 1 × 0.14 × 0.10 mm after 4–6 d of equilibration against 500 µl reservoir solution consisting of 24–26%(*w*/*v*) polyethylene glycol (PEG) 3350, 0.2 *M* potassium sodium tartrate (PST) at pH 7.2.

The pseudo-apo crystal form was obtained after failing to co-crystallize the protein with chitin (chitin from shrimp shells; Sigma). A saturated stock solution of chitin was prepared in DMSO and dissolved in the reservoir solution (also containing a low concentration of micro-seeds), which was added to the drop. The X-ray data obtained from this crystal were collected at 0.95 Å resolution but chitin was not observed; this structure was therefore considered to be a pseudo-apo form.

Crystals containing the artificial substrate 4-MU-NAG$_3$ (Sigma) were obtained in the same crystallization condition as

Files: d/lp5005/lp5005.3d d/lp5005/lp5005.sgml **LP5005** FA IU-1512/21(26)5 1512/20(26)5 () LP5005 PROOFS D:FA:2015:71:5:0:0–0

the apo form of CHIT1 combined with micro-seeding by mixing 1.5 µl protein solution with an equal volume of reservoir solution containing a range of final concentrations of 4-MU-NAG_3 of between 0.3 and 2.5 m*M*. Hydrolysis of the substrate occurred in the drop; therefore, a crystal of CHIT1 complexed with only two *N*-acetylglucosamine monomers, *i.e.* chitobiose, appeared after 10–30 d. All of the crystals obtained of the apo form, the pseudo-apo form or the complex with chitobiose were cryocooled in liquid nitrogen using a solution containing 35% PEG 3350, 0.2 *M* PST.

X-ray diffraction data were collected on the X06DA (PXIII) beamline of the Swiss Light Source (SLS), Villigen, Switzerland. All data sets were integrated, merged and scaled using *HKL*-2000 (Otwinowski & Minor, 1997). The structures were solved by molecular replacement with *Phaser* (McCoy *et al.*, 2007) using the coordinates of the native structure of the same protein as an initial search model (PDB entry 1guv; Fusetti *et al.*, 2002). The model was improved by alternating cycles of manual model building using *Coot* (Emsley *et al.*, 2010) with refinement using *REFMAC*5 (Murshudov *et al.*, 2011) and using *PHENIX* (Adams *et al.*, 2010). The stereochemical quality of the final model was assessed with *MolProbity* (Chen *et al.*, 2010). To determine the protonation states of some of the polar residues in the active site, an additional refinement with removed stereochemical restraints for the C—O bond lengths was performed using *SHELXL* full-matrix least-squares refinement (Sheldrick, 2008). The protonation states were determined by measuring and analyzing the C—O bond lengths. Generally, neutral carboxyls have unequal lengths of around 1.21 and 1.32 Å for the C═O and C—OH bonds, respectively. In contrast, negatively charged carboxyls are expected to have identical C—O bond lengths of around 1.26 Å (Erskine *et al.*, 2003; Howard *et al.*, 2004). Structural figures were prepared using *PyMOL* (http://www.pymol.org). A summary of the data-collection and structure-refinement statistics is given in Table 1.

2.4. Quantum mechanics/molecular mechanics (QM/MM) calculations

For combined QM/X-ray refinement we used the program *QMRx* based on the general-purpose *DYNGA* driver (Parker *et al.*, 2003). This program is similar to other available codes capable of hybrid QM/X-ray refinement, such as *ChemShell* (Metz *et al.*, 2014), *ComQum* (Ryde *et al.*, 2002) and *DivCon* in *PHENIX* (Borbulevych *et al.*, 2014). By using *QMRx*, we performed the QM calculations *via* the PM7 Hamiltonian (Borbulevych *et al.*, 2014; Stewart, 2009). The accuracy of PM7 has been discussed in the literature in great detail (Stewart, 2013). We use the *MOZYME* (Stewart, 2009) implementation of PM7 as available in *MOPAC*2012. In our calculations, we address the charge assignments in two significant ways: by drastically increasing the QM integral cutoffs and by introducing a dielectric screening function as recommended in the literature (Andersson *et al.*, 2013). The effect of these additions is a more accurate description of the system but at the cost of a large increase in the total CPU time, limiting the total number of conformations that can be explored in a reasonable amount of time. To compensate for this high computational cost the QM integral cutoff was varied, following *MOPAC* software recommendations, from CUTOFF = 6.0 for GNORM = 20 to CUTOFF = 12.0 for GNORM = 1.0 in four steps. For the X-ray section *QMRx* either uses *XPLOR-NIH* (Wick *et al.*, 2014) or *SHELX* as a slave program to obtain the first-order derivatives of the X-ray restraint function. Mixing of the QM and X-ray restraints was performed using the standard protocols as described in the literature (Metz *et al.*, 2014; Ryde *et al.*, 2002; Yu *et al.*, 2006; Falklöf *et al.*, 2012). The procedure uses the standard approach for constraint weight evaluation as described in MM/X-ray methods for the evaluation of the restraint (Brünger & Rice, 1997), albeit applied to smaller random regions of the structure owing to the much larger computational cost for the QM calculation. The restraint weight factor in our case was confined to the range 0.2–0.3. The convergence during optimization was monitored by following the norm of the total gradient (ng) until we achieved ng < 1.25. The procedure applied to proton transfer has been described in the literature (see, for example, tutorial A10 at http://ambermd.org/tutorials/) and requires the active coordinates to be expressed in internal coordinates, for which the *BABEL* program was used. 20 trajectories (200 ps each) using QM/MD calculations were performed for every transition (I–IV) using the *DYNGA* program as previously described (Parker *et al.*, 2003) and using *Gaussian*03 as a slave program (Foresman & Frisch, 1996). Residues Tyr27, Asp136, Asp138, Glu140 and Tyr141 and the chitobiose appear as observed in the structures of CHIT1 in complex with chitobiose. Note that the total number of atoms needs to be constant; therefore, this motif is present in all calculations. All of the water molecules in direct contact with these residues and pointing towards the reaction centre were included. Calculations were performed using the PBEPBE1 Hamiltonian as implemented in *Gaussian*03 and the 6-31G* atomic basis sets (Frisch *et al.*, 1998) for all calculations. The step of integration was 0.2 fs with T = 298 K using a Nose thermostat as implemented in *DYNGA*.

3. Results and discussion

In this study, we present a detailed structural and mechanistic analysis of the CHIT1 catalytic domain based on several X-ray crystal structures at atomic resolution. To the best of our knowledge, this is the first report in which a GH18 chitinase member has been studied at such a resolution. This family, which is structurally characterized by a conserved three-dimensional fold consisting of an $(\alpha/\beta)_8$ TIM barrel, does not show a particularly high sequence similarity (the average pairwise identity is 21%; http://www.sanger.ac.uk/Software/Pfam; Synstad *et al.*, 2004). Nonetheless, the catalytic triad with the D*X*D*X*E signature is fully conserved in all active GH18 chitinases (Bussink *et al.*, 2007). Accordingly, CHIT1 is a reliable model to understand the detailed structural basis of the catalytic mechanism of GH18 chitinases. Our work has focused on studying the geometry changes within the catalytic triad of CHIT1 during different stages of enzyme catalysis.

Files: d/lp5005/lp5005.3d d/lp5005/lp5005.sgml **LP5005** FA IU-1512/21(26)5 1512/20(26)5 () LP5005 PROOFS D:FA:2015:71:5:0:0–0

Additionally, extensive analyses were applied to combine the observed geometric shifts with the repercussive protonation-state modifications within the side chains of the main residues in the catalytic site. Consequently, correlations between the geometric rearrangements and the proton translocations have been established.

3.1. Atomic resolution structures of the catalytic domain of CHIT1 reveal a double conformation of key catalytic residues

The previously reported crystal structures of CHIT1 in the apo form (PDB entry 1guv) or in complex with chitobiose (PDB entry 1lg1) determined by Fusetti *et al.* (2002) showed the catalytic domain at 2.35 and 2.78 Å resolution, respectively. Even in the presence of an inhibitor, the highest resolution obtained was 1.65 Å (in complex with the inhibitor argifin; PDB entry 1waw; Rao *et al.*, 2005). Since our aim was to perform detailed structural mechanistic studies and protonation-state analysis, we needed to improve the resolution of these structures to subatomic resolution. We therefore set up new crystallization conditions combined with micro-seeding, which allowed us to obtain crystals of CHIT1 with higher diffraction quality extending to approximately 1.0 Å resolution (§2.3).

As expected, the final models of both the apo and pseudo-apo forms of CHIT1 comprised 365 amino acids with 2–3 additional residues observed belonging to the thrombin site. The His tag did not appear and is presumed to be disordered. The overall tertiary structure, determined at 0.95 Å resolution for the pseudo-apo form and at 1.0 Å resolution for the apo form, is a TIM barrel and resembles that previously determined at 2.35 Å resolution (Fusetti *et al.*, 2002). Increasing the resolution of the apo form of CHIT1 allowed us to gain further insight into the structural features of the catalytic groove and of the active site, in particular concerning its flexibility. Interestingly, at this resolution several residues in the active site are detected in multiple conformations; in particular, the residues of the catalytic triad (D_1XD_2XE) adopt double conformations (Figs. 1*a* and 2*a*). In contrast, at 2.35 Å resolution (PDB entry 1guv) the multiple conformations of many of these residues were unresolved.

