This is a repository copy of *Structure of the GH9 glucosidase/glucosaminidase from Vibrio cholerae*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/155225/

Version: Published Version

**Article:**
Wu, Liang orcid.org/0000-0003-0294-7065 and Davies, Gideon J. orcid.org/0000-0002-7343-776X (2018) Structure of the GH9 glucosidase/glucosaminidase from Vibrio cholerae. Acta Crystallographica Section F: Structural Biology Communications. pp. 512-523. ISSN 2053-230X

https://doi.org/10.1107/S2053230X18011019

---

**Reuse**
This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:
https://creativecommons.org/licenses/

**Takedown**
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.
**Structure of the GH9 glucosidase/glucosaminidase from Vibrio cholerae**

Liang Wu and Gideon J. Davies*

York Structural Biology Laboratory, Department of Chemistry, The University of York, York YO10 5DD, England.
*Correspondence e-mail: gideon.davies@york.ac.uk

Glycoside hydrolase family 9 (GH9) of carbohydrate-processing enzymes primarily consists of inverting endoglucanases. A subgroup of GH9 enzymes are believed to act as exo-glucosidases or exo-glucosaminidases, with many being found in organisms of the family Vibrionaceae, where they are proposed to function within the chitin-catabolism pathway. Here, it is shown that the GH9 enzyme from the pathogen *Vibrio cholerae* (hereafter referred to as VC0615) is active on both chitosan-derived and /C12-glucoside substrates. The structure of VC0615 at 3.17 Å resolution is reported from a crystal form with poor diffraction and lattice disorder. VC0615 was highly refractory to crystallization efforts, with crystals only appearing using a high protein concentration under conditions containing the precipitant poly-/C13-glutamic acid (PGA). The structure is highly mobile within the crystal lattice, which is likely to reflect steric clashes between symmetry molecules which destabilize crystal packing. The overall tertiary structure of VC0615 is well resolved even at 3.17 Å resolution, which has allowed the structural basis for the exo-glucosidase/glucosaminidase activity of this enzyme to be investigated.

1. Introduction

*Vibrio cholerae* is the pathogen responsible for cholera, a disease characterized by severe diarrhoea that is estimated to affect ~3 million people worldwide and causes nearly 100 000 deaths per annum (Ali et al., 2015). Bacteria of the family Vibrionaceae, which includes the *Vibrio* genus, occur naturally as members of the marine bacterioplankton community, where they form part of the bacterial flora of chitinous zooplankton. Vibrionaceae readily utilize chitin as a source of organic carbon and nitrogen, and their attachment to chitinous zooplankton such as copepods has been hypothesized to provide a nutrient-rich habitat for the bacteria (Heidelberg et al., 2002). The association of *V. cholerae* with copepods contributes significantly to the cholera disease burden, such that an increased copepod number in water sources is positively correlated with increased outbreaks of cholera (Huq et al., 2005). Chitin has also been suggested to directly affect the physiology of *V. cholerae*, for example by protecting the bacteria from cold stress (Amako et al., 1987) or from killing by gastric acid (Nalin et al., 1979). The key importance of chitin in the life cycle of *V. cholerae* and other Vibrionaceae demands a detailed understanding of the biochemical pathways responsible for chitin utilization by these organisms, many of which contribute significantly to the global disease burden.

The *V. cholerae* enzyme VC0615 was initially identified as a putative inverting exoglucosidase (termed BglA) based on its
clear exo activity on cello-oligosaccharides and aryl-β-glucosides and its apparent lack of activity on β-1,4-linked disaccharides of GlcN and GlcNAc (Park et al., 2002). Based on sequence homology, VC0615 was assigned to glycoside hydrolase family 9 (GH9) of the Carbohydrate-Active enZymes (CAZy) classification scheme (Cantarel et al., 2009; The CAZYpedia Consortium, 2018), a family primarily associated with endoglucanases (‘cellulases’). Given that VC0615 was found in a chitin-utilization operon, this apparent inconsistency between specificity and operon perplexed the authors, as they understood that V. cholerae was unable to metabolize cellobiose or other higher cello-oligosaccharides.

