Acta Crystallographica Section D
Biological Crystallography
ISSN 1399-0047

Crystal structures of substrate-bound chitinase from the psychrophilic bacterium *Moritella marina* and its structure in solution

Piotr H. Malecki,[a] Constantinos E. Vorgias,[b] Maxim V. Petoukhov,[c] Dmitri I. Svergun[c] and Wojciech Rypniewski[a]*

[a]Institute of Bioorganic Chemistry, Polish Academy of Sciences, Noskowskiego 12/14, 61-704 Poznan, Poland, [b]Department of Biochemistry and Molecular Biology, National and Kapodistrian University of Athens, Zografou Campus, 157 01 Athens, Greece, and [c]Hamburg Unit, EMBL c/o DESY, European Molecular Biology Laboratory, Notkestrasse 85, 22607 Hamburg, Germany

Correspondence e-mail: wojtekr@ibch.poznan.pl

Received 8 November 2013
Accepted 26 November 2013

PDB references: chitinase, E153Q mutant, 4mb3; complex with NAG_4, 4mb4; complex with NAG_5, 4mb5

The four-domain structure of chitinase 60 from *Moritella marina* (*Mm*Chi60) is outstanding in its complexity. Many glycoside hydrolases, such as chitinases and cellulases, have multi-domain structures, but only a few have been solved. The flexibility of the hinge regions between the domains apparently makes these proteins difficult to crystallize. The analysis of an active-site mutant of *Mm*Chi60 in an unliganded form and in complex with the substrates NAG_4 and NAG_5 revealed significant differences in the substrate-binding site compared with the previously determined complexes of most studied chitinases. A SAXS experiment demonstrated that in addition to the elongated state found in the crystal, the protein can adapt other conformations in solution ranging from fully extended to compact.

1. Introduction

Glycoside hydrolases (EC 3.2.1.–) hydrolyse the glycosidic bond between carbohydrates or between a carbohydrate and a non-carbohydrate moiety. Glycoside hydrolases belonging to family 18 (GH18) include chitinases (EC 3.2.1.14), endo-β-*N*-acetylglucosaminidases (EC 3.2.1.96) and catalytically inactive proteins that function as enzyme inhibitors or lectins.

Many members of GH18 display a multi-domain architecture with various arrangements of the domains (Carbohydrate Active Enzymes Database; CAZy; Cantarel *et al.*, 2009). The common feature is a catalytic domain with a β/α-barrel fold, the function of which is to hydrolyse the β-1,4-linkage between *N*-acetyl-D-glucosamine (NAG) residues in chitin. Chitinases of the GH18 family contain a characteristic D*X*D*X*E motif in their amino-acid sequence in which the glutamate is the catalytic residue essential for activity (Henrissat & Davies, 1997). Many catalytic domains contain an $\alpha+\beta$ domain or extended loops inserted into the TIM barrel which participate in the formation of the substrate-binding groove, making it deeper or even tunnel-like. Carbohydrate-binding modules (CBMs) designed to bind chitin (ChBDs; chitin-binding domains) are characteristic of chitinolytic enzymes (Akagi *et al.*, 2006). They are found in plant, fungal and bacterial proteins and can be structurally diverse. In addition to those, many chitinases also contain fibronectin type III (FnIII) or immunoglobulin-like (Ig-like) domains (Malecki *et al.*, 2013; Toratani *et al.*, 2006). One or more of these can be found between the catalytic domain and the ChBD. They are postulated to serve as spacers to adjust the position of the ChBD relative to the catalytic domain (Toratani *et al.*, 2006). However, the exact role of the linking domains is still unclear.

Although many multi-domain chitinases have been identified in gene sequences, crystallographic studies are few, presumably because the molecules are flexible and difficult to crystallize. The most complex X-ray structure within GH18 is that of chitinase 60 from *Moritella marina* (*Mm*Chi60; Stefanidi & Vorgias, 2008). This cold-adapted bacterium produces chitinase to utilize chitin as a source of carbon, nitrogen and energy. *Mm*Chi60 consists of four domains: a catalytic domain, two Ig-like domains and a C-terminal ChBD. Another characteristic of *Mm*Chi60 is a shallow substrate-binding groove, as opposed to the most studied chitinases from *Serratia marcescens*, in which the substrate is accommodated in a deep long cleft.

The previously reported *Mm*Chi60 structures revealed the enzyme in an unliganded form and in complexes with oxazolinium ion (a reaction intermediate) and with the NAG_2 reaction product (Malecki *et al.*, 2013). The activity of the enzyme in the crystalline form prevented us from examining the enzyme–substrate complex. Consequently, only interactions in the −1 and −2 subsites could be described. No clear substrate-binding groove could be discerned on the + side; therefore, we decided to investigate the interactions of substrate with the enzyme using a low-activity mutant.

In this work, we present crystal structures of the *Mm*Chi60 E153Q mutant in an unliganded form and in complex with NAG_4 and NAG_5. This is the first report of a full-length shallow-binding-groove chitinase in complex with these substrates. We also examined the wild-type enzyme in solution using small-angle X-ray scattering (SAXS) in order to determine whether the protein is a monomer or a dimer in solution and to investigate its conformation and flexibility.

2. Methods

2.1. Biological source

The chitinase under study is from the bacterial strain *M. marina* (synonym *Vibrio marinus*; ATCC 15381), which was originally isolated from a sample raised from a depth of 1200 m in the Northern Pacific Ocean at a temperature of 3.2°C (Morita & Haight, 1964). The cells were grown at 15°C for 48 h in Bacto Marine broth medium 2216 (Difco, USA).

2.2. Plasmid and DNA manipulations

The chitinase gene named *Mm*Chi60 from *M. marina* was cloned as described in Stefanidi & Vorgias (2008). Mutations were introduced into a plasmid isolated from transformed *Escherichia coli* using a Plasmid Miniprep DNA Purification Kit (EURx, Poland). The active-site mutant was produced by introducing an E153Q mutation by site-directed mutagenesis using the forward primer 5′-CGATATTGATTTA**CAA**CAAGCAGCGATCACAGC-3′ and reverse primer 5′-GCTGTGATCGCTGCTTG**TTG**TAAATCAATATCG-3′, in which the mutated codon GAA to CAA is shown in bold. The PCR product showed tandem repeats of the primer in the cloned DNA. To avoid this, two single-primer extension reactions were performed according to the procedure of Edelheit *et al.* (2009). The PCR products were combined, heated to 95°C and cooled slowly to promote re-annealing of denatured products. Subsequently, a 20 µl sample was exposed to 0.5 µl *Dpn*I restriction enzyme (Fermentas) to digest the parental methylated plasmid strains. The plasmids were used to transfer *E. coli* BL21 MAGIC host cells. The resulting recombinant plasmids were verified by DNA sequencing.