In the 1.0 Å resolution apo structure the catalytic glutamate shows a double conformation: Glu140-conf*A*, a slightly rotated nonplanar rotamer (16% occupancy), and Glu140-conf*B*, a major planar rotamer (84% occupancy) (Figs. 1*a* and 1*b*). In addition, the middle aspartate (Asp138) adopts two conformations (Figs. 1*a*, 1*c*, 1*d* and 1*e*). Asp138-conf*A* is oriented towards Asp136 (45% occupancy) and forms a short hydrogen bond to it (2.55 Å), suggesting the presence of a low-barrier hydrogen bond (LBHB; Figs. 1*d* and 1*e*). Asp138-conf*B* (55% occupancy) is flipped

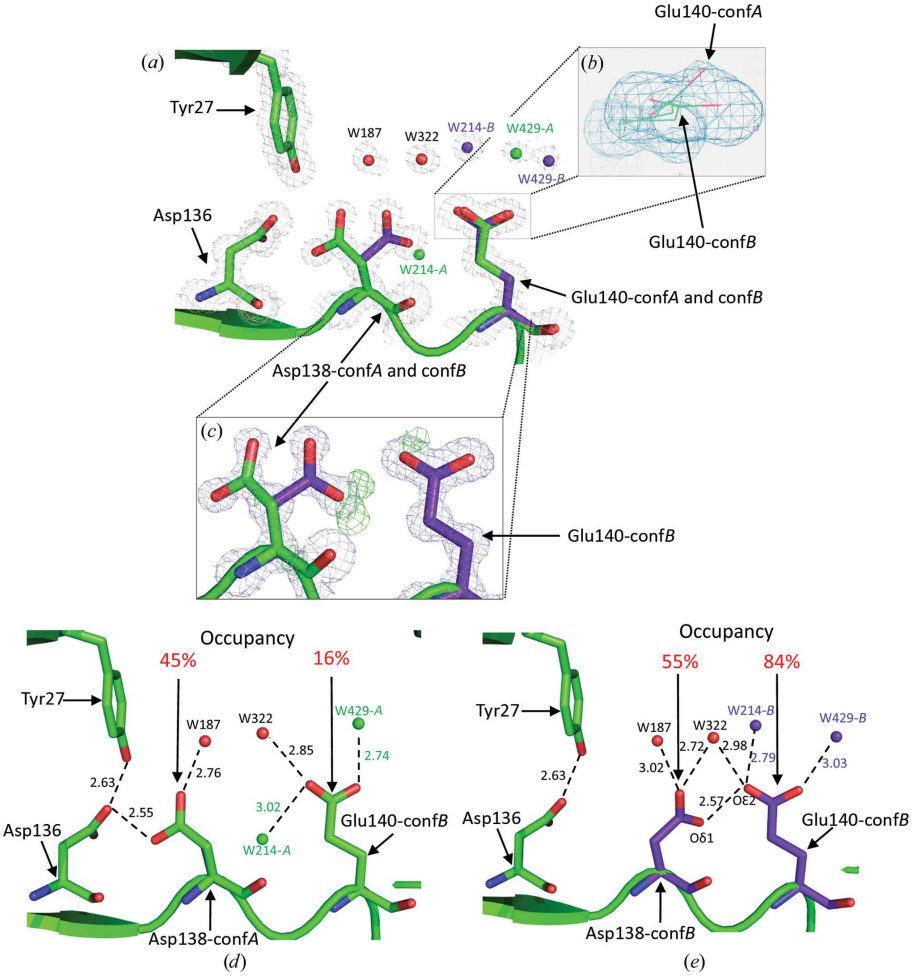


Figure 1
CHIT1 active site in the apo form. (*a*) Cluster of Asp136, Asp138, Glu140 and Tyr27 shown as sticks and water molecules W187, W322, W214-*A* and W214-*B*, and W429-*A* and W429-*B* shown as spheres. Tyr27, Asp136, Asp138-conf*A*, Glu140-conf*A*, W214-*A* and W429-*A* are coloured green. Asp138-conf*B*, Glu140-conf*B*, W214-*B* and W429-*B* are coloured blue/purple. The $2mF_o - F_c$ electron-density map (1σ cutoff) of the cluster is shown as a mesh and coloured grey. (*b*) Two conformations (rotated Glu140-conf*A* and planar Glu140-conf*B*) shown as lines within a $2mF_o - F_c$ electron-density map (1σ cutoff). (*c*) $2mF_o - F_c$ electron-density map (grey; 1σ cutoff) of Asp138 and Glu140-conf*B* and $mF_o - F_c$ electron-density map of Asp138 and Glu140 (green; 3σ cutoff). (*d*, *e*) Stereoview of the hydrogen-bonding network, percentage occupancies and distances in each conformation of the cluster shown in (*a*). Ser181 is also near the catalytic triad, but is not shown for the sake of clarity.

Files: d/lp5005/lp5005.3d d/lp5005/lp5005.sgml **LP5005** FA IU-1512/21(26)5 1512/20(26)5 () LP5005 PROOFS D:FA:2015:71:5:0:0–0

towards Glu140, to which it forms a short hydrogen bond (2.57 Å, again suggesting an LBHB), stabilizing its conformation (Fig. 1*e*). Thus, Asp138-conf*A* is linked to Asp136, while Asp138-conf*B* is linked to Glu140-conf*B* (Figs. 1*d* and 1*e*). Moreover, the outer O atom of Glu140-conf*A* forms a hydrogen bond to conformation *A* of the water molecule W429 (W429-*A*, 49% occupancy). A second water molecule with 51% occupancy appears 1.57 Å from W429-*A*. Since it is not possible to have two water molecules at such a short distance, we can hypothesize that W429-*A* is displaced by 1.57 Å and thereby most likely adopts a second conformation (W429-*B*, 51% occupancy; Figs. 1*d* and 1*e*). This slight positional shift of water molecule 429 between conformations *A* and *B* is consistent with the presumed movement of residue Glu140, with which it is in contact.

During our analysis, we compared the overall occupancy distributions between the residues of the catalytic triad and the surrounding water molecules. Indeed, Glu140 seemed to be practically unaffected by the conformational changes of Asp138. This can be explained by the presence of a second water molecule which occupies the same position as Asp138-conf*B* (W214-*A*, 40% occupancy; Fig. 1*c*). Interestingly, W214-*B* (60% occupancy) is interacting with Glu140-conf*B*, which will occur only when Asp138-conf*B* is interacting with Glu140-conf*B*. On the other hand, when Asp138, *i.e.* Asp138-conf*A*, is flipped towards Asp136, W214-*B* might move from this frontal position to the location occupied by the inner O atom of Asp138-conf*B* (Figs. 1*d* and 1*e*).

It is worth noting that Asp138-conf*B* was not detected in the previously published 2.35 Å resolution apo structure and the corresponding lower resolution density was modelled as a water molecule. It was then considered that the catalytic triad adopted only one stable state (corresponding to Asp138-conf*A*, Glu140-conf*A* and W214-*A* in our 1.0 Å resolution apo structure). In contrast, our data suggest the existence of a dynamic equilibrium of the catalytic triad in the apo form. Note also that the conformation of Asp136 remains unchanged because it is stabilized by hydrogen bonds to Tyr27 and Ser181 (Figs. 1*d* and 1*e*). When Asp138 turns towards Asp136, W214 moves down and mimics the position of the inner O atom of Asp138, forming a hydrogen bond to the inner O atom of the Glu140 side chain. Simultaneously, W429 stabilizes the outer O atom of Glu140 (Fig. 1*d*). Simply put, the position and the reorganization of water molecules occurring in coordination with Asp138 flipping limits the flexibility of Glu140, which explains the dominant planar conformation.

In most of the apo structures of GH18 members only a single conformation of the middle aspartate (Asp_2) is observed. This is mainly owing to limitations owing to resolutions of lower than 2 Å and to the use of molecules such as glycerol or ethylene glycol in the cryo-solution, which are capable of interacting with the catalytic glutamate (Yang *et al.*, 2010; Fusetti *et al.*, 2002). Interestingly, an Asp_2 double conformation has been described before in other apo structures of GH18 members, such as ChiA from *Serratia marcescens* (Papanikolau *et al.*, 2001) and chitinase D from *S. proteamaculans* (Madhuprakash *et al.*, 2013), with resolutions of 1.55 and 1.49 Å, respectively. Besides, mutation of Asp_2 has been shown to abolish the enzymatic activity (Papanikolau *et al.*, 2001; Bokma *et al.*, 2002). Indeed, the X-ray crystal structure of the mutant archaeal chitinase complexed with NAG_4 (PDB entry 3a4x), in which Asp_2 was substituted by an alanine, showed an altered conformation of the catalytic glutamate. This resulted in its displacement by 5 Å away from the scissile O atom of the glycosidic bond (Bokma *et al.*, 2002).

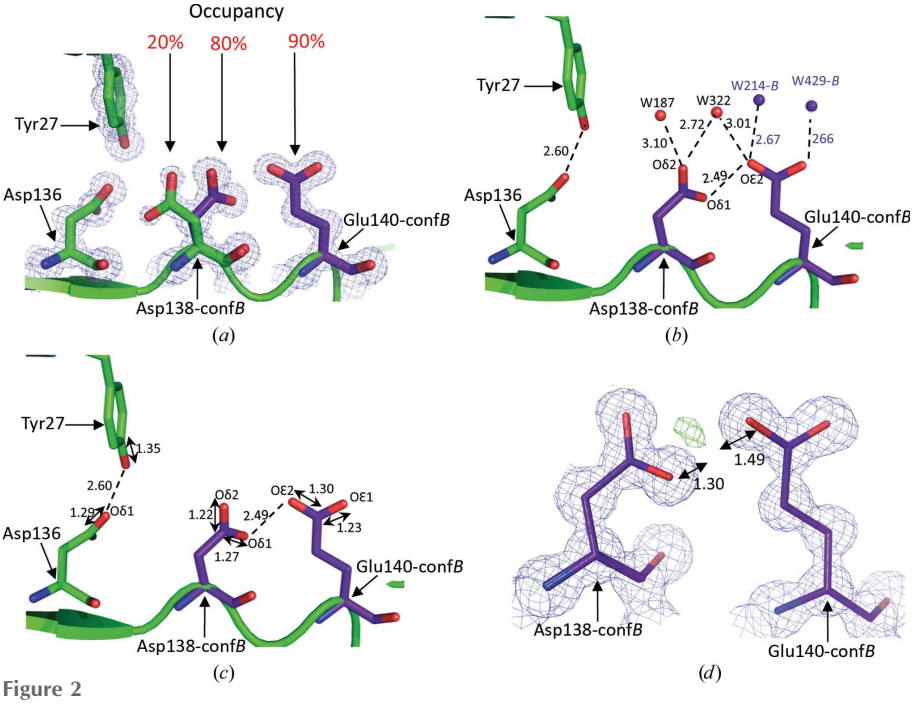


Figure 2
CHIT1 in the pseudo-apo form. (*a*) Cluster of Asp136, Asp138, Glu140 and Tyr27 shown as sticks with percentage occupancies. The $2mF_o - F_c$ electron-density map (1σ cutoff) of the cluster is shown as a mesh and coloured blue. Tyr27, Asp136, Asp138-conf*A* and Glu140-conf*A* are in green. Asp138-conf*B* and Glu140-conf*B* are in blue/purple. (*b*) The hydrogen-bonding network with the distances in conformation *B* of the cluster shown in (*a*). (*c*) Cluster of Tyr27, Asp136, Asp138-conf*B* and Glu140-conf*B* shown as sticks with the C—O bond length obtained by *SHELXL* refinement. (*d*) The $2mF_o - F_c$ electron-density map (1σ cutoff) of Glu140-conf*B* and Asp138-conf*B* in blue and the $mF_o - F_c$ map (3σ cutoff) in green shows a signal (the green blob) that could correspond to a shared H atom between Glu140 and Asp138 in conformation *B*. The two arrows indicate the distances from the inner O atom of Glu140 and the outer O atom of Asp138 to the green blob.