In 2008, a large-scale assessment of chitin-utilization pathways in the Vibrionaceae suggested that VC0615 was in fact a exo-acting glucosaminidase, the substrate of which was the GlcN-GlcN-6-O disaccharide resulting from the import of chitobiose (GlcN)2 through a phosphotransferase transport system (Hunt et al., 2008). [Note: here we use chitobiose to refer to the β-1,4 (GlcN)2 disaccharide and N,N'-diacetylechitobiose to refer to the β-1,4 (GlcNAc)2 disaccharide, as outlined in the IUPAC–IUB naming convention guidelines (Nomenclature Committee of IUB and IUPAC–IUB Joint Commission on Biochemical Nomenclature, 1988).]

More recently, Honda and coworkers showed that a close (60% identity) homologue of VC0615, PBPR0520 from Photobacterium profundum S9, was indeed a multifunctional exo-acting glucosaminidase/β-glucosidase, with a $k_{cat}/K_m$ at least ten times greater for chito-oligosaccharides compared with the corresponding cello-oligosaccharides, supporting the functional assignment of this enzyme as an exo-acting glucosaminidase active on chitobiose substrates (Honda et al., 2011). Subsequently, a crystal structure of PBPR0520 confirmed this enzyme to be an exo-acting GH9, illustrating that loop extensions (compared with endo-acting GH9 enzymes) blocked the −4 to −2 subsites of the enzyme, rendering it unable to bind substrates in an endo fashion (Honda et al., 2016).

To date, 16 different three-dimensional structures are known for CAZy family GH9. 13 of these known structures are endoglucanases, reflecting their dominance within the GH9 family. One structure of a Clostridium thermocellum cellobiohydrolase (CbhA) has also been reported. Whilst not strictly an exo-acting enzyme, as it hydrolyses terminal cellobiose units rather than terminal glucose units, Schubot and coworkers classified CbhA as an ‘exocellulase’ based on the closed topology of the active site of the enzyme, which blocked the −4 to −2 subsites of the enzyme, rendering it unable to bind substrates in an endo fashion (Honda et al., 2016).

hydrolysis of glucosides by VC0615 was still sufficient to determine Michaelis–Menten kinetic parameters for aryl β-glucoside substrates. Despite the inferior data, the 3.17 Å resolution structure was sufficient to confirm the structural basis of how this enzyme functions as an exo-glycosidase through extended loops from α-helices 1–2 and α-helices 9–10, which block substrate binding beyond the −1 subsite.

2. Materials and methods
2.1. Macromolecule production

A synthesized cDNA encoding VC0615, cloned into the NdeI (5′) and XhoI (3′) sites of the pET-21a vector, was purchased from GenScript. This construct contains the full VC0615 gene followed by a C-terminal hexahistidine tag. The correctness of the cloned plasmid was verified by sequencing with the T7-fwd and pET-RP primers (Table 1) prior to use in protein production.

The pET-21a-VC0615 plasmid was used to transform chemically competent Escherichia coli BL21 Gold (DE3) cells. Cells were grown for protein production in Terrific Broth (TB) containing ampicillin (100 μg ml⁻¹) at 37°C with shaking. Upon reaching an optical density at 600 nm (OD₆₀₀) of 0.8, the cultures were induced with 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG), followed by incubation at 16°C overnight with shaking.