2.3. Overproduction and purification of mutated enzyme

The overproduction and purification steps differ from those reported previously for the native form of *Mm*Chi60. Briefly, the cell culture was grown in Luria–Bertani (LB) medium containing 100 µg ml^{-1} ampicillin and 50 µg ml^{-1} kanamycin at 37°C. Enzyme expression was induced with 1 m*M* isopropyl β-D-1-thiogalactopyranoside (IPTG) at the mid-exponential growth phase and the culture was further incubated at 18°C overnight. Advantage was taken of the fact that the protein of interest is localized in the periplasm. Cells were broken by osmotic shock according to a previously published procedure with some alternations (Poole & Hancock, 1984). *E. coli* cells were suspended and mixed for 10 min in osmotic buffer consisting of 20%(*w*/*v*) sucrose, 0.03 *M* Tris, 3 m*M* EDTA and were spun down at 6000 rev min^{-1}. The cells were resuspended in ice-cold water and the periplasmic proteins were released into the solution by mixing the cold solution for 10 min. The cells were centrifuged at 10 000 rev min^{-1}. The supernatant was adjusted to 1 *M* ammonium sulfate, 20 m*M* sodium phosphate buffer pH 8.0 and directly applied onto a 10 ml Phenyl-Sepharose 6 Fast Flow column previously equilibrated in the same buffer (GE Healthcare). The column was washed with the same buffer and the bound proteins were eluted, concentrated and applied onto a Superdex 200 gel-filtration column (1.6 × 60 cm, Pharmacia) in 20 m*M* Tris buffer, 200 m*M* NaCl pH 8.0. The enzyme purity was assessed by SDS–PAGE (0.1% SDS, 12.5% polyacrylamide; Laemmli, 1970) followed by Coomassie Brilliant Blue staining. The enzyme concentration was determined from the absorption at 280 nm using an UV spectrophotometer. The molar extinction coefficient was calculated using the *ProtParam* tool on the ExPasy server (Gasteiger *et al.*, 2005) to be 106 355 M^{-1} cm^{-1}.

2.4. Crystallization and X-ray data collection

All *Mm*Chi60 E153Q crystals were grown using the hanging-drop vapour-diffusion method at 291 K. Initial crystallization conditions were found in the Morpheus protein crystallization screen (Molecular Dimensions; solution H4) using an Art Robbins Gryphon crystallization robot by mixing 0.4 µl 10 mg ml^{-1} protein solution with 0.4 µl well solution. The well solution consisted of a mixture of precipitants [12.5%(*w*/*v*) PEG 1000, 12.5%(*w*/*v*) PEG 3350 and 12.5%(*v*/*v*) MPD], a mixture of amino acids [sodium L-glutamate, alanine (racemic), glycine, lysine–HCl (racemic) and serine (racemic) each at 20 m*M*] and 0.1 *M* MES–imidazole pH 6.5. NAG ligands (Carbosynth) were soaked in by adding the ligand to the cryosolution at a concentration that exceeded the protein concentration by approximately 5:1. Three soaking

Table 1
SAXS experimental parameters.

Data-collection parameters	
Beamline	P12, EMBL
Beam dimensions (mm)	0.2 × 0.12
Wavelength (nm)	0.124
s range (nm^{-1})	0.07–4.6
Exposure time (s)	1
Concentration range (mg ml^{-1})	0.6–8.8
Temperature (K)	283
Structural parameters	
$I(0)$ (relative) [from $P(r)$]	3000 ± 200
R_g (nm) [from $P(r)$]	3.2 ± 0.1
$I(0)$ (relative) (from Guinier)	3000 ± 200
R_g (nm) (from Guinier)	3.2 ± 0.1
D_{max} (nm)	11 ± 0.5
Porod volume estimate (nm^3)	75 ± 10
Dry volume calculated from sequence (nm^3)	70
Molecular-mass determination	
$I(0)$ (relative) for BSA	4500 ± 200
Molecular weight MW (kDa) [from $I(0)$]	47 ± 10
Calculated monomeric MW from sequence (kDa)	58.6

experiments were performed with *N*,*N*′,*N*″,*N*‴-tetraacetyl chitotetraose (NAG_4), *N*,*N*′,*N*″,*N*‴,*N*⁗-pentaacetyl chitopentose (NAG_5) and *N*,*N*′,*N*″,*N*‴,*N*⁗,*N*⁗′-hexaacetyl chitohexose (NAG_6). The time of immersion of the crystal into the cryosolution containing 25% glycerol and the ligand depended on the length of the NAG polymer. In the case of NAG_4 the crystal was immersed for 10 min, whereas crystals soaked in NAG_5 and NAG_6 were incubated overnight. The crystal without a ligand did not require additional cryosolution. X-ray diffraction data were collected on beamline 14.2 at the BESSY synchrotron in Berlin using a MAR Research MX-225 detector. The data were integrated and scaled using the *XDS* software (Kabsch, 2010).

2.5. Structure determination and refinement

The phase problem was solved by molecular replacement using *Phaser* (McCoy *et al.*, 2007) with the crystal structure of *Mm*Chi60 (PDB entry 4hmc; Malecki *et al.*, 2013) as the search model. The model was refined with the *PHENIX* structure-refinement software (Afonine *et al.*, 2012). The NAG ligands were only modelled if clear difference electron density was observed for the ligand.

2.6. Small-angle X-ray scattering (SAXS) measurements and analysis

Synchrotron X-ray solution-scattering data were collected on the EMBL P12 beamline (DESY, Hamburg; Blanchet *et al.*, 2012) using a Pilatus detector and automated filling (Round *et al.*, 2008). The scattering profiles were measured at several solute concentrations ranging from 0.6 to 8.8 mg ml^{-1}. Data acquisition was performed at a sample-to-detector distance of 2.7 m, covering the momentum-transfer range $0.07 < s < 4.6$ nm^{-1} [$s = 4\pi \sin(\theta)/\lambda$, where 2θ is the scattering angle and $\lambda = 0.12$ nm is the X-ray wavelength] in 20 frames of 0.05 s each to check for possible radiation damage. SAXS experimental parameters are summarized in Table 1.

Primary data reduction was performed using an automated pipeline (Franke *et al.*, 2012), and comprehensive analysis of the scattering profiles was accomplished using the *ATSAS* software package (Petoukhov *et al.*, 2012). The forward scattering $I(0)$ and the radius of gyration R_g were evaluated using the Guinier approximation, assuming that at very small angles ($s < 1.3/R_g$) the intensity can be represented as $I(s) = I(0)\exp[-(sR_g)^2/3]$. The maximum dimension D_{max} was computed using the indirect transform package *GNOM* (Svergun, 1992). Molecular-weight estimation was made by comparison of the forward scattering with $I(0)$ of a bovine serum albumin (BSA) standard. The scattering from the crystallographic monomer of *Mm*Chi60, $I_{calc}(s)$, was calculated using *CRYSOL* (Svergun *et al.*, 1995), which also fitted the experimental data $I_{exp}(s)$ by adjusting the excluded volume and the contrast of the hydration layer to minimize the discrepancy,

$$\chi^2 = \frac{1}{N-1}\sum_{j=1}^{N}\left[\frac{I_{exp}(s_j) - cI_{calc}(s_j)}{\sigma(s_j)}\right]^2, \quad (1)$$

where c is a scaling factor, N is the number of points and σ denotes the experimental error.