Files: d/lp5005/lp5005.3d d/lp5005/lp5005.sgml **LP5005** FA IU-1512/21(26)5 1512/20(26)5 () LP5005 PROOFS D:FA:2015:71:5:0:0–0

Altogether, this behaviour of Asp_2 seems to be a conserved feature of this residue in GH18 chitinases and reinforces the idea that these flipping conformational changes could be essential to stabilize the catalytic glutamate in the apo form. This geometric stabilization could in fact maintain a favourable conformation for substrate cleavage.

3.2. The protonation states of the catalytic triad residues of the apo form of CHIT1 provide insight into the hydrolysis mechanism

To gain insight into the physicochemical properties of the catalytic triad in CHIT1, we decided to investigate its protonation pattern in the apo form. Usually, ultrahigh-resolution X-ray crystal structures have the potential to reveal exact H-atom positions (Afonine *et al.*, 2010) or alternatively to measure and analyze the C—O bond lengths (Ahmed *et al.*, 2007; Coates *et al.*, 2008; Adachi *et al.*, 2009). The advantage of the latter approach is that it avoids the technical difficulty of locating H atoms in weak electron density. C—O bond lengths can be determined by performing a supplementary refinement with no stereochemical restraints using *SHELXL* (Sheldrick, 2008). This refinement optimizes the carboxyl bond lengths according to the diffraction data. However, in the case of our apo structure at 1.0 Å resolution, this strategy turned out to be insufficient. Indeed, our data correspond to a snapshot which reflects two states of the catalytic triad with Asp138 and Glu140 adopting partial conformations. Therefore, the measured carboxyl bond lengths of Glu140 and Asp136 represent an average of the two states and do not reflect the situation in one of the two conformations. Moreover, the standard average error for the carboxyl bond length of Asp138 in each conformation *A* is relatively high (0.035 Å).

Unexpectedly, when we attempted to co-crystallize CHIT1 with crystalline chitin, we improved the resolution to 0.95 Å but we did not find any electron density corresponding to such a polymer in the structure; therefore, we consider it as a pseudo-apo form. Crystalline chitin is known not to interact with the catalytic domain in the absence of the chitin-binding domain, as is the case here. This structure shows Asp138-conf*B* as the dominant conformation (80% occupancy) as well as Glu140-conf*B* (90% occupancy) (Fig. 2*a*). Moreover, the organization of the water-molecule network in the catalytic groove was closely similar to the 1.0 Å resolution apo structure, with the exception that water molecules W214 and W429 showed only one conformation (for the sake of clarity regarding the residue conformations, we will call them *B*), consistent with a predominant Asp138-conf*B*. We then employed the 0.95 Å resolution pseudo-apo structure as an improved model for analyzing the protonation pattern of the catalytic triad through the C—O bond-length measure.

Following refinement in *phenix.refine* (Afonine *et al.*, 2012), the 0.95 Å resolution pseudo-apo structure was refined by *SHELXL* using full-matrix least-squares refinement (Sheldrick, 2008) with removed stereochemical restraints on the carboxyl moieties of all glutamate and aspartate residues. Our results show that for the Asp136 inner O atom the C—$O^{\delta1}$ bond was refined to 1.26 ± 0.02 Å, while for the outer O atom $O^{\delta2}$ the C—$O^{\delta2}$ bond was refined to 1.29 ± 0.02 Å, suggesting partial protonation of this $O^{\delta2}$. In parallel, the phenol O atom of Tyr27 reveals a C—O bond of 1.35 ± 0.02 Å, suggesting also partial to total protonation of the hydroxyl of Tyr27 (Figs. 2*b* and 2*c*). Taken together and since Tyr27 and Asp136 make a short hydrogen bond (2.60 Å), these data suggest that these two residues share a proton within an LBHB (Figs. 2*b* and 2*c*). To further validate our results, QM/MM was performed with the 0.95 Å pseudo-apo structure. The aforementioned interpretation was also supported and expanded by the QM/MM calculations, which showed a low barrier energy for proton sharing between these two residues (see Fig. 3 and Supplementary Movie S1), where the intermediate point *b* (Fig. 3) is

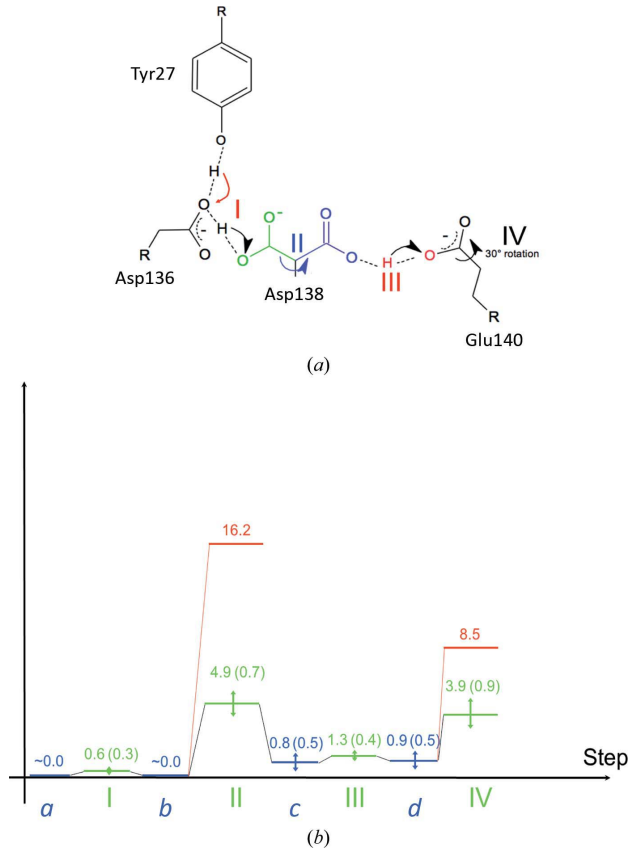


Figure 3
Proposed catalytic triad torsion and proton-exchange pathway. (*a*) I–IV show the proposed steps of the reference mechanism for the QM/MD barriers calculation. (*b*) Proposed energy profile of proton exchange and catalytic triad torsion. Error bars represent a range including 90% of the trajectories. The intermediate point *b* is very close in total energy to the initial structure and an energy gap could not be properly evaluated (there was at least one point near the minima with similar energy to the starting point *a*). The barriers for the rotation of Asp138 and Glu140 from a single 500 ps trajectory are shown in red. The barriers are much higher than in any of the other processes explored. Note that owing to the limitations of the steered dynamics approach, the barriers computed should be considered to be an upper limit and not an accurate value for the process. The QM model explored suggests that the process-limiting step is the rotation of the carboxylate groups (Asp138 and Glu140) and not the actual proton-transport steps.

Files: d/lp5005/lp5005.3d d/lp5005/lp5005.sgml **LP5005** FA IU-1512/21(26)5 1512/20(26)5 () LP5005 PROOFS D:FA:2015:71:5:0:0–0

very close in total energy to the initial structure and an energy gap could not be properly evaluated (there was at least one point near the minima with a similar energy to the starting point *a*).

In the case of Asp138-conf*B* the C—$O^{\delta2}$ bond displayed a bond length of 1.22 ± 0.02 Å, indicating that it is deprotonated, while the C—$O^{\delta1}$ bond of Asp138-conf*B* was refined to 1.27 ± 0.02 Å, indicating that it is partially protonated (Figs. 2*b* and 2*c*). In parallel, the C—$O^{\varepsilon1}$ bond of the dominant conformation Glu140-conf*B* was refined to 1.23 ± 0.02 Å, while the C—$O^{\varepsilon2}$ bond showed a bond length of 1.30 ± 0.02 Å. The carboxylate bond lengths between Glu140 and Asp138 are consistent with the presence of a short LBHB of 2.49 Å between these two residues and underline the surprising finding that the outer O atom is not protonated in the apo form when Asp138 flips towards Glu140. In fact, the values obtained led us to conclude that the inner O atom of Asp138-conf*B* ($O^{\delta1}$) shares a proton with the inner O atom of Glu140 ($O^{\varepsilon2}$), where the affinities of these two residues for the proton are closely similar. In this regard, a round electron-density signal appears between Asp138-conf*B* and Glu140-conf*B*, which may correspond to this shared proton (Fig. 2*d*).

Furthermore, the QM/MM calculations confirm the X-ray observations and also indicate that a proton is shared between $O^{\delta1}$ of Asp138-conf*B* and $O^{\varepsilon2}$ of Glu140-conf*B* in the apo form of CHIT1 (see Figs. 3*a* and 3*b* and Supplementary Movie S1).

To investigate the protonation state of the catalytic triad when Asp138-conf*B* flips to Asp138-conf*A*, we based our study on data from QM/MM calculations. This is because in the 0.95 Å resolution pseudo-apo structure Asp138-conf*A* has an occupancy of 20%, which is not sufficient to make an accurate bond-length analysis as the standard errors are high (see Table 2). Interestingly, QM/MM calculations revealed that when Asp138-conf*A* forms an LBHB to Asp136, there is a proton-sharing phenomenon between Tyr27, Asp136 and Asp138. Moreover, in this conformation the outer O atom of Glu140 $O^{\varepsilon1}$ stays deprotonated and the inner O atom $O^{\varepsilon2}$ forms a hydrogen bond to water molecule W214-*A* as shown in Fig. 1(*d*) for the apo structure at 1.0 Å resolution. Combining the X-ray and QM/MM data from both structures (apo and pseudo-apo), we propose that CHIT1 possesses an unusual system to 'stock' the proton before hydrolysis. This system involves at least four residues (Tyr27, Asp136, Asp138 and Glu140), in which Asp138, by flipping constantly, swings the proton to each side of the catalytic site from Asp136 to Glu140. Remarkably, our finding reveals that in contrast to the previously reported data (van Aalten *et al.*, 2001; Papanikolau *et al.*, 2001; Fusetti *et al.*, 2002; Jitonnom *et al.*, 2011, 2014), CHIT1 maintains the outer O atom $O^{\varepsilon1}$ of Glu140, which is supposed to donate the proton to cleave the glycosidic bond, deprotonated. Taken together, our data questions the previously published model of the hydrolysis mechanism (van Aalten *et al.*, 2001) as well as the proposed transglycosylation model (Zakariassen *et al.*, 2011).