The cells were harvested by centrifugation at 4000g for 15 min at 4°C. The pellet was resuspended in HisTrap buffer A (20 mM Tris pH 8.0, 500 mM NaCl, 20 mM imidazole, 1 mM DTT) supplemented with EDTA-free cComplete protease-
inhibitor cocktail (Roche) and bovine pancreatic DNase I (Sigma), and the cells were lysed by passage through a cell disruptor (Constant Systems) at a pressure of 172 MPa. The cell lysate was cleared by centrifugation at 50 000 g for 45 min at 4 °C and the cleared lysate was immediately applied onto a 5 ml HisTrap FF crude column (GE Healthcare) pre-equilibrated with HisTrap buffer A. Following the application of the lysate, the column was washed with five column volumes (CV) of HisTrap buffer A to remove unbound protein, followed by elution of the bound protein using an increasing linear gradient of HisTrap buffer B (20 mM Tris pH 8.0, 500 mM NaCl, 1000 mM imidazole, 1 mM DTT) over 20 CV. SDS–PAGE was used to identify fractions containing VC0615, which were pooled and concentrated to a volume of ~5 ml using a 30 kDa cutoff Vivaspin centrifugal concentrator (Millipore). This concentrated protein from HisTrap purification was purified by size-exclusion chromatography (SEC) using a S75 16/600 column (GE Healthcare) pre-equilibrated in SEC buffer (20 mM HEPS pH 7.4, 200 mM NaCl, 1 mM DTT). SEC fractions containing VC0615 were identified by SDS–PAGE, pooled and diluted 20-fold in ion-exchange buffer A (20 mM HEPS pH 7.4, 1 mM DTT). The diluted protein solution was applied onto a 5 ml Hitrap Q HP column (GE Healthcare) pre-equilibrated in IEX buffer A. The loaded column was washed with 5 CV IEX buffer A and eluted with a 20 CV increasing linear gradient of IEX buffer B (20 mM HEPS pH 7.4, 1500 mM NaCl, 1 mM DTT). SDS–PAGE identified the main IEX elution peak to contain pure VC0615, whilst a small shoulder peak eluting at higher NaCl concentrations contained contaminant proteins.

Fractions containing purified VC0615 were pooled and diluted with IEX buffer A to adjust the NaCl concentration to 100 mM. Purified protein samples were concentrated using a 30 kDa cutoff Vivaspin centrifugal concentrator and the final concentrations were calculated by UV absorbance at 280 nm (A280) using a calculated molar extinction coefficient of 129 720 M$^{-1}$ cm$^{-1}$ and a calculated molecular mass of 66 130.81 g mol$^{-1}$.

2.2. Kinetic assays

Enzyme kinetics were determined by measuring the hydrolysis of the artificial fluorogenic substrate 4-methylumbelliferyl-$\beta$-d-glucopyranoside (4MU-$\beta$-d-Glc; Sigma). In brief, 50 μl solutions of 4MU-$\beta$-d-Glc at 2× final concentration in 20 mM sodium phosphate pH 6.5 were prepared in a 96-well microplate (Nunc). To initiate the reaction, 50 μl of VC0615 at a concentration of 2 μM in 20 mM sodium phosphate pH 6.5 was added to each substrate solution, giving final reaction volumes of 100 μl containing 1× 4MU-$\beta$-d-Glc substrate and 1 μM VC0615. Reactions were monitored continuously for release of the fluorescent 4MU product over 600 s using a Polarstar microplate reader (BMG Labtech). The initial reaction rates plotted against the 4MU-$\beta$-d-Glc substrate concentration were fitted by nonlinear regression to the Michaelis–Menten equation $V = (V_{max}[4MU-\beta$-d-Glc]) / (K$\text{m} + [4MU-\beta$-d-Glc])$. $k_{\text{cat}}$ was determined from $V_{\text{max}}$ through the relationship $V_{\text{max}} = k_{\text{cat}}[E]$. All reactions were carried out in triplicate.

Reactions using other 4MU glycosides were carried out as above.

2.3. Thin-layer chromatography (TLC) analyses

All oligosaccharide substrates and standards for digests and TLC analyses were purchased from Megazyme, except for (GlcN)$_2$, which was purchased from Sigma. Digest reactions were carried out in McIlvaine phosphate–citrate buffer pH 6.5 using an oligosaccharide substrate concentration of 10 mg ml$^{-1}$ and 0.5 μM VC0615. 20 μl reaction volumes containing enzyme and substrate were incubated at 37 °C with shaking. At set time points, 2 μl of the reaction mixture was removed and ‘paused’ by flash-freezing in liquid N$_2$. Following completion of the reaction time course, all frozen aliquots were thawed and spotted onto an aluminium foil-backed silica TLC plate (Sigma).

The TLC plate was run in a 50:25:25 n-butanol:water:acetic acid solvent system until the solvent front reached ~1–2 cm from the top of the plate. In order to improve the separation between spots, the TLC plate was dried and rerun a second time in the same solvent system. Following the second run, the TLC plate was dried and developed using a p-anisaldehyde (Sigma) stain (3.7 ml p-anisaldehyde, 1.5 ml acetic acid, 5 ml concentrated sulfuric acid, 135 ml ethanol) with mild heating.