An ensemble-optimization method (EOM; Bernadó *et al.*, 2007) was applied to assess the flexibility of *Mm*Chi60 in solution. In the EOM approach, a large pool (of 10 000 *Mm*Chi60 chains) with random orientations of individual domains with respect to each other is generated and the optimized ensemble is selected by a genetic algorithm which averaged-intensity fits the experimental SAXS data. The distribution of the structural descriptors (*i.e.* R_g, D_{max}) in the optimized ensemble with respect to the original random pool provides an idea of the conformation variability of the macromolecule in solution.

3. Results

Three crystal structures were solved and examined in order to evaluate and understand the process of ligand binding to *Mm*Chi60. Point mutation of the active glutamate to glutamine served to debilitate the activity of the enzyme in order to observe a stable enzyme–substrate complex. The mutated active-site residue is part of the D*X*D*X*E signature of glycosyl hydrolases belonging to family 18 (Henrissat & Davies, 1997). The structures were compared with the previously reported wild-type *Mm*Chi60 in an unliganded form and in complex with shorter NAG ligands (Malecki *et al.*, 2013). A SAXS experiment was performed to analyse the flexibility of the protein molecule in solution and to resolve the issue of whether the protein is a monomer or a dimer in solution.

3.1. Structure of the E153Q mutant of *Mm*Chi60

The enzymatic activity of the mutated *Mm*Chi60 E153Q is 10^5 times lower than that of the wild-type enzyme. The mutant was prepared to study the interactions of NAG substrates with the chitinase. Although the crystallization conditions were different, the mutant protein crystallized in the same space

Table 2
Summary of X-ray data-collection and refinement statistics.

Values in parentheses are for the highest resolution shell.

	E153Q	E153Q–NAG_4	E153Q–NAG_5
PDB code	4mb3	4mb4	4mb5
Space group	$P3_112$	$P3_112$	$P3_112$
Unit-cell parameters			
a (Å)	67.66	67.32	67.06
b (Å)	67.66	67.32	67.06
c (Å)	255.64	255.76	256.63
Beamline	14.2, BESSY	14.2, BESSY	14.2, BESSY
Wavelength (Å)	0.91841	0.91841	0.91841
Resolution (Å)	48.3–1.55 (1.64–1.55)	48.12–1.48 (1.57–1.48)	48.05–1.64 (1.74–1.64)
R_{merge}†	0.07 (0.88)	0.05 (0.85)	0.07 (0.90)
R_{meas}‡	0.07 (0.95)	0.06 (0.92)	0.08 (0.98)
$R_{p.i.m.}$§	0.028 (0.39)	0.022 (0.36)	0.031 (0.40)
$CC_{1/2}$	99.9 (80.3)	99.9 (78.3)	99.9 (79.4)
Completeness (%)	99.8 (98.7)	99.8 (99.4)	99.9 (99.3)
Observed reflections	609238	758729	534266
Unique reflections	97956	110782	81639
$\langle I/\sigma(I)\rangle$	16.4 (1.8)	20.5 (2.0)	17.7 (1.9)
Multiplicity	6.2	6.9	6.5
R¶	0.14	0.14	0.15
R_{free}††	0.18	0.17	0.18
Protein atoms	4205	4217	4200
Ligand atoms	23	102	119
Water molecules	909	711	719
Average *B* factor (Å^2)	26	25	24
Error in Luzzati plot (Å)	0.15	0.14	0.16
R.m.s. deviation from ideal			
Bond lengths (Å)	0.02	0.02	0.02
Bond angles (°)	1.63	1.76	1.79
Ramachandran plot (%)			
Favoured	97.6	98.16	97.77
Allowed	2.4	1.84	2.23
Outliers	0	0	0

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the integrated intensity of a given reflection and $\langle I(hkl)\rangle$ is the mean intensity of multiple corresponding symmetry-related reflections. ‡ $R_{meas} = \sum_{hkl}\{N(hkl)/[N(hkl)-1]\}^{1/2} \times \sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$. § $R_{p.i.m.} = \sum_{hkl}\{1/[N(hkl)-1]\}^{1/2} \times \sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$. ¶ $R = \sum_{hkl} ||F_{obs}| - |F_{calc}|| / \sum_{hkl} |F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure factors, respectively. †† R_{free} is *R* calculated using a set of randomly chosen reflections excluded from the refinement.

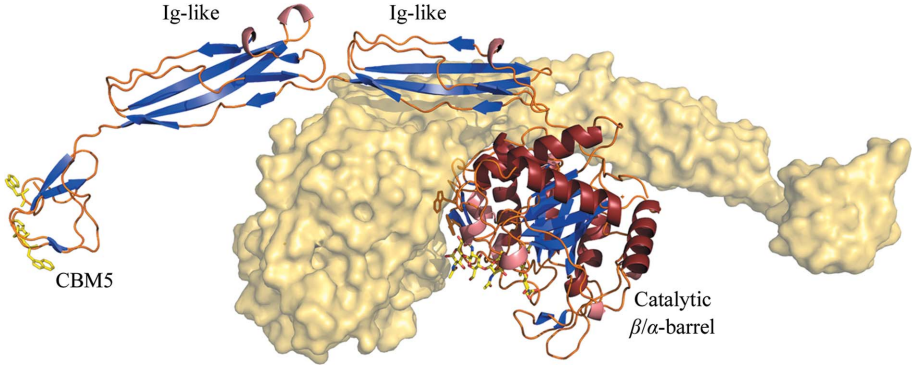

Figure 1
Two *Mm*Chi60 molecules related by the crystallographic twofold axis shown as a cartoon plot and surface representations. The four protein domains are labelled. The secondary-structure elements are coloured as follows: α-helices, red; β-sheets, blue; 3_{10}-helices, pink. NAG_5 can be seen in the substrate-binding site (coloured sticks). Tryptophan residues proposed to be significant in chitin binding are shown as yellow sticks in the chitin-binding domain and as brown sticks in the Trp patch on the insertion domain.

group as the wild type, with very similar unit-cell parameters (Table 2). Consequently, the overall structure of the *Mm*Chi60 E153Q mutant with and without ligands resembles the wild-type structure reported previously (Malecki *et al.*, 2013; Fig. 1).

Two complexes have been obtained: with NAG_4 and NAG_5 bound in the substrate-binding groove of the enzyme. To examine the effect of the mutation on the architecture of the active site in its resting state, the structure of the E153Q mutant was also solved without a ligand. In addition, a structure of the E153Q mutant soaked with NAG_6 was examined, but unfortunately electron density was only observed for five NAG moieties.