3.3. Atomic resolution structures of the catalytic domain of CHIT1 in complex with chitobiose provide insight into the catalytic mechanism

Studies of the complex of CHIT1 with a substrate were conducted to determine how the catalytic residue Glu140 could protonate the O atom of the scissile glycosidic bond. The limiting step for comparison with the apo CHIT1 structure was again to obtain a resolution of 1.0 Å or better. As the soaking experiments were destabilizing the crystals, resulting in the loss of high-resolution diffraction, we conducted co-crystallization experiments of CHIT1 with different concentrations of the synthetic substrate 4-MU-NAG_3. Fortunately, we succeeded in developing a robust protocol to co-crystallize CHIT1 with 4-MU-NAG_3 by means of micro-seeding. This methodology allowed us to control crystal growth and obtain high-quality crystals that diffracted to atomic resolution. Effectively, co-crystals with 4-MU-NAG_3 at concentrations of 0.3, 1 and 2.5 m*M* reached X-ray data resolutions of 1.10, 1.05 and 1.10 Å, respectively (Table 1 and Fig. 4).

However, all of the structures (solved by molecular replacement using the apo structure as an initial model) revealed CHIT1 complexed to a dimer of *N*-acetylglucosamine (chitobiose) located in the −1 and −2 subsites. This indicates that hydrolysis occurred in the drop and thus allows us to analyze the post-hydrolysis three-dimensional structure of CHIT1. All of our $2mF_o - F_c$ electron-density maps of the three structures show the −2 NAG in a chair conformation, while the −1 NAG adopts a boat conformation (Fig. 4*b*), which disagrees with the previously published structure of CHIT1–chitobiose at 2.78 Å resolution in which the −1 NAG was modelled in a chair conformation (PDB entry 1lg1; Fusetti *et al.*, 2002). Most probably, the low resolution of this structure impeded clear determination of the −1 NAG configuration and could thus explain this disagreement. Moreover, the boat conformation of the −1 NAG observed in our structures is consistent with the substrate-distortion event described in other GH18 chitinases and reported to be required for the substrate-assisted mechanism in this enzyme family (Brameld & Goddard, 1998; Songsiriritthigul *et al.*, 2008; van Aalten *et al.*, 2001).

Interestingly, on comparing the three structures, our data indicate a gradual increase in chitobiose occupancies in the catalytic groove consistent with the augmentation of substrate in the drop. As a result, the occupancy of chitobiose was refined to 50, 69 and 80% for the structures obtained at 0.3, 1 and 2.5 m*M* substrate concentrations, respectively (Fig. 4*b*). Remarkably, in the condition with the lowest substrate concentration (0.3 m*M*) and occupancy (50%) we noticed that the occupancy of Glu140-conf*A*, which is minimal (16%) in the apo form, significantly increases to 41%, becoming closer to the occupancy of the Glu140 planar conformation (Glu140-conf*B*) (Figs. 4*a* and 4*b*). In the same condition, Asp138 also shows a quasi-equal occupancy of conformations *A* (55%) and *B* (45%) (Fig. 4*c*). Notably, by supplementing the substrate amount in the drop the occupancy of chitobiose in the binding site gradually increases (69% in the structure with 1 m*M* substrate concentration and 80% in the structure with 2.5 m*M*

Files: d/lp5005/lp5005.3d d/lp5005/lp5005.sgml **LP5005** FA IU-1512/21(26)5 1512/20(26)5 () LP5005 PROOFS D:FA:2015:71:5:0:0–0

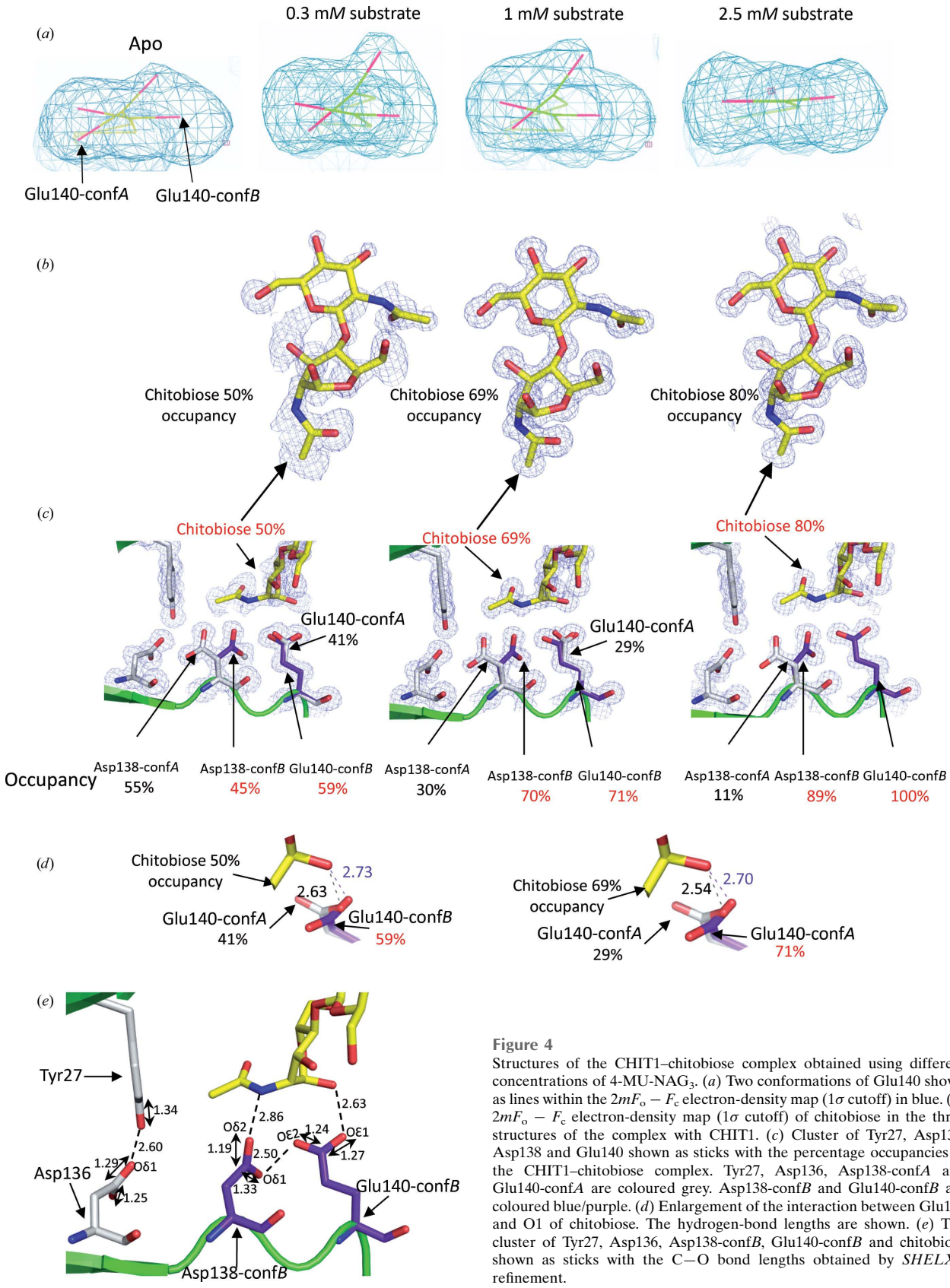


Figure 4
Structures of the CHIT1–chitobiose complex obtained using different concentrations of 4-MU-NAG_3. (*a*) Two conformations of Glu140 shown as lines within the $2mF_o - F_c$ electron-density map (1σ cutoff) in blue. (*b*) $2mF_o - F_c$ electron-density map (1σ cutoff) of chitobiose in the three structures of the complex with CHIT1. (*c*) Cluster of Tyr27, Asp136, Asp138 and Glu140 shown as sticks with the percentage occupancies in the CHIT1–chitobiose complex. Tyr27, Asp136, Asp138-conf*A* and Glu140-conf*A* are coloured grey. Asp138-conf*B* and Glu140-conf*B* are coloured blue/purple. (*d*) Enlargement of the interaction between Glu140 and O1 of chitobiose. The hydrogen-bond lengths are shown. (*e*) The cluster of Tyr27, Asp136, Asp138-conf*B*, Glu140-conf*B* and chitobiose shown as sticks with the C—O bond lengths obtained by *SHELXL* refinement.

Files: d/lp5005/lp5005.3d d/lp5005/lp5005.sgml **LP5005** FA IU-1512/21(26)5 1512/20(26)5 () LP5005 PROOFS D:FA:2015:71:5:0:0–0

substrate concentration). This augmentation was also accompanied by higher occupancies of Glu140-conf*B* (71 and 100% in the structures with 1 and 2.5 m*M* substrate concentration, respectively) and Asp138-conf*B* (70 and 89% in the structures with 1 and 2.5 m*M* substrate concentration, respectively), as well as a decrease in the slightly rotated nonplanar conformation of Glu140 (Glu140-conf*A*; 29 and 0% in the structures with 1 and 2.5 m*M* substrate concentration, respectively) (Fig. 4*c*). On one hand, this confirms the previously reported data indicating that the presence of the substrate induces the rotation of Asp138 towards Glu140 (van Aalten *et al.*, 2000). On the other hand, our observations reveal for the first time two clear conformations of Glu140 in the presence of a hydrolyzable substrate, one planar and one rotated, indicating that such movement occurs (Fig. 4*a*). This observation prompted us to believe that we have detected at least two states of the enzyme. In the state where the planar conformation is adopted by Glu140 (Glu140-conf*B*), this conformation is stabilized by hydrogen-bond contacts provided *via* the chitobiose on the outer side ($O^{\varepsilon1}$) and Asp138-conf*B* on the inner side ($O^{\varepsilon2}$) (Fig. 4*e*). In contrast, it seems that in conformation *A* of our CHIT1–chitobiose structures, in which there is no stable interaction with the catalytic triad, Glu140 displays a rotated conformation (Glu140-conf*A*; Fig. 4*a*). This idea is reinforced by the fact that when the occupancy of chitobiose increases the occupancy of the rotated conformation of Glu140 decreases until it becomes negligible in the co-crystal grown at 2.5 m*M* substrate.