2.4. SEC-MALLS

SEC-MALLS experiments were run in 20 mM HEPS pH 7.4, 200 mM NaCl buffer. The injected sample comprised 100 μl VC0615 at 2.5 mg ml$^{-1}$ in 20 mM HEPS pH 7.4, 100 mM NaCl, 1 mM DTT. Experiments were conducted on a system comprising a Wyatt HELEOS II multi-angle light-scattering detector and a Wyatt REX refractive-index detector linked to a Shimadzu HPLC system (SPD-20A UV detector, LC20-AD isocratic pump system, DGU-20A3 degasser and SIL-20A autosampler). Work was conducted at room temperature (20 ± 2 °C). All solvents and buffers were 0.2 μm filtered before use and a further 0.1 μm filter was present in the flow path. The Shimadzu LC Solutions software was used to control the HPLC and ASTRA V software was used for the HELEOS II and REX detectors.

All data were analysed using the ASTRA V software. Molecular masses were estimated using the Zimm fit method with a degree of 1. A value of 0.18 ml g$^{-1}$ was used for the protein refractive-index increment (dn/dc).

2.5. Crystallization

Crystallization trials were carried out in sitting-drop vapour-diffusion format using 96-well Swissci SD-2 plates (‘MRC’ plates) against a range of commercial screens, with drops consisting of 120 nl protein and 120 nl reservoir solution. Initial trials with PEG- and/or salt-based crystallization screens using an ~30 mg ml$^{-1}$ stock of VC0615 protein did not yield any hits. Subsequently, the protein concentration was increased to 67.5 mg ml$^{-1}$ and a further range of commercial
screens were tested. Crystals were found in the PGA screen (Molecular Dimensions; Hu et al., 2008) under conditions containing 0.1 M Tris buffer pH 7.8 and a combination of PEG and PGA-LM precipitants. Optimization yielded crystals at basic pH values across a range of PEG and PGA-LM concentrations in both sitting-drop and hanging-drop vapour-diffusion formats. The final crystallization conditions were 0.1 M Tris pH 7.5, 5% PGA-LM, 8% PEG 20K in a hanging-drop vapour-diffusion format using 1 µl protein solution and 1 µl reservoir solution.

Substantial difficulty with cryoprotection was experienced during the crystal-harvesting process. Harvesting crystals into a ‘standard’ cryoprotectant consisting of the crystallization well solution (0.1 M Tris pH 7.5, 5% PGA-LM, 8% PEG 20K) supplemented with 25% ethylene glycol resulted in the crystals rapidly dissolving. Extensive experimentation found that the stable cryoprotection of crystals necessitated the use of high concentrations of PGA-LM, although the removal of PEG 20K did not adversely affect the stability or diffraction of the crystals. Thus, crystals were harvested into an optimized cryoprotectant solution consisting of 0.1 M Tris pH 7.5, 13% PGA-LM, 25% ethylene glycol before flash-cooling in liquid nitrogen for data collection. Crystallization information is summarized in Table 2.

2.6. Data collection and processing

Data were collected on beamline I03 at Diamond Light Source (DLS), UK. Images were indexed and integrated with XDS (Kabsch, 2010), followed by data reduction and scaling with AIMLESS (Evans & Murshudov, 2013). All calculations were carried out within the new CCP4i2 interface to the CCP4 software suite (Potterton et al., 2018). Data-collection statistics are summarized in Table 3.