In the unliganded E153Q mutant the substrate-binding groove is occupied by solvent water molecules and a sodium ion in a position equivalent to the O7 atom of the acetyl group of NAG in the −1 subsite. In the crystals soaked with NAG_4 and NAG_5, electron density for the ligands was clearly identified in the substrate-binding site. In each structure, a partially occupied (0.6) sulfate ion was found coordinating water molecules bound to the NAG moiety in the −2 site.

Glu153, the active residue in wild-type *Mm*Chi60, was observed previously with the O atoms of the side chain stationed against the N atoms of the main chains of Asp119 and Ala120 in the nearby flexible loop (Malecki *et al.*, 2013). In the E153Q mutant the glutamine residue does not occupy the 'away' position observed in the native protein but points towards the substrate-binding groove and is hydrogen-bonded to the second Asp of the D*X*D*X*E motif. If a ligand is present it is also hydrogen-bonded to the NAG moiety occupying the −1 site. The same position of residue 153 was observed in the active form of the enzyme in complex with the reaction product. The conformations of Asp149 and Asp151 of the D*X*D*X*E motif are unchanged in all of the presented structures. In the E153Q mutant, Asp151 is in close proximity (2.7–3 Å) to Gln153. The interaction between the two residues is *via* O atoms. The correctness of the modelling was checked by swapping $O^{\varepsilon1}$ and $N^{\varepsilon2}$: the *B* factors changed from no difference to a difference of 5 Å^2.

3.2. Substrate accommodation

The location of the substrate-binding groove is similar to that in other chitinases with the β/α-barrel fold, but the groove is shallow and is only clearly defined in the immediate vicinity of the active site. NAG_4 was observed in the −2, −1, +1 and +2 subsites (Fig. 2). In the crystal soaked in NAG_5 the +3 site was also occupied by a NAG moiety (Fig. 2). A similar positioning of the NAG_5 substrate was observed in chitinase B from *S. marcescens* (PDB entry 1e6n), in which two NAG subunits were modelled in the glycon (−) subsite and three in the aglycon (+) subsite (van

Aalten *et al.*, 2001). This is a different substrate-binding mode from chitinase A (PDB entry 1ehn), in which only two NAG units were observed in the aglycon site, whereas in the glycon subsite five or more were modelled (Papanikolau *et al.*, 2001).

All available chitinase structures in complex with ligands have a tunnel-like aglycon site. In *Mm*Chi60 E153Q there is no tunnel or clear groove on the aglycon side, and the second and third NAG moieties are bound on the flat surface of the protein; they are hydrogen-bonded to the enzyme only on one side, with the other face accessible to the solvent (Fig. 3). All interactions between the NAG units and the enzyme are summarized in Table 3. NAG_4 is closely superimposable on the NAG_5 ligand, with an almost identical conformation of the superposed parts (Fig. 2). One difference can be noted in the +2 subsite, in which the terminal sugar ring of the NAG_4 ligand is statically disordered, showing both anomeric states. In addition to the β anomer common to NAG polymers, the alternative configuration of the +2 sugar is α and its O1 atom is hydrogen-bonded to the O atom of Asn221. The terminal moiety of NAG_5 is also in the α anomeric state.

The conformations of all of the sugar rings can be clearly identified based on electron-density maps. In the −2, +1, +2 and +3 subsites the sugars adopt the standard chair conformation, whereas the sugar ring in the −1 subsite is distorted into a boat conformation. Similar distortions were previously observed in other inactive chitinase structures with a bound ligand or in an active enzyme in complex with the product (Malecki *et al.*, 2013; Papanikolau *et al.*, 2001; Songsiriritthigul *et al.*, 2008; Tsuji *et al.*, 2010; van Aalten *et al.*, 2001). The boat conformation makes the sugar chain bent and twisted and helps to position the acetyl group of the NAG in the −1 subsite at the bottom of the active-site groove, which is essential for substrate-assisted formation of the reaction intermediate (an oxazolinium ion). The angle at which the substrate is bent (measured between C4 and C1 of NAG in the −1 subsite and C4 of NAG in the +1 subsite) is 98°. The average temperature factor of the NAG units shows that the moieties in the −1 and +1 sites have the lowest values, indicating that they are the best ordered parts of the NAG_5 ligand: 13 Å^2 for the −2 subsite, 11 Å^2 for the −1 subsite, 12 Å^2 for the +1 subsite, 14 Å^2 for the +2 subsite and 20 Å^2 for the +3 subsite. It is interesting to note that in the substrate the best

Table 3
NAG_5 ligand interaction with the protein molecule.

NAG moiety	Ligand atom	Bond length (Å)	Residue atom	Residue
−2	O6	3.1	N	Ala118
	O6	2.8	O	Ala118
	O7	3.0	N^{ε}	Trp311
	O7	2.9	N^{ε}	Trp315
−1	O7	2.6	OH	Tyr220
	N2	3.1	O^{δ}	Asp151
	N2	3.2	O^{ε}	Gln153
	O3	3.1	N	Ala118
	O6	2.9	N^{δ}	Asn221
	O6	2.7	O	Ala279
1	O6	2.7	O^{ε}	Gln218
	O6	2.9	N^{δ}	Asn221
	O7	2.7	O^{ε}	Tyr46
2	N2	2.9	O^{ε}	Gln222
	N2	3.2	$O^{\varepsilon 1}$	Glu191
	O3	2.6	$O^{\varepsilon 1}$	Glu191
3	O7	2.9	N	Gly224
	O6	2.5	O^{δ}	Asp225
	O1	3.0	N^{ε}	Gln154 (symmetry-related molecule)

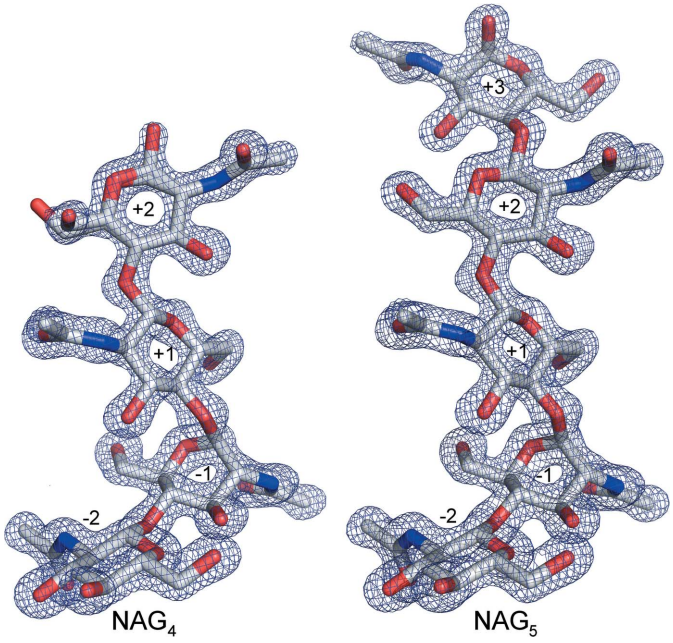

Figure 2
Weighted $2F_o - F_c$ maps contoured at the 2σ level for ligands bound in the substrate-binding groove. The ligands occupy subsites −2 to +2 for the crystal soaked with NAG_4 (left) and −2 to +3 after soaking with NAG_5 (right).