Regarding substrate recognition and binding, Songsiriritthigul *et al.* (2008) have reported that the chito-oligosaccharide chain is in a linear form during the initial step of substrate recognition. In the next step, the substrate chain performs a bending step leading to distortion of the −1 NAG to a boat conformation (Songsiriritthigul *et al.*, 2008). According to Songsiriritthigul and coworkers, the bending is accompanied by a twist of the glycosidic bond to make it accessible to cleavage by the catalytic glutamate. Consistent with their report, in our structures with 0.3 and 1 m*M* substrate the rotated Glu140-conf*A* establishes a distance of 2.63 and 2.54 Å, respectively, to the chitobiose (Glu140 $O^{\varepsilon2}$—chitobiose O1B). This distance is increased to 2.73 Å (in the structure at 0.3 m*M* substrate) or 2.70 Å (in the structure at 1 m*M* substrate) when Glu140 adopts the planar Glu140-conf*B* conformation, indicating that the rotated conformation (Glu140-conf*A*) favours cleavage of the substrate (Fig. 4*d*). In this regard, a similar rotation of Glu140 has previously been detected in CHIT1 and AMCase complexes with allosamidin derivatives mimicking the intermediate (at a lower resolution than the previously mentioned CHIT1–chitobiose complex), meaning that Glu140 adopts this rotated position in the transition state (Fusetti *et al.*, 2003; Olland *et al.*, 2009). Altogether, our results suggest that upon substrate bending and twisting the catalytic glutamate also rotates to gain access to the glycosidic bond. Such a rotation would not be possible if Asp138 is in conformation *B* as it stabilizes the inner-side O atom $O^{\varepsilon2}$ of Glu140. Based on this analysis, we propose that the arrival and distortion of the −1 NAG displaces the water molecules which interact with Glu140 and Asp138 in the apo form. Thus, Glu140-conf*A* should rotate when Asp138 adopts conformation *A*. The displacement of the water network leads to the loss of the dynamic equilibrium described in §§3.1 and 3.2 (Figs. 1*d*, 1*e* and 2*b*) which was limiting the mobility of Glu140. Hence, Glu140 is 'free' to rotate when Asp138 turns towards Asp136.

Overall, one can say that upon the arrival of substrate the 'apo' dynamic equilibrium is destabilized, allowing Glu140 to rotate so that the scissile O atom of the glycosidic bond in an accessible position.

3.4. Structural analysis of the catalytic triad residues in the CHIT1–chitobiose structure reveals the coexistence of two enzymatic states in the same crystal form

Having highlighted the importance of the rotation of Glu140 in the process of hydrolysis, we next wanted to investigate its role in the substrate complex. We therefore studied the protonation state of the catalytic triad based on analysis of the carboxyl bond-length distances and QM/MM in CHIT1 complexed with chitobiose.

As previously mentioned, the chitobiose present in the CHIT1–chitobiose complexes obtained with 0.3 and 1 m*M* substrate shows a lower occupancy than the structure obtained with 2.5 m*M* substrate. The latter has a single conformation of Glu140, while in the other two structures Glu140 and Asp138 display double conformations, leading to a decrease in the electron-density peak for each conformation, and thereby increasing the average error for the conformation with lower occupancy. As a result, we could not determine the protonation pattern of Glu140-conf*A* since this conformation did not reach more than 41% occupancy in all of the solved structures with chitobiose. This occupancy is not sufficient to obtain a low standard error using *SHELXL* refinement. Consequently, to overcome this problem we performed QM/MM calculations using the CHIT1–chitobiose structure obtained with 1 m*M* substrate, because this structure contains two conformations of Asp138 and Glu140. Therefore, this structure in combination with the QM/MM calculations allowed the determination of the charge of the catalytic triad when the conformation of Glu140 is rotated (Glu140-conf*A*) and planar (Glu140-conf*B*). Importantly, our QM/MM calculations have shown that the rotated Glu140-conf*B* is protonated while Asp138-conf*A* is deprotonated. This indicates that Asp138-conf*A* transfers a proton to Glu140 before it flips towards Asp136 when deprotonated (see Supplementary Movie S1). Once Glu140 has been protonated, it rotates to gain access to the O atom of the scissile bond. In addition to the QM/MM analysis performed with the structure obtained using a 1 m*M* concentration of substrate, we performed an unrestrained refinement with *SHELXL* on the structure obtained using a 2.5 m*M* concentration, as it represents the most accurate model of the post-hydrolysis state among our three structures (100% planar conformation of Glu140 and Asp138-conf*B* as a dominant conformation at 89%; Table 2).

Files: d/lp5005/lp5005.3d d/lp5005/lp5005.sgml **LP5005** FA IU-1512/21(26)5 1512/20(26)5 () LP5005 PROOFS D:FA:2015:71:5:0:0–0

Table 2
C—O bond length, standard error and pK_a for the structures of CHIT1 in the pseudo-apo form and in complex with chitobiose obtained using 2.5 m*M* substrate.

Residue	Bond	Bond length (Å)	Error (Å)	Bond	Bond length (Å)	Error (Å)	pK_a Conformation *A*	pK_a Conformation *B*
Pseudo-apo structure (0.95 Å resolution)								
Glu140-conf*B*	C—O$^{\varepsilon1}$	1.2311	0.0200	C—O$^{\varepsilon2}$	1.3075	0.0192	8.13	6.55
Asp138-conf*B*	C—O$^{\delta1}$	1.2180	0.0175	C—O$^{\delta2}$	1.2710	0.0175	13.07	12.34
Asp136	C—O$^{\delta1}$	1.2660	0.0126	C—O$^{\delta2}$	1.2877	0.0125	4.64	6.02
Tyr27	C—OH	1.3420	0.0127	—	—	—	—	—
CHIT1–chitobiose, 2.5 m*M* substrate (1.1 Å resolution)								
Glu140-conf*B*	C—O$^{\varepsilon1}$	1.2736	0.0200	C—O$^{\varepsilon2}$	1.2397	0.0207	13.4	
Asp138-conf*B*	C—O$^{\delta1}$	1.3261	0.0208	C—O$^{\delta2}$	1.1907	0.0179	7.6	
Asp136	C—O$^{\delta1}$	1.2506	0.0167	C—O$^{\delta2}$	1.2878	0.0170	6.12	
Tyr27	C—OH	1.3470	0.0158	—	—	—	—	

We then compared the *SHELXL* results for the protonation states of the CHIT1–chitobiose complex obtained at 2.5 m*M* substrate with the QM/MM calculations performed on this complex obtained at 1 m*M* substrate. Hence, the unrestrained refinement of the CHIT1–chitobiose structure obtained using 2.5 m*M* substrate shows that the planar conformation (equivalent to Glu140-conf*B* in the other structures) has a C—O$^{\varepsilon1}$ bond length of 1.27 Å and a C—O$^{\varepsilon2}$ bond length of 1.24 ± 0.02 Å, indicating that both O atoms share the charge. On the other hand, Asp138-conf*B* reveals a C—O$^{\delta1}$ bond length of 1.33 ± 0.02 Å and a C—O$^{\delta2}$ bond length of 1.19 Å, indicating that O$^{\delta1}$ is protonated (Fig. 4*e*). In this state, the Glu140-conf*B* carboxyl side-chain O$^{\varepsilon1}$ is stabilized by the scissile O atom of the −1 NAG and forms a short hydrogen bond of 2.63 Å to it, while the other carboxyl O atom of Glu140-conf*B*, O$^{\varepsilon2}$, forms a strong hydrogen bond (2.50 Å) to the O$^{\delta1}$ of Asp138-conf*B* (Fig. 4*e*). Even though the interatomic distance between the two O atoms (O$^{\varepsilon2}$ of Glu140-conf*B* and O$^{\delta1}$ of Asp138-conf*B*) is similar between this and the pseudo-apo structure at 0.95 Å resolution, an important difference was revealed when we measured the carboxylate bond length of Glu140-conf*B* (Table 2). Thus, our results show that Glu140-conf*B* in the CHIT1–chitobiose complex becomes charged under the condition of 2.5 m*M* substrate. In fact, the presence of such a short O$^{\varepsilon2}$—O$^{\delta1}$ interatomic distance and an ionic profile for Glu140 together with a protonated O atom in the C—O$^{\delta1}$ bond of Asp138-conf*B* prompt us to suggest that the short hydrogen bond between O$^{\varepsilon2}$ and O$^{\delta2}$ is not an LBHB but a strong ionic hydrogen bond (SIHB; Meot-Ner, 2012). Such a suggestion needs to be confirmed by neutron diffraction or NMR. Indeed, such strong unusual nonstandard short hydrogen bonds have recently been revealed by neutron crystallography in elastase and photoactive yellow protein (Yamaguchi *et al.*, 2009; Tamada *et al.*, 2009). Regarding the 2.63 Å interatomic distance between Glu140 O$^{\varepsilon1}$ and O1 of chitobiose, it is not possible to determine whether it is an LBHB or an SIHB by X-ray crystallography. Nonetheless, we can propose that after a hydrolysis cycle the carboxylate of Glu140 becomes charged and bordered by two short hydrogen bonds on each side of its carboxylate.

To further support our data, we estimated the pK_a of all of the polar residues within CHIT1 in both the apo-form and chitobiose-bound structures using the *PROPKA* server (http://propka.ki.ku.dk/; Table 2). When we compare the estimated pK_a values for the Asp138 and Glu140 residues in the apo and the chitobiose-bound structures we obtain a ΔpK_a of −4.74 for Asp138, while Glu140 shows a ΔpK_a of +6.85. This significant pK_a shift indicates that an inversion of the acid/base profiles of Glu140 and Asp138 occurs in the presence of the substrate. In fact, Glu140, which was acidic in the apo form, becomes basic in the presence of chitobiose, whereas Asp138, which was basic, is converted into an acidic residue. The decrease in the pK_a of Asp138 in the presence of chitobiose is most likely owing to the formation of a hydrogen bond between the *N*-acetyl group of the −1 NAG moiety and the outer O atom of Asp138.

Altogether, the pK_a shift which occurs upon the arrival of the substrate supports our notion that Asp138-conf*B* transfers a proton to Glu140. This is then followed by flipping of Asp138 and rotation of Glu140 to gain access to and protonate the O atom of the glycosidic bond.