3. Results and discussion

3.1. Biochemical characterization of recombinant VC0615

We first characterized the enzymatic activity of purified recombinant VC0615 using the fluorogenic substrates 4MU-β-glucose and 4MU-β-d-cellobiose. In line with previous analyses of this enzyme (Park et al., 2002), 4MU-β-d-glucose...
was hydrolysed by VC0615, with Michaelis–Menten kinetic parameters $K_m = 768 \, \mu M$ and $k_{cat} = 1.3 \, \text{min}^{-1}$ (Fig. 1a). Consistent with its reported activity as an exo-acting enzyme, we were unable to model Michaelis–Menten kinetics for the hydrolysis of 4MU-β-d-cellobiose by VC0615, as the release of 4MU was preceded by a lag period reflecting prior hydrolysis of a →2 glucose before the release of 4MU. No hydrolysis was observed against 4MU-β-d-xylose, 4MU-β-d-GlcNAc or 4MU-β-d-mannose substrates (Fig. 1b). Honda and coworkers have previously shown PBRA0520 from *P. profundum* (a close homologue of VC0615) to be a competent exo-glucosaminidase (Honda et al., 2011). However, we were unable to measure kinetic parameters for the hydrolysis of 4MU-β-d-glucosamine by VC0615 owing to a lack of commercial availability of this substrate.

In order to obtain a semi-quantitative measure of the efficiency of glucosidase versus glucosaminidase activity by VC0615, we next investigated the hydrolysis of various oligosaccharides by VC0615 using thin-layer chromatography (TLC). When cellobiose, cellohexaose or cellooligosaccharides were used as substrates for VC0615, we observed a gradual increase in glucose monomers and shorter cello-oligosaccharides in the reaction mixture. The cleavage pattern observed was consistent with the processive removal of single glucose units from cello-oligosaccharide chains, rather than internal chain hydrolysis by an endo-glucosidase (Fig. 2a).

When chitobiose (GlcN$_2$) was used as the substrate for VC0615, cleavage to the monosaccharide was also observed (Fig. 2b). Strikingly, a large amount of product was already visible after 5 min reaction time and VC0615 effected a complete digestion of the chitobiose substrate within 30 min. In contrast, a substantial amount of starting material remained in the cellobiose digest after 1 h (both reactions were initiated with 10 mg ml$^{-1}$ substrate). These experiments demonstrate that VC0615 is a substantially more effective exo-glucosaminidase than exo-glucosidase, which is in line with previous observations on PBRA0520 (Honda et al., 2011) and the lack of cellulose metabolism shown by Vibrionaceae in general. Consistent with the 4MU substrate experiments, no digestion of N,N'-diacetylchitobiose (GlcNAc)$_2$ was observed even after 1 h incubation (Fig. 2c).

3.2. Crystallization and crystal-packing interactions of VC0615

A first round of crystallization trials of VC0615 were carried out at ~30 mg ml$^{-1}$ protein concentration using a range of PEG and/or salt precipitant screens. We found no protein crystals in any of the conditions tested. Subsequently, the VC0615 concentration was increased to ~67.5 mg ml$^{-1}$ and additional crystallization screens were tested. In this second round of screening we observed several hits in the PGA screen (Molecular Dimensions) under conditions containing Tris

![Figure 1](image)

**Figure 1**
Hydrolysis of fluorogenic substrates by VC0615. (a) Michaelis–Menten kinetics for the hydrolysis of 4MU-β-d-glucose by VC0615. Error bars are standard deviations from three technical replicates. (b) Hydrolytic release of 4MU from various substrates. Only 4MU-β-d-glucose and 4MU-β-d-cellobiose show any hydrolysis. The release of 4MU from 4MU-β-d-cellobiose follows a lag period, consistent with processive hydrolysis by an exo-acting enzyme. (RFU, relative fluorescence units.)
buffer and a combination of PGA-LM and PEG precipitants (Fig. 3a). Optimization trials indicated that crystallization favoured basic pH values, with a range of PGA-LM and PEG concentrations being tolerated. Both sitting-drop and hanging-drop vapour-diffusion techniques yielded crystals of similar visual quality. The final crystallization conditions were 0.1 M Tris pH 7.5, 5% PGA-LM, 8% PEG 20K. Crystals typically appeared within several hours and reached maximum size within 2 d (Table 2).

Despite the apparent visual quality of the optimized VC0615 crystals, it became apparent that their suitability for X-ray diffraction studies was limited. Even after extensive optimization of the crystallization and cryoprotectant conditions, all crystals of VC0615 tested showed highly smeared diffraction patterns that were indicative of substantial internal disorder within the crystal. The most promising crystals of VC0615, cryoprotected using 13% PGA-LM and 25% ethylene glycol, were sent to beamline I03 at Diamond Light Source, UK for data collection. The synchrotron diffraction images were similar in appearance to those collected in-house (Fig. 3b), although the mosaicity of the data was somewhat lower than expected (0.26° for the best data set). The best data set for VC0615 was processed to 3.17 Å resolution and was used for all further calculations (Tables 3 and 4).