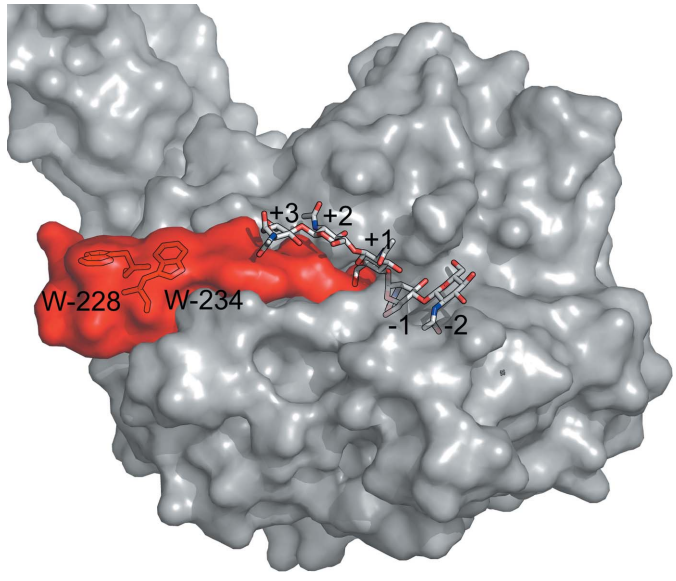

Figure 3
Surface representation of the E153Q *Mm*Chi60 molecule. The insertion domain (residues 218–235) forms a platform-like structure on the aglycon side (red). Trp228 and Trp 234 forming the 'Trp patch' are shown. NAG_5 is represented as sticks.

ordered NAG moiety is at the active centre, *i.e.* in the −1 subsite, whereas in the product the NAG unit in this position has a higher *B* factor than the moiety in the −2 subsite 1 (Malecki *et al.*, 2013).

3.3. Flexibility in the ligand-binding region

In the previous study, the catalytic Glu153 and the flexible loop Gly116–Ile122 were observed to shift towards the ligand upon its binding (Malecki *et al.*, 2013). In this study of the E153Q mutant, even though the mutated residue points toward the binding groove in all cases, as observed previously in the product-bound state, the loop adopts different conformations depending on whether or not a ligand is present (Fig. 4*a*). In the complex of E153Q *Mm*Chi60 with NAG_5, the O3 atom of NAG in the −1 subsite is stabilized by the main-chain N atom of Ala118 which is shifted by 1.8 Å in comparison to the unliganded structure. The same N atom is hydrogen-bonded to the O6 atom of the NAG moiety in the −2 subsite. The carbonyl O atom of Ala118 also moves by 2.2 Å to partake in binding to the O6 atom of the NAG moiety in the −2 subsite. The flexible loop in the unliganded E153Q structure is located in an intermediate position between the unliganded native enzyme and the liganded mutant (Fig. 4*a*).

The amino-acid segment 218–235 participates in ligand binding on the aglycon side of the substrate-binding site. With respect to the overall topology, this is an insertion domain in the β/α-barrel consisting mainly of two antiparallel β-strands connected by a loop, classified as a β-hairpin. The insertion domain forms a platform-like structure on the surface of the catalytic domain (Fig. 3). The β-hairpin displays a mobility of up to 2 Å with respect to the catalytic domain, as observed on superposition of the present structures and the previously published *Mm*Chi60 structures (Malecki *et al.*, 2013; Fig. 4*b*). The residues Gln218, Asn221, Gln222, Gly224 and Asp225 are

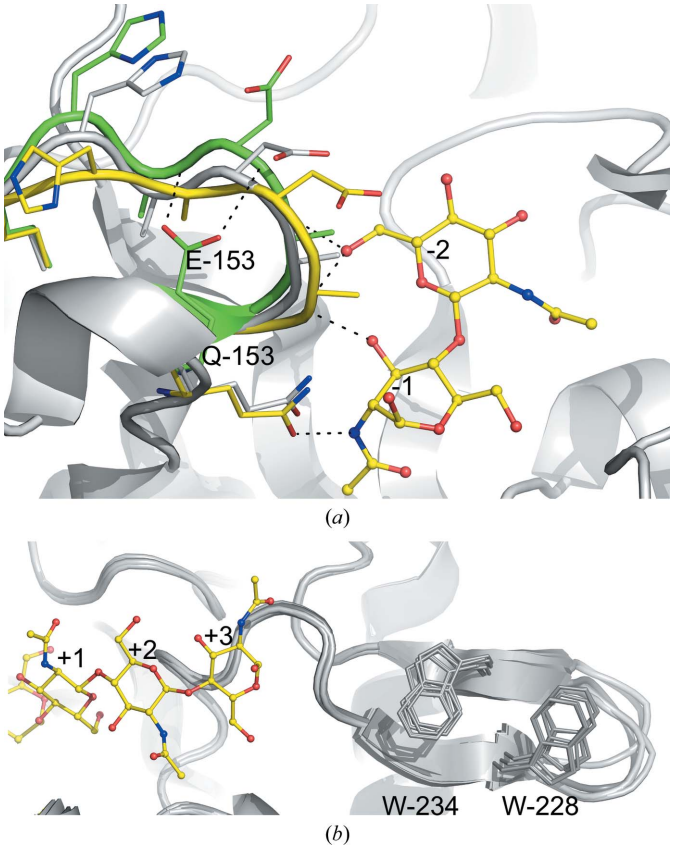

Figure 4
(*a*) Superposition showing the flexible loop on the glycon side of the substrate-binding site: the wild-type enzyme with the catalytic Glu153 'parked' against the loop, as in Malecki *et al.* (2013), is shown in green, the unliganded mutant E153Q is shown in grey and the liganded mutant is shown in yellow, showing part of the NAG_5 substrate bound in the −2 and −1 subsites. (*b*) Superposition showing the insertion subdomain forming a tongue-like structure with the Trp patch on its surface. The three structures presented in this paper are superposed as well as the three structures from Malecki *et al.* (2013). The subdomain exhibits considerable mobility. Part of the NAG_5 substrate is shown on the aglycon side of the substrate-binding site (coloured sticks).