3.5. Detailed structural analysis sheds new light onto the hydrolytic step

All of the proposed hydrolysis mechanisms for GH18 chitinases have reported that after the cleavage of the glycosidic bond through its protonation by the catalytic glutamate (Glu140 in CHIT1) an oxazolinium-ion intermediate is generated and a water molecule is activated by the same catalytic glutamate (van Aalten *et al.*, 2001; Songsiriritthigul *et al.*, 2008; Tews *et al.*, 1997; Papanikolau *et al.*, 2001; Jitonnom *et al.*, 2011, 2014). According to these proposed hydrolysis models, during the activation of the hydrolytic water molecule the catalytic glutamate receives an H atom and the –OH group of the water molecule performs a nucleophilic attack on the anomeric C atom, leading to the reformation of the −1 NAG moiety with retention of the initial configuration (van Aalten *et al.*, 2001; Songsiriritthigul *et al.*, 2008). In contrast to the generally accepted mechanism, our results show that after hydrolysis Glu140 is ionic and forms a hydrogen bond to the acquired –OH on the anomeric carbon C1. This leads us to believe that the activation of the hydrolytic water might have not been carried out by Glu140 since Glu140 is not protonated

Files: d/lp5005/lp5005.3d d/lp5005/lp5005.sgml **LP5005** FA IU-1512/21(26)5 1512/20(26)5 () LP5005 PROOFS D:FA:2015:71:5:0:0–0

as a result of the catalysis. Even though the deprotonation of Glu140 could be explained by the fact that the reaction occurred on a short substrate containing three sugar units, the product chitobiose could simply dissociate from the active site and then re-enter it, and Glu140 would simply release its proton to the bulk water in the process. We can therefore still hypothesize that another residue could participate in the activation of the hydrolytic water molecule. One candidate residue for this task is probably Asp213, which is located opposite to Glu140 and possesses an outer O atom forming a short contact with the water molecule (W300) (Figs. 5*a*, 5*b* and 5*c*). The contact of Asp213 with the water molecule (W300) occurs in the apo form and in the presence of chitobiose, where it forms a hydrogen bond to –O6H of the −1 NAG moiety. However, when we superimposed a long-chain NAG polymer in subsites −4 to +2 based on the crystal structure of mutant ChiA, the active site of which is highly similar to that of CHIT1, W300 overlaps with the +1 NAG. This means that

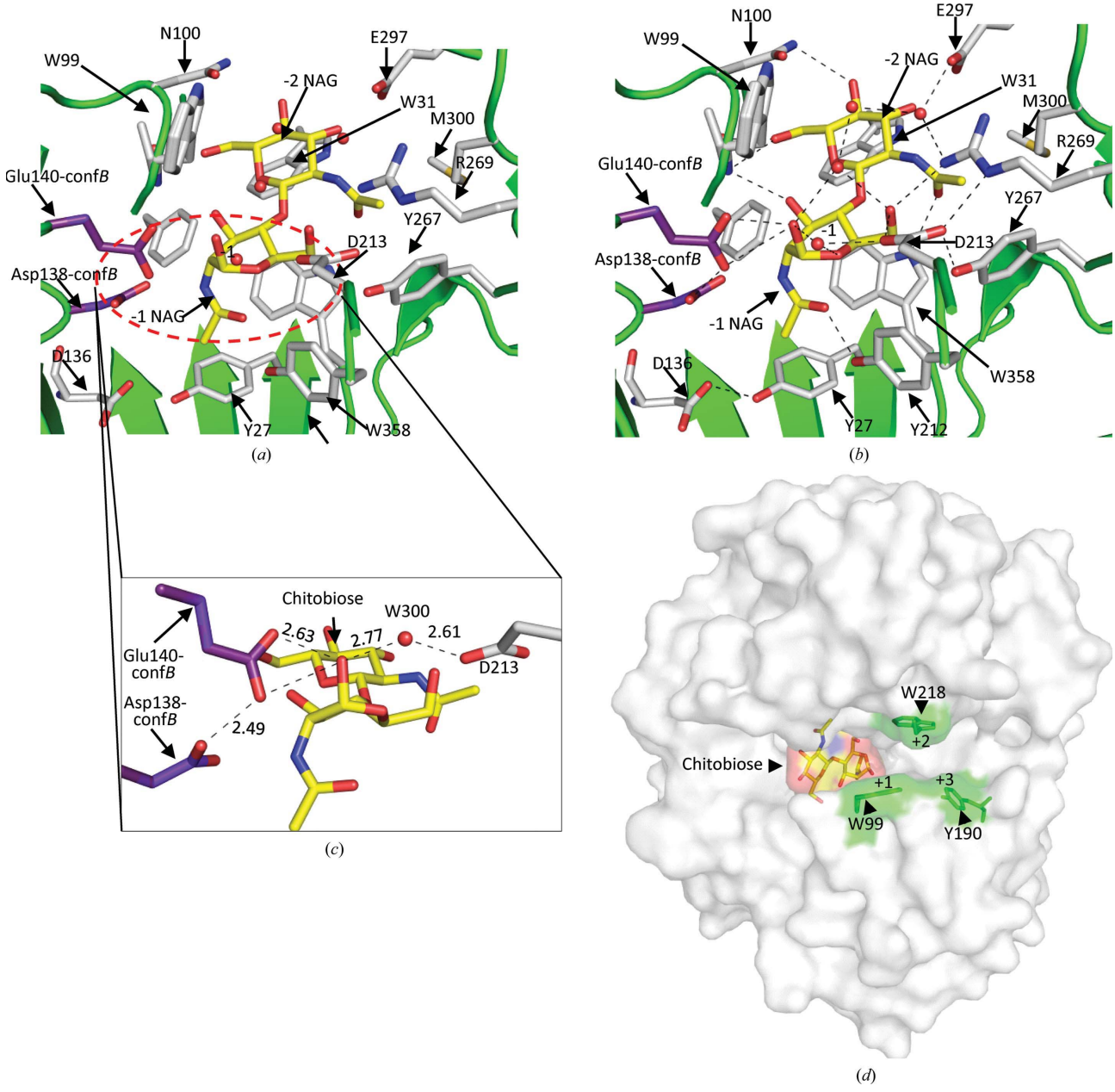


Figure 5
Stereoviews of the interactions of chitobiose in the binding and active sites. (*a*, *b*) Different residues of the binding and active sites involved in the interaction with chitobiose shown as sticks. (*c*) Enlargement of the position of Asp138-conf*B*, Glu140-conf*B*, Asp213 and the water molecule w300 in the presence of chitobiose. Hydrogen-bond distances are indicated. (*d*) CHIT1 and chitobiose in surface representation with the aromatic residues Trp99, Trp218 and Tyr190 in subsites +1, +2 and +3 represented as sticks and coloured green.

Files: d/lp5005/lp5005.3d d/lp5005/lp5005.sgml **LP5005** FA IU-1512/21(26)5 1512/20(26)5 () LP5005 PROOFS D:FA:2015:71:5:0:0–0

upon the sliding of the substrate to the +1 and +2 subsites it displaces this water molecule which was present in the apo form, whereas after the cleavage this water molecule regains its +1 position and is stabilized by Asp213 after the displacement of the aglycon. It is worth noting that Asp213 is highly conserved in GH18 chitinases and its stabilization of the water molecule (W300) also appears in the crystal structure of chitinase D from *S. proteamaculans* at 1.49 Å resolution (Madhuprakash *et al.*, 2013). In ChiB from *S. marcescens*, mutation of Asp215 (Asp213 in CHIT1) to alanine resulted in mild activity of this enzyme and to an acidic shift in its pH optimum (Synstad *et al.*, 2004). Strikingly, several studies have demonstrated that the mutation of the equivalent Asp213 to an alanine is deleterious for the chitinase activity (Synstad *et al.*, 2004; Papanikolau *et al.*, 2001). Taken together, these data suggest that the activation of the water molecule might not be performed by the catalytic glutamate (as it becomes charged after a full hydrolysis cycle) but by another residue, which could be Asp213 in CHIT1.

3.6. A new scenario for the processive hydrolysis

Based on the analysis of the different observed occupancies, hydrogen bonds and C—O bond lengths as well as QM/MM calculations, we have joined it to the scenario of the hydrolysis reaction linked to the enzymatic processivity of CHIT1. The mechanism that we suggest consists of the following.

(i) In the apo form of CHIT1, a dynamic equilibrium is established within the catalytic triad together with Tyr27, allowing the storage of the catalytic proton by a flipping conformational change of Asp138 as well as a back-and-forth movement of the water molecules W214 and W429 (Fig. 6*a*).

(ii) Upon arrival of the substrate, the water molecules in the binding site are displaced and owing to the loss of the dynamic equilibrium as well as the shift in the pK_a, Asp138 transfers a proton to Glu140 and, once deprotonated, turns towards Asp136 (Fig. 6*b*). Simultaneously, Glu140 rotates towards the twisted glycosidic bond and owing to the elevation in its pK_a Glu140 protonates the scissile O atom, leading to the formation of the oxazolinium-ion intermediate (Figs. 6*c* and 6*d*). This is accompanied by the displacement of the aglycon sugar, which allows the hydrolytic water molecule to access the vicinity of the active site (Fig. 6*d*). At this point, after receiving a new proton from Asp136, Asp138 turns towards Glu140, stabilizing its rotated conformation in a planar conformation by an SIHB and establishes a hydrogen bond to the N atom of the *N*-acetyl group in the −1 NAG moiety (Figs. 6*b*, 6*c* and 6*d*).

(iii) A nucleophilic residue (probably Asp213) activates the hydrolytic water molecule, which in turn performs a nucleophilic attack on the anomeric C atom C1 of the intermediate ion, leading to formation of the −1 NAG with retention of its configuration (Fig. 6*e*).

(iv) A hydrolysis cycle is completed at this point and the substrate slides along the enzyme. During the substrate sliding, the −2 NAG arrives at the −1 subsite, resulting in an unproductive binding as the *N*-acetyl group is placed in the opposite direction to the catalytic triad. This allows Asp138 to transfer the proton which was stocked between $O^{\epsilon 2}$ and $O^{\delta 2}$ *via* an SIHB to Glu140. Moreover, as the outer O atom $O^{\delta 2}$ of Asp138 is not stabilized by the *N*-acetyl group, this residue can now flip towards Asp136, acquire a new proton and turn again towards Glu140 to continue a new hydrolysis cycle as described in (ii). The energy landscape of this mechanism has been calculated by a steered dynamics calculation including quantum modelling which proceeded through the steps of the reaction.

This process repeats itself during the hydrolysis of the chitin chain. Consequently, the flipping conformational change of Asp138 in the apo form is important for storage of the proton, where it swings the proton. However, in the presence of a substrate chain it functions as a shuttle of the proton from Tyr27 and Asp136 to Glu140. The repeating cycle in the processive hydrolysis involving many amino acids brings to mind that this enzyme processes the chitin chain according to a 'Fordist model'.