The diffraction of the VC0615 crystals indicated a trigonal space group with a single screw axis along the $a$ axis, matching either $P\bar{3}21$ or $P3_2121$. Matthews coefficient analysis suggested six or seven molecules of protein per asymmetric unit to be the most likely composition of the VC0615 crystals (with 30 and 40% probability and 53.3 and 45.5% solvent content, respectively), although five or eight molecules per asymmetric unit were also plausible (with 10 and 15% probability and 61.1 and 37.7% solvent content, respectively).

The VC0615 structure was solved by molecular replacement using the GH9 enzyme VP2484 from V. parahaemolyticus, which shares 68% identity with VC0615. The best molecular-replacement solution from Phaser indicated the correct space group to be $P3_2121$, with five molecules of VC0615 in the asymmetric unit. Surprisingly, when the same molecular replacement was conducted with MOLREP using default...
settings, only four molecules of VC0615 were placed. Closer inspection revealed that whilst electron density for four of the five chains of VC0615 was clear and well defined, the density for the fifth chain (hereafter referred to as chain E) was substantially poorer, bordering on uninterpretable in many regions. Consistent with this poorer electron density, the B factors for chain E were substantially higher than for the other four chains (hereafter referred to as chains A–D) (Fig. 4a).

To establish the degree of oligomerization for VC0615 in solution, we analysed purified protein by size-exclusion chromatography coupled to multi-angle laser light scattering (SEC-MALLS). Purified VC0615 eluted as a single major peak with a calculated molecular mass of ~118 kDa, indicating that the species in solution was likely to be a dimer (monomer mass of ~66 kDa; Fig. 4b). PISA analysis (Krissinel & Henrick, 2007) suggested stable dimers between molecules of chain A and chain D of an adjacent asymmetric unit, and also between chain B and chain C of an adjacent asymmetric unit. Calculated interface areas were 1294.9 Å² per molecule for the A–D dimer and 1289.6 Å² per molecule for the B–C dimer. Given the essentially identical calculated interface areas between the A–D and B–C dimers, as well as their strong similarity (r.m.s.d. of 0.32 Å over 1129 residues as calculated by SSM superposition in CCP4mg), this arrangement of VC0615 is likely to be the biologically relevant dimer present in solution (Fig. 4c).

Analysis of protein-packing interactions within the VC0615 crystal lattice may provide some explanation for the substantially poorer electron density observed for chain E. The crystals of VC0615 appeared to contain a large central channel parallel to the c axis at the intersection of four adjacent unit cells, which is bordered by molecules of chain E (Fig. 4d). Given the dimeric nature of VC0615, we speculated on the possibility of a sixth ‘chain F’ of VC0615 in crystallo, which might lie within this central channel as the dimer partner of chain E. Although some positive Fc – Fo difference density was observed in the central channel, this density was too poor to model even a small portion of VC0615. Thus, we examined the effect of adding ‘chain F’ to the VC0615 crystal structure by modelling its predicted position through the superposition of the A–D dimer onto chain E. The resulting hypothetical six-chain model revealed a large steric clash between ‘chain F’ and a symmetry equivalent along the c axis, demonstrating that an additional molecule of VC0615 cannot be readily accommodated in this crystal-packing arrangement (Fig. 4e). Solving the structure of VC0615 in the lower symmetry space group P32 did not produce a structure in which the ‘chain F’ molecules were well resolved, indicating that the observed symmetry clash was unlikely to result from a pseudosymmetry operator being mistaken for crystallographic symmetry in the P321 structure. Given that ‘chain F’ is the dimer partner of chain E, difficulty in packing this chain readily explains the disorder observed for molecules of chain E within the VC0615 crystal structure.