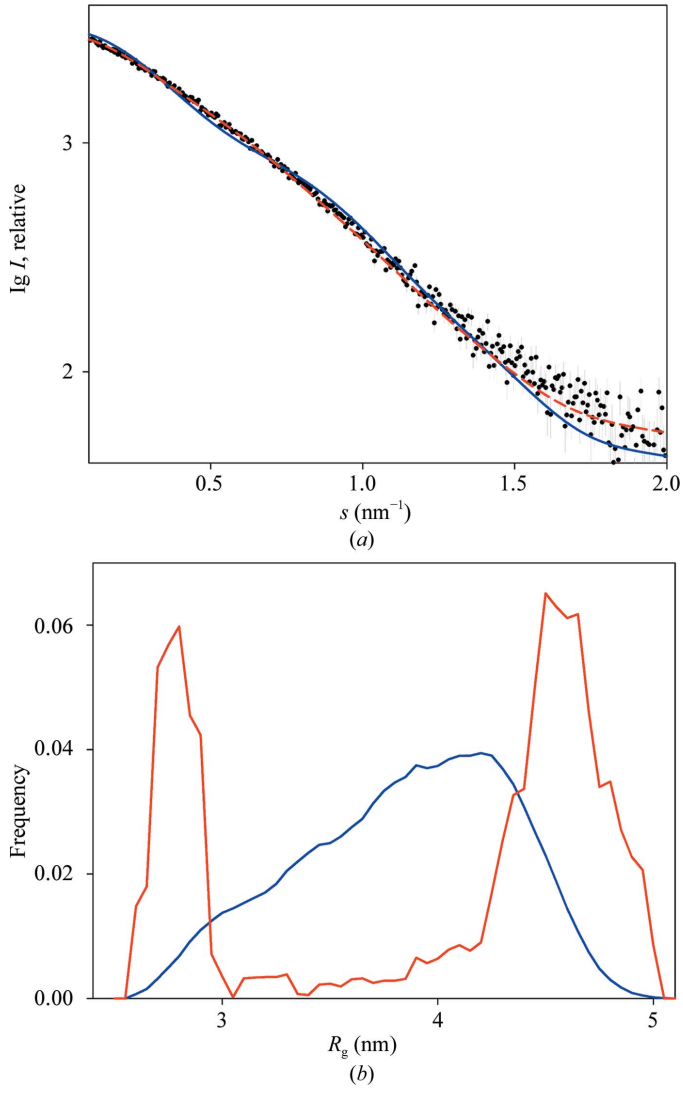

Figure 5
(*a*) SAXS profiles of *Mm*Chi60. Experimental data are denoted by black dots with the fit from the crystallographic monomer shown as a blue solid line and the fit from the EOM ensemble as a red dashed line. (*b*) R_g distribution of *Mm*Chi60 in solution obtained by EOM. Random pool, blue; distribution in the optimized ensemble, red.

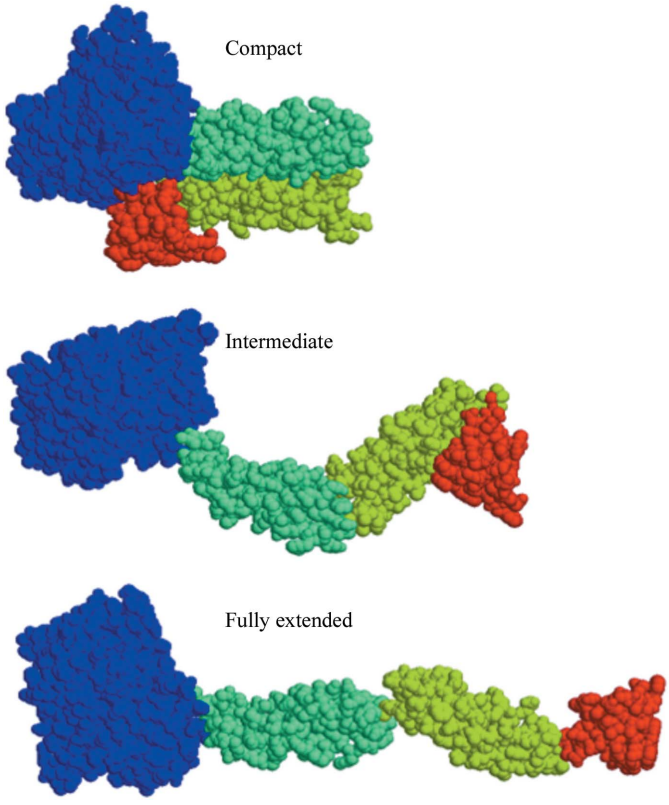

Figure 6
Three conformations of *Mm*Chi60 from the EOM set consistent with the SAXS data.

hydrogen-bonded to the NAG moieties of NAG_5 occupying the aglycon subsites (Supplementary Fig. S1[1]). It appears that longer ligands would be blocked from interactions on this side by close contacts with a symmetry-related protein molecule.

The bound substrate on the + site points in the direction of the β-hairpin (residues 226–235). The substrate-binding area extends towards two solvent-exposed tryptophan residues, Trp228 and Trp234, which can be considered as a secondary chitin-binding element (with the main chitin-binding element being the chitin-binding domain).

3.4. Conformational freedom in solution

The experimental SAXS profile of *Mm*Chi60 is presented in Fig. 5(*a*) and the corresponding structural parameters are summarized in Table 1. The SAXS data point to a monomeric state of *Mm*Chi60 in solution, even though the crystal packing suggests a dimeric form (Fig. 1). The overall size and radius of gyration (R_g) computed from the rather elongated crystallographic monomer are noticeably larger than the experimentally deduced values in solution. The scattering profile computed by *CRYSOL* from the atomic model of *Mm*Chi60 displays several shoulders (Fig. 5*a*), whereas the experimental curve demonstrates a rather smooth behaviour. These deviations between the experimental and computed profiles yield a discrepancy of $\chi = 1.7$. As previously proposed (Malecki *et al.*, 2013), *Mm*Chi60, being a multi-domain protein, could possess considerable conformational freedom. The discrepancy between the computed and experimental SAXS profiles, in which the former curve oscillates around the latter curve, is a typical result of flexibility in solution. Indeed, it has been shown in a theoretical study (Bernadó, 2010) that domain motion often results in 'smearing' of the SAXS curve.

To check the hypothesis of inter-domain flexibility, the ensemble-optimization method (EOM) was applied. The full-length protein was divided into four individual domains and these were allowed to adopt random orientations with respect to each other to sample the conformational space. The distribution of the R_g values in the original random pool is presented in Fig. 5(*b*) (blue line). The optimized ensemble consisting of compact, completely extended and intermediate conformations yielded a good fit to the experimental data for *Mm*Chi60 with $\chi = 1.2$ (Fig. 5*a*). The corresponding R_g distribution presented in Fig. 5(*b*) (red line) displays two major populations: completely extended (R_g around 4.7 nm) and rather compact (R_g around 2.7 nm). Overall, our results suggest a high degree of flexibility in the hinge regions of *Mm*Chi60 in solution (Fig. 6). Interestingly, the intermediate conformation displays an overall shape rather similar to the crystal structure (Figs. 1 and 6). The intermediate conformation appears to be present in solution in smaller amounts and manifests itself as a shoulder of the peak corresponding to the completely open structure.