3.7. Detection of product in the catalytic groove provides insight into the transglycosylation mechanism

The detection of chitobiose in the CHIT1 active site is not surprising as it has previously also been detected by soaking crystals of CHIT1 itself (Fusetti *et al.*, 2002) or other native bacterial GH18 chitinases with chito-oligosaccharides (Malecki *et al.*, 2013; Perrakis *et al.*, 1994). The presence of chitobiose in the catalytic groove could be explained by the stacking interactions made by two tryptophans (Trp31, which is found in the −3 and −2 subsites, and Trp358 in the −1 subsite) as well as hydrogen-bond contacts with polar residues and water molecules in the binding site (Figs. 5*b* and 5*c*). Consistently, the in-depth studies conducted by Eide *et al.* (2012) have also shown a high binding affinity for NAG moieties in the −2 and −1 subsites of CHIT1. Moreover, CHIT1 is known to be processive, as are other bacterial GH18 chitinases. Parenthetically, the equivalents of Trp31 and Trp358 in other GH18 chitinases are key residues that are relevant to ensuring the processive capacity of these enzymes. For example, mutation of Trp137 in chitinase A (Trp31 in CHIT1) of *S. marcescens* has been reported to strongly affect processivity (Zakariassen *et al.*, 2009). Therefore, it is believed that the presence of such trytophans in the binding site is important to prevent the chito-oligosaccharide chain from leaving processive GH18 chitinases, thereby allowing the polymer to slide along these enzymes. On the other hand, it is known that chito-oligomer substrates have a successive alternation of the *N*-acetyl group position as each NAG unit is rotated by 180° (Fig. 5*b*) in relation to the next. Thus, the sliding of two NAG units is sufficient to obtain an *N*-acetyl group accommodated in the −1 subsite on the side of the catalytic triad. Once accommodated, this fulfills the condition for the substrate-assisted mechanism to be carried out, thereby leading to cleavage of the glycosidic bond. As a result, the products of processive hydrolysis are disaccharides. Given that CHIT1 cleaves by dimers, chitobiose is the last cleavage unit which cannot be

Files: d/lp5005/lp5005.3d d/lp5005/lp5005.sgml **LP5005** FA IU-1512/21(26)5 1512/20(26)5 () LP5005 PROOFS D:FA:2015:71:5:0:0–0

further cleaved nor slide as it is stabilized by interactions along the dimer. Hence, the fact that we obtained the CHIT1–chitobiose complex not by soaking but by long-duration co-crystallization experiments (four weeks) together with the observation that the chitobiose occupancy increases proportionally to the substrate concentration makes us suggest that the CHIT1–chitobiose complex is a relatively stable complex.

We propose that the high affinity of the −2 and −1 subsites, which causes a relatively high stability of the dimeric product (chitobiose) in the catalytic groove, blocks the −*n* subsite and thus represents the basis for a low-saturation enzymatic capacity. Furthermore, CHIT1 is known to display a high affinity for NAG moieties at the +*n* subsite owing to the presence of aromatic residues (Trp99, Trp218 and Tyr190; Fig. 5*d*). This suggests that a combination of both the obstruction at the −*n* subsites with the high substrate affinity at the +*n* subsites turns these subsites into substrate acceptors. After substrate cleavage and since the catalytic groove is obstructed by the intermediate oxazolinium at the −*n* subsite, either the returning aglycon or another chito-oligosaccharide can be positioned at the +*n* subsite as a new substrate. This lies at the foundation of the re-polymerization phenomenon known as transglycosylation (Taira *et al.*, 2010; Zakariassen *et al.*, 2011) when instead of a water molecule an alcohol attacks the anomeric C atom of the intermediate. Together, these data provide further structural insight into the previously reported high transglycosylation rate of CHIT1 (Aguilera *et al.*, 2003).

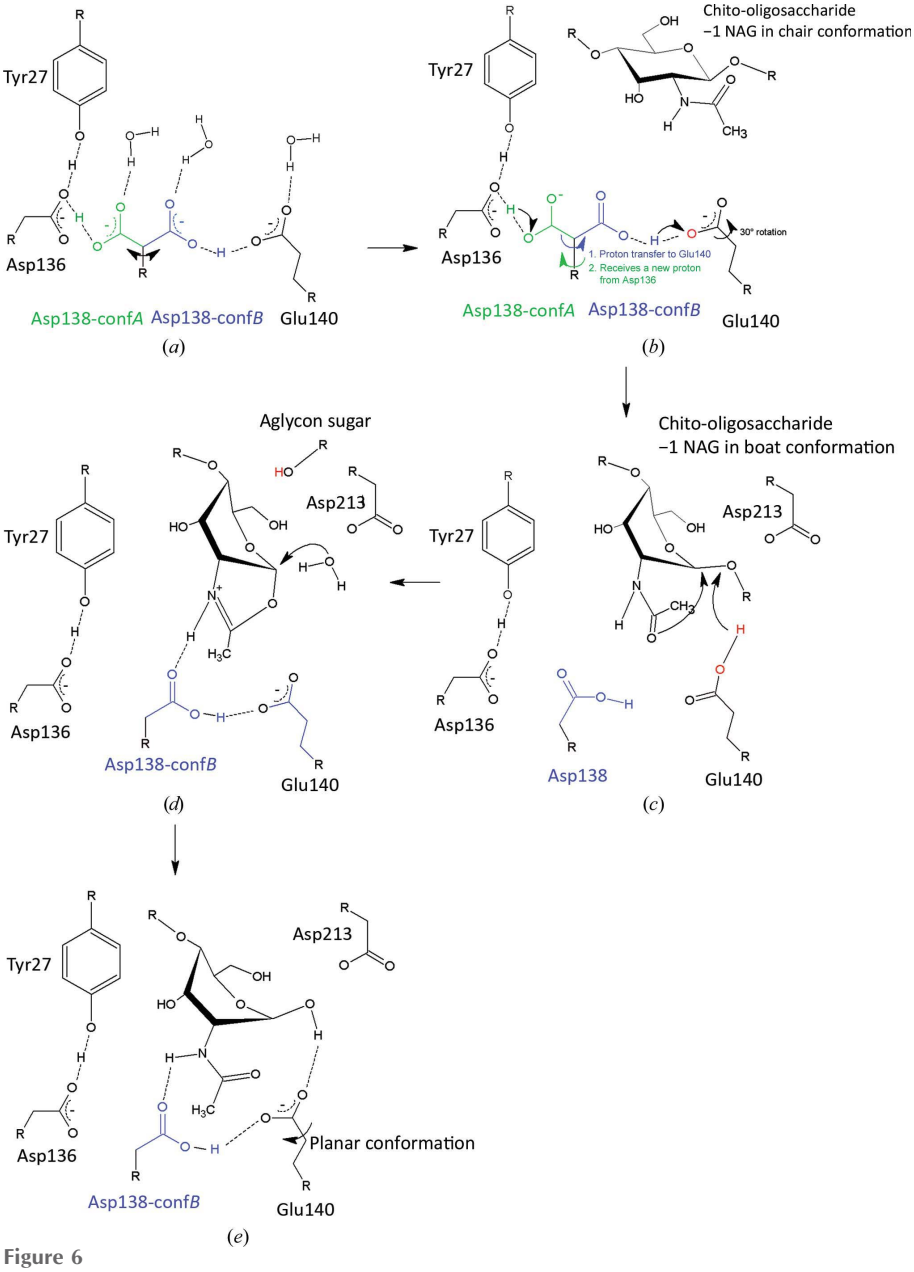


Figure 6
Proposed hydrolysis mechanism of CHIT1 (see text). (*a*) Dynamic equilibrium of the active site of CHIT1 in the apo form. (*b*) Substrate arrival. (1) Proton transfer from Asp138-conf*B* to Glu140. (2) Asp138 flips towards Asp136 to receive a new proton from the latter. (*c*) The −1 NAG twists to a boat conformation and Glu140 protonates the scissile O atom of the glycosidic bond. (*d*) Aglycon sugar displacement, formation of the oxazolinium-ion intermediate and stabilization of protonated Asp138-conf*B*. (*e*) End of the hydrolysis cycle. Regeneration of the −1 NAG with retention of its configuration. SIHB formation between Asp138-conf*B* and Glu140 in the planar form.

4. Conclusion

In this study, and for the first time, we report subatomic resolution structures of the apo form of CHIT1 and its complex with chitobiose by means of X-ray crystallography. We have extended our study to the protonation state of the catalytic residues by the combined use of partially unrestrained refinement with *SHELXL* full-matrix least-squares refinement and QM/MM calculations, which have revealed new insights regarding the catalytic mechanism of the hydrolysis reaction in CHIT1, the main features of which are conserved in the GH18 chitinase family. Indeed, we provide new findings regarding the role of Asp138 as a swing in the apo form and as a

Files: d/lp5005/lp5005.3d d/lp5005/lp5005.sgml **LP5005** FA IU-1512/21(26)5 1512/20(26)5 () LP5005 PROOFS D:FA:2015:71:5:0:0–0

proton shuttle during hydrolysis. Strikingly and in contrast to what was previously assumed, our study of the protonation state of the key catalytic residue Glu140 reveals that the outer O atom of Glu140 is deprotonated in the apo form and adopts an ionic state after hydrolysis. Furthermore, our investigation into the geometry of Glu140 showed for the first time a rotation that liberates Glu140 from Asp138 and therefore allows the protonated O atom to better access the glycosidic bond and to cleave it. Importantly, we indicate a shift in the type of hydrogen bond established between Asp138 and Glu140 from an LBHB in the apo form to an SIHB in the complex with chitobiose, which could be important to maintain the ability to perform many hydrolytic cycles. Moreover, our results underline the low-barrier phenomenon of proton sharing taking place between Tyr27 together with Asp136 and Asp138 in the apo form for proton storage. In addition, our findings highlight the putative role of Tyr27 and Asp136 in 'supplying' protons to Asp138 thanks to a low energy barrier for proton translocation between these three residues during the hydrolysis cycle. Besides providing a deeper understanding of the hydrolytic mechanism, our structures of the CHIT1–chitobiose complex have provided additional insights regarding the structural basis of the high rate of transglycosylation in CHIT1. Finally, this work offers the perspective of conducting joint atomic X-ray plus neutron diffraction studies to obtain further insight into our newly proposed CHIT1 catalytic mechanism. Overall, the data presented here provide new structural knowledge which could serve as a basis for the design of more specific and powerful inhibitors of CHIT1 and GH18 chitinases.

Acknowledgements

We thank the IGBMC Structural Biology and Genomics Platform staff for technical assistance. We are grateful to Dr Irene Yujnovsky and Dr Sergey Melnikov for helpful discussions. The crystallographic experiments were performed on the X06DA (PXIII) beamline at the Swiss Light Source synchrotron, Paul Scherrer Institut, Villigen, Switzerland. We thank, in particular, Ezequiel Panepucci and Vincent Olieric for their help with data collection. This work has been funded by the CNRS, the INSERM, Strasbourg University, BiostructX (FP7, contract 283570), Instruct (European Strategy Forum of Research Infrastructures; ESFRI) and the French Infrastructure for Integrated Structural Biology (FRISBI).