4. Discussion

4.1. Effect of the E153Q mutation

Kinetic and structural studies have shown that the E153Q mutation significantly debilitates the enzymatic activity and also affects the position of residue 153. The replacement of the carbonyl group by a neutral amide group entailed a considerable structural rearrangement. The flipping of the active-site residue upon ligand binding observed in previously reported structures (Malecki *et al.*, 2013) is abolished in the mutated *Mm*Chi60. The absence of one carbonyl group prevents it from 'parking' against the main-chain amine groups of the flexible loop nearby. This indicates that both O atoms are essential for the 'away' conformation in the resting state. The mutated residue in the unliganded structure is directed towards the active site and is hydrogen-bonded to the neighbouring Asp151. In the unliganded and liganded structures Gln153 interacts with Asp151 *via* the O atoms (3 Å), indicating the presence of an H atom in between. This is somewhat surprising because in the interaction between the Gln and Asp side chains one would expect the amine group to participate in the hydrogen bond. The evident preference for the O atom of Gln153 suggests that the O atom of Asp151 is already protonated.

4.2. Substrate binding

The shallowness of the substrate-binding site makes it easily accessible to the substrate. Upon substrate binding, confor-

[1] Supporting information has been deposited in the IUCr electronic archive (Reference: DZ5316).

mational changes can be observed in the glycon and aglycon binding sites of the protein.

4.2.1. Induced fit of the Gly116–Ile122 loop. Three stable positions of the Gly116–Ile122 loop can be distinguished (Fig. 4*a*). In the resting state of the wild-type enzyme this flexible loop is in the distal position with respect to the substrate glycon binding site. The active Glu153 is stationed against the main-chain N atoms of the loop. This conformation makes the glycon substrate-binding site wide and accessible to the substrate (Malecki *et al.*, 2013). In the unliganded state of the E153Q mutant, with Gln153 detached from the flexible loop, the loop is shifted towards the substrate-binding groove to a position which can be considered intermediate. Upon the binding of NAG oligomers by the mutated *Mm*Chi60, the loop is in the proximal position with respect to the substrate, as in the complex of the active enzyme with the reaction product NAG_2. It seems that the binding of the substrate induces conformational changes that are favourable for interaction with the protein. Owing to the approach of the flexible loop three hydrogen bonds are created, positioning the substrate in the −1 and −2 subsites. The closing movement of the loop could be considered an induced fit because it is only observed in the presence of substrate.

4.2.2. Extent of the binding site on the aglycon (insertion-domain) and glycon sides. The ligand-binding study reveals the role of the subdomain composed of a β-hairpin that forms a platform-like structure on the surface of the protein on the aglycon side of the substrate-binding site (Fig. 3). This can be contrasted with ChiA or ChiB from *S. marcescens*, which have a deep substrate-binding groove in this area. Similar insertion domains to that in *Mm*Chi60 have been observed in ChiC from *S. marcescens* (Payne *et al.*, 2012) and a chitinase from *Lactococcus lactis* (PDB entry 3ian; New York SGX Research Center for Structural Genomics, unpublished work), but these structures are unliganded. The platform in *Mm*Chi60 is made up of amino acids 218–235, some of which are engaged in substrate binding *via* hydrogen bonds (Table 3). The platform-like structure on the aglycon side extends from the +1 subsite towards the nearly coplanar solvent-exposed Trp228 and Trp234. The last observed NAG moiety, in the +3 subsite, is approximately two NAG units short of the Trp patch (Figs. 3 and 4*b*).

Unfortunately, we were not able to observe longer ligands binding to this platform, although they were also used in soaking experiments. Short soaking with NAG_6 resulted in diffraction, but only a NAG pentamer was clearly visible in the electron-density map, spanning sites +3 to −2. Overnight incubation with NAG_6 resulted in cracking of the crystals and no diffraction. It seems that crystal contacts on the + (aglycon) side obstruct longer NAG oligomers from binding. Crystal damage after longer soaking times suggests that the ligand exhibits affinity for an area extending beyond the +3 subsite.

It is interesting why NAG_6 did not simply occupy subsites +3 to −3, as there seems to be unobstructed space for at least one more NAG residue on the − side. Comparison of the liganded chitinase A from *S. marcescens* with *Mm*Chi60 revealed that residues involved in ligand binding in the −3 subsite are not conserved in *Mm*Chi60. It is possible that substrate binding in *Mm*Chi60 does not extend beyond subsite −2.

4.3. Solution studies

SAXS measurements helped to resolve the issue of possible dimerization suggested by crystal contacts (Fig. 1), demonstrating that *Mm*Chi60 is a monomer in solution. In addition, the SAXS experiment indicated that *Mm*Chi60 adopts a wide range of conformations in an aqueous environment ranging from compact, with all of the domains close to each other, through an elongated sea-horse shape similar to that found in the crystal, to fully extended (Fig. 6). The ensemble of conformations generated in modelling the SAXS profile gives a rather rough idea of the dynamic behaviour of the molecule, but allowing inter-domain flexibility is reasonable and there is also some evidence for it in the crystallographic model. In particular, the linker between the chitin-binding domain and the rest of the molecule has a clearly elevated temperature factor and poorly defined electron density.

'Flexibility' has been postulated as a crucial factor in the adaptation of enzymes to cold. In common understanding, it pertains primarily to the binding/active site, improving the ability of the enzyme to recognize and admit the substrate under unfavourable thermodynamic conditions. Given the extended architectures of *Mm*Chi60 and its substrate, flexibility of the whole protein molecule should also be advantageous, allowing the enzyme to cling to chitin and probe its surface simultaneously.

A similar multi-domain protein architecture is found in many bacterial chitinases, but apart from *Mm*Chi60 only the spatial arrangement of chitinase A1 from *Bacillus circulans* has been experimentally investigated. The latter was determined by a combination of X-ray crystallography, SAXS and NMR (Toratani *et al.*, 2006). There is no similarity in the primary structures of the two enzymes; however, the arrangement of the domains is similar: an N-terminal catalytic domain, two linking domains and a C-terminal chitin-binding domain. The convergent evolution of this kind of chitinase suggests advantages of such an architecture. The flexibility of the molecule and its distributed chitin-binding elements could enable it to disrupt the surface of insoluble chitin.

4.4. Reaction mechanism and substrate-breaking strategy

The enzymatic reaction mechanism is probably the same among chitinases belonging to the GH18 family (van Aalten *et al.*, 2001), but different enzymes use various strategies to maximize their turnover rate. Chitin is a 'difficult' substrate: in the crystalline form the NAG chains are highly interconnected by hydrogen bonds and form a rather featureless surface. For the enzyme to act, the surface needs to be disrupted. In addition to the catalytic domain, chitinases usually contain other elements that interact with chitin, providing affinity and facilitating access to the substrate. Some chitinases have been found to process the substrate sequentially. In processive enzymes, a single carbohydrate chain slides along the

substrate-binding cleft while disaccharides are cleaved off successively, as observed for ChiA and ChiB from *S. marcescens* (Davies & Henrissat, 1995). In processive chitinases the binding groove is deep or even tunnel-shaped. On the other hand, the nonprocessive enzymes are characterized by shallow substrate-binding grooves, as in the three-domain exochitinase ChiC from *S. marcescens* (Horn *et al.*, 2006; Suzuki *et al.*, 1999; Vaaje-Kolstad *et al.*, 2005).

The reaction mechanism of *Mm*Chi60 and the role of the flexible active Glu153 have been discussed previously based on three crystal structures: unliganded wild-type enzyme, a complex with an oxazolinium ion (the reaction intermediate) and a complex with the reaction product NAG_2 (Malecki *et al.*, 2013). The crystal structures reported in the present paper add to this the initial stage of the enzymatic cycle: substrates bound to a mutant of reduced activity. This is the first structure of a chitinase with a shallow binding groove and a β-hairpin insertion domain in complex with substrate. Among chitinases with shallow binding grooves it is also the only one to clearly show a bound NAG moiety in the +3 subsite. The NAG chain bound in the aglycon region points towards the β-hairpin, which forms an elevated platform with the tryptophan patch on its surface. This suggests a role of this secondary chitin-binding element (after the chitin-binding CMB5 domain) in the enzyme's mechanism of disrupting the structure of chitin before it can be hydrolysed (Payne *et al.*, 2012).

The overall structure of *Mm*Chi60 is suggestive as to the mode of action of the enzyme as it skims the surface of chitin. The role of the chitin-binding domain would be to provide general affinity of the protein for chitin, and the long, flexibly hinged Ig-like domains would afford the catalytic domain the freedom to probe the surface of chitin, while the tongue-like insertion domain with the Trp patch on the surface of the catalytic domain (Fig. 4*b*) would help to orient it towards the substrate and dislodge NAG chains from the surface of the chitin prior to hydrolysis. The demonstrated catalytic efficiency of *Mm*Chi60 at low temperature could make it useful as a biocatalyst (Stefanidi & Vorgias, 2008).

The project was co-funded by the European Union within the European Regional Development Fund (grant UDA-POIG.02.01.00-30-182/09 and project MPD). This work was supported by the European Commission (the Seventh Framework Programme) BioStruct-X project (contract No. 283570)

References

Aalten, D. M. F. van, Komander, D., Synstad, B., Gåseidnes, S., Peter, M. G. & Eijsink, V. G. H. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 8979–8984.

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.

Akagi, K., Watanabe, J., Hara, M., Kezuka, Y., Chikaishi, E., Yamaguchi, T., Akutsu, H., Nonaka, T., Watanabe, T. & Ikegami, T. (2006). *J. Biochem.* **139**, 483–493.

Bernadó, P. (2010). *Eur. Biophys. J.* **39**, 769–780.

Bernadó, P., Mylonas, E., Petoukhov, M. V., Blackledge, M. & Svergun, D. I. (2007). *J. Am. Chem. Soc.* **129**, 5656–5664.

Blanchet, C. E., Zozulya, A. V., Kikhney, A. G., Franke, D., Konarev, P. V., Shang, W., Klaering, R., Robrahn, B., Hermes, C., Cipriani, F., Svergun, D. I. & Roessle, M. (2012). *J. Appl. Cryst.* **45**, 489–495.

Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V. & Henrissat, B. (2009). *Nucleic Acids Res.* **37**, D233–D238.

Davies, G. & Henrissat, B. (1995). *Structure*, **3**, 853–859.

Edelheit, O., Hanukoglu, A. & Hanukoglu, I. (2009). *BMC Biotechnol.* **9**, 61.

Franke, D., Kikhney, A. G. & Svergun, D. I. (2012). *Nucl. Instrum. Methods Phys. Res. A*, **689**, 52–59.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D. & Bairoch, A. (2005). *The Proteomics Protocols Handbook*, edited by J. M. Walker, pp. 571–607. Totowa: Humana Press.

Henrissat, B. & Davies, G. (1997). *Curr. Opin. Struct. Biol.* **7**, 637–644.

Horn, S. J., Sikorski, P., Cederkvist, J. B., Vaaje-Kolstad, G., Sørlie, M., Synstad, B., Vriend, G., Vårum, K. M. & Eijsink, V. G. (2006). *Proc. Natl Acad. Sci. USA*, **103**, 18089–18094.

Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.

Laemmli, U. K. (1970). *Nature (London)*, **227**, 680–685.

Malecki, P. H., Raczynska, J. E., Vorgias, C. E. & Rypniewski, W. (2013). *Acta Cryst.* D**69**, 821–829.

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.

Morita, R. Y. & Haight, R. D. (1964). *Limnol. Oceanogr.* **9**, 103–106.

Papanikolau, Y., Prag, G., Tavlas, G., Vorgias, C. E., Oppenheim, A. B. & Petratos, K. (2001). *Biochemistry*, **40**, 11338–11343.

Payne, C. M., Baban, J., Horn, S. J., Backe, P. H., Arvai, A. S., Dalhus, B., Bjørås, M., Eijsink, V. G., Sørlie, M., Beckham, G. T. & Vaaje-Kolstad, G. (2012). *J. Biol. Chem.* **287**, 36322–36330.

Petoukhov, M. V., Franke, D., Shkumatov, A. V., Tria, G., Kikhney, A. G., Gajda, M., Gorba, C., Mertens, H. D. T., Konarev, P. V. & Svergun, D. I. (2012). *J. Appl. Cryst.* **45**, 342–350.

Poole, K. & Hancock, R. E. (1984). *Eur. J. Biochem.* **144**, 607–612.

Round, A. R., Franke, D., Moritz, S., Huchler, R., Fritsche, M., Malthan, D., Klaering, R., Svergun, D. I. & Roessle, M. (2008). *J. Appl. Cryst.* **41**, 913–917.

Songsiriritthigul, C., Pantoom, S., Aguda, A. H., Robinson, R. C. & Suginta, W. (2008). *J. Struct. Biol.* **162**, 491–499.

Stefanidi, E. & Vorgias, C. E. (2008). *Extremophiles*, **12**, 541–552.

Suzuki, K., Taiyoji, M., Sugawara, N., Nikaidou, N., Henrissat, B. & Watanabe, T. (1999). *Biochem. J.* **343**, 587–596.

Svergun, D. I. (1992). *J. Appl. Cryst.* **25**, 495–503.

Svergun, D., Barberato, C. & Koch, M. H. J. (1995). *J. Appl. Cryst.* **28**, 768–773.

Toratani, T., Kezuka, Y., Nonaka, T., Hiragi, Y. & Watanabe, T. (2006). *Biochem. Biophys. Res. Commun.* **348**, 814–818.

Tsuji, H., Nishimura, S., Inui, T., Kado, Y., Ishikawa, K., Nakamura, T. & Uegaki, K. (2010). *FEBS J.* **277**, 2683–2695.

Vaaje-Kolstad, G., Horn, S. J., van Aalten, D. M. F., Synstad, B. & Eijsink, V. G. H. (2005). *J. Biol. Chem.* **280**, 28492–28497.