References

Aalten, D. M. F. van, Komander, D., Synstad, B., Gaseidnes, S., Peter, M. G. & Eijsink, V. G. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 8979–8984.
Aalten, D. M. F. van, Synstad, B., Brurberg, M. B., Hough, E., Riise, B. W., Eijsink, V. G. & Wierenga, R. K. (2000). *Proc. Natl Acad. Sci. USA*, **97**, 5842–5847.
Adachi, M. *et al.* (2009). *Proc. Natl Acad. Sci. USA*, **106**, 4641–4646.
Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.
Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H. & Adams, P. D. (2012). *Acta Cryst.* D**68**, 352–367.
Afonine, P. V., Mustyakimov, M., Grosse-Kunstleve, R. W., Moriarty, N. W., Langan, P. & Adams, P. D. (2010). *Acta Cryst.* D**66**, 1153–1163.
Aguilera, B., Ghauharali-van der Vlugt, K., Helmond, M. T., Out, J. M., Donker-Koopman, W. E., Groener, J. E., Boot, R. G., Renkema, G. H., van der Marel, G. A., van Boom, J. H., Overkleeft, H. S. & Aerts, J. M. (2003). *J. Biol. Chem.* **278**, 40911–40916.
Ahmed, H. U., Blakeley, M. P., Cianci, M., Cruickshank, D. W. J., Hubbard, J. A. & Helliwell, J. R. (2007). *Acta Cryst.* D**63**, 906–922.
Andersson, M. P., Jensen, J. H. & Stipp, S. L. (2013). *PeerJ*, **1**, e198.
Aricescu, A. R., Lu, W. & Jones, E. Y. (2006). *Acta Cryst.* D**62**, 1243–1250.
Bokma, E., Rozeboom, H. J., Sibbald, M., Dijkstra, B. W. & Beintema, J. J. (2002). *Eur. J. Biochem.* **269**, 893–901.
Borbulevych, O. Y., Plumley, J. A., Martin, R. I., Merz, K. M. & Westerhoff, L. M. (2014). *Acta Cryst.* D**70**, 1233–1247.
Brameld, K. A. & Goddard, W. A. (1998). *J. Am. Chem. Soc.* **120**, 3571–3580.
Brünger, A. T. & Rice, L. M. (1997). *Methods Enzymol.* **277**, 243–269.
Bussink, A. P., Speijer, D., Aerts, J. M. & Boot, R. G. (2007). *Genetics*, **177**, 959–970.
Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* D**66**, 12–21.
Coates, L., Tuan, H. F., Tomanicek, S., Kovalevsky, A., Mustyakimov, M., Erskine, P. & Cooper, J. (2008). *J. Am. Chem. Soc.* **130**, 7235–7237.
Davies, G. & Henrissat, B. (1995). *Structure*, **3**, 853–859.
Eide, K. B., Norberg, A. L., Heggset, E. B., Lindbom, A. R., Vårum, K. M., Eijsink, V. G. & Sørlie, M. (2012). *Biochemistry*, **51**, 487–495.
Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.
Erskine, P. T., Coates, L., Mall, S., Gill, R. S., Wood, S. P., Myles, D. A. A. & Cooper, J. B. (2003). *Protein Sci.* **12**, 1741–1749.
Falklöf, O., Collyer, C. & Reimers, J. (2012). *Theor. Chem. Acc.* **131**, 1076.
Foresman, J. B. & Frisch, Æ. (1996). *Exploring Chemistry With Electronic Structure Methods: A Guide to Using Gaussian.* Pittsburgh: Gaussian Inc.
Frisch, M. J. *et al.* (1998). *Gaussian* 98. Gaussian Inc., Pittsburgh, Pennsylvania, USA.
Fusetti, F., Pijning, T., Kalk, K. H., Bos, E. & Dijkstra, B. W. (2003). *J. Biol. Chem.* **278**, 37753–37760.
Fusetti, F., von Moeller, H., Houston, D., Rozeboom, H. J., Dijkstra, B. W., Boot, R. G., Aerts, J. M. & van Aalten, D. M. F. (2002). *J. Biol. Chem.* **277**, 25537–25544.
Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M., Appel, R. & Bairoch, A. (2005). *The Proteomics Protocols Handbook*, edited by J. Walker, pp. 571–607. Totowa: Humana Press.
Gooday, G. (1990). *Adv. Microb. Ecol.* **11**, 387–430.
Henrissat, B. & Davies, G. (1997). *Curr. Opin. Struct. Biol.* **7**, 637–644.
Hollak, C. E., van Weely, S., van Oers, M. H. & Aerts, J. M. (1994). *J. Clin. Invest.* **93**, 1288–1292.
Howard, E. I., Sanishvili, R., Cachau, R. E., Mitschler, A., Chevrier, B., Barth, P., Lamour, V., Van Zandt, M., Sibley, E., Bon, C., Moras, D., Schneider, T. R., Joachimiak, A. & Podjarny, A. (2004). *Proteins*, **55**, 792–804.
Jitonnom, J., Lee, V. S., Nimmanpipug, P., Rowlands, H. A. & Mulholland, A. J. (2011). *Biochemistry*, **50**, 4697–4711.
Jitonnom, J., Limb, M. A. & Mulholland, A. J. (2014). *J. Phys. Chem. B*, **118**, 4771–4783.
Laemmli, U. K. (1970). *Nature (London)*, **227**, 680–685.
Madhuprakash, J., Singh, A., Kumar, S., Sinha, M., Kaur, P., Sharma, S., Podile, A. R. & Singh, T. P. (2013). *Int. J. Biochem. Mol. Biol.* **4**, 166–178.
Malecki, P. H., Raczynska, J. E., Vorgias, C. E. & Rypniewski, W. (2013). *Acta Cryst.* D**69**, 821–829.

Files: d/lp5005/lp5005.3d d/lp5005/lp5005.sgml **LP5005** FA IU-1512/21(26)5 1512/20(26)5 () LP5005 PROOFS D:FA:2015:71:5:0:0–0

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
Meot-Ner, M. (2012). *Chem. Rev.* **112**, PR22–PR103.
Metz, S., Kästner, J., Sokol, A. A., Keal, T. W. & Sherwood, P. (2014). *WIREs Comput. Mol. Sci.* **4**, 101–110.
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.* D**67**, 355–367.
Olland, A. M., Strand, J., Presman, E., Czerwinski, R., Joseph-McCarthy, D., Krykbaev, R., Schlingmann, G., Chopra, R., Lin, L., Fleming, M., Kriz, R., Stahl, M., Somers, W., Fitz, L. & Mosyak, L. (2009). *Protein Sci.* **18**, 569–578.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Papanikolau, Y., Prag, G., Tavlas, G., Vorgias, C. E., Oppenheim, A. B. & Petratos, K. (2001). *Biochemistry*, **40**, 11338–11343.
Parker, C. L., Ventura, O. N., Burt, S. K. & Cachau, R. E. (2003). *Mol. Phys.* **101**, 2659–2668.
Perrakis, A., Tews, I., Dauter, Z., Oppenheim, A. B., Chet, I., Wilson, K. S. & Vorgias, C. E. (1994). *Structure*, **2**, 1169–1180.
Rao, F. V., Houston, D. R., Boot, R. G., Aerts, J. M. F. G., Hodkinson, M., Adams, D. J., Shiomi, K., Ōmura, S. & van Aalten, D. M. F. (2005). *Chem. Biol.* **12**, 65–76.
Ryde, U., Olsen, L. & Nilsson, K. (2002). *J. Comput. Chem.* **23**, 1058–1070.
Sheldrick, G. M. (2008). *Acta Cryst.* A**64**, 112–122.
Songsiriritthigul, C., Pantoom, S., Aguda, A. H., Robinson, R. C. & Suginta, W. (2008). *J. Struct. Biol.* **162**, 491–499.
Stewart, J. J. (2009). *J. Mol. Model.* **15**, 765–805.
Stewart, J. J. (2013). *J. Mol. Model.* **19**, 1–32.
Suginta, W. & Sritho, N. (2012). *Biosci. Biotechnol. Biochem.* **76**, 2275–2281.
Synstad, B., Gaseidnes, S., van Aalten, D. M. F., Vriend, G., Nielsen, J. E. & Eijsink, V. G. (2004). *Eur. J. Biochem.* **271**, 253–262.
Taira, T., Fujiwara, M., Dennhart, N., Hayashi, H., Onaga, S., Ohnuma, T., Letzel, T., Sakuda, S. & Fukamizo, T. (2010). *Biochim. Biophys. Acta*, **1804**, 668–675.
Tamada, T., Kinoshita, T., Kurihara, K., Adachi, M., Ohhara, T., Imai, K., Kuroki, R. & Tada, T. (2009). *J. Am. Chem. Soc.* **131**, 11033–11040.
Terwisscha van Scheltinga, A. C., Hennig, M. & Dijkstra, B. W. (1996). *J. Mol. Biol.* **262**, 243–257.
Tews, I., Terwisscha van Scheltinga, A. C., Perrakis, A., Wilson, K. S. & Dijkstra, B. W. (1997). *J. Am. Chem. Soc.* **119**, 7954–7959.
White, A. & Rose, D. R. (1997). *Curr. Opin. Struct. Biol.* **7**, 645–651.
Wick, C. R., Hennemann, M., Stewart, J. J. & Clark, T. (2014). *J. Mol. Model.* **20**, 2159.
Yamaguchi, S., Kamikubo, H., Kurihara, K., Kuroki, R., Niimura, N., Shimizu, N., Yamazaki, Y. & Kataoka, M. (2009). *Proc. Natl Acad. Sci. USA*, **106**, 440–444.
Yang, J., Gan, Z., Lou, Z., Tao, N., Mi, Q., Liang, L., Sun, Y., Guo, Y., Huang, X., Zou, C., Rao, Z., Meng, Z. & Zhang, K.-Q. (2010). *Microbiology*, **156**, 3566–3574.
Yu, N., Li, X., Cui, G., Hayik, S. A. & Merz, K. M. II (2006). *Protein Sci.* **15**, 2773–2784.
Zakariassen, H., Aam, B. B., Horn, S. J., Varum, K. M., Sorlie, M. & Eijsink, V. G. (2009). *J. Biol. Chem.* **284**, 10610–10617.
Zakariassen, H., Hansen, M. C., Jøranli, M., Eijsink, V. G. & Sørlie, M. (2011). *Biochemistry*, **50**, 5693–5703.
Zhao, Y., Bishop, B., Clay, J. E., Lu, W., Jones, M., Daenke, S., Siebold, C., Stuart, D. I., Jones, E. Y. & Aricescu, A. R. (2011). *J. Struct. Biol.* **175**, 209–215.

Files: d/lp5005/lp5005.3d d/lp5005/lp5005.sgml **LP5005** FA IU-1512/21(26)5 1512/20(26)5 () LP5005 PROOFS D:FA:2015:71:5:0:0–0