When the VC0615 structure was refined with only chains A–D in the asymmetric unit, the Rwork and Rfree factors were ~2–3% higher and substantial positive Fc – Fo difference density could be observed in the region corresponding to

Figure 3
Crystallization and diffraction of VC0615. (a) Bright-field and UV fluorescence images of VC0615 crystals from PGA screen well G8 (0.1 M Tris pH 7.8, 5% PGA-LM, 8% PEG 20K). (b) Diffraction image of a VC0615 crystal on beamline I03 at Diamond Light Source.
Figure 4
Subunit interactions and crystal packing of VC0615. (a) Asymmetric unit of VC0615 coloured by $B$ factor, showing the substantially higher disorder of chain $E$. The map shown is $F_o - F_c$ OMIT density contoured at 0.22 e Å$^{-3}$ calculated from refinement of VC0615 without chain $E$. (b) SEC-MALLS trace showing a single peak for purified VC0615, containing a species with an estimated molecular mass of $\sim$118 kDa, indicating that VC0615 forms a dimer in solution. (c) Superposition of chains $A$ and $D$ and chains $B$ and $C$ dimers of VC0615. (d) Packing of VC0615 crystals. A single asymmetric unit is highlighted in blue (chains $A$–$D$) and purple (chain $E$). Chains $A$–$D$ and chain $E$ in other asymmetric units are coloured grey and red, respectively. A central channel surrounded by molecules of chain $E$ can be observed at the intersection of four unit cells. (e) Steric clashes which arise between two symmetry-related molecules of a hypothetical ‘chain $F$’ (red and purple) in the VC0615 crystal lattice.
Thus, we elected to leave chain E modelled in the final deposited structure (PDB entry 6gdt).

3.3. The structure of VC0615 and its relationship to other GH9 enzymes

The structure of VC0615 is highly similar to those solved for GH9 enzymes from \textit{P. profundum} (PBPRA0520; PDB entry 5dgr; Honda \textit{et al.}, 2016) and \textit{V. parahaemolyticus} (VP2484; PDB entry 3h7l; New York SGX Research Center for Structural Genomics, unpublished work), sharing 60 and 68\% sequence identity, respectively. Superposition using the SSM superpose function of CCP4 gives r.m.s.d.s of 0.66 Å over 554 residues for VC0615 and PBPRA0520, and 0.59 Å over 563 residues for VC0615 and VP2484 (comparing chain A of each structure).

VC0615 protomers are comprised of two domains: an N-terminal fibronectin type III (Fn3) domain (amino acids 1–92) and a C-terminal \((\alpha/\alpha)_6\)-barrel domain (93–566) which contains the active site. The active site of VC0615 (position inferred from homology to a PBPRA0520–ligand complex; PDB entry 5dgr) is situated within a series of loops which emanate from the core helices of the \((\alpha/\alpha)_6\)-barrel (Fig. 5). Long linkers from \((\alpha/\alpha)_6\)-helices 1–2 (residues 112–162) and \((\alpha/\alpha)_6\)-helices 9–10 (404–442) of VC0615 (features that are well resolved at 3.17 Å resolution) contribute to the formation of a closed-off steric block towards the ‘rear’ face of the binding pocket, which restricts the active-site binding pocket to only a single sugar-binding subsite. These long loops distinguish VC0615 and related exo-acting GH9s from the endo-acting GH9s, such as Tf-Cel9A cellulase from \textit{Thermobifida fusca}, which has much shorter loops linking its \(\alpha\)-helices (PDB entry 4tf4; Sakon \textit{et al.}, 1997; Fig. 6).

Although not strictly an exo-acting enzyme, the GH9 cellobiohydrolase CbhA from \textit{C. thermocellum} (PDB entry 1rq5; Schubot \textit{et al.}, 2004) has been reported to contain a structurally closed binding pocket, albeit one which hydrolyses terminal cellobiose rather than terminal glucose units. Structural comparison of VC0615 and related exo-acting GH9s with CbhA shows that the linkers from \(\alpha\)-helices 1 to 2 and 9 to 10 contribute more to steric blocking of the −2 subsite in the true exo-acting enzymes (Fig. 7). Notably, whilst the loop from \(\alpha\)-helices 1 to 2 is actually longer in CbhA (117–183) compared with the exo-acting GH9s (112–162 in VC0615), the \(\alpha\)-helix 1–2 loop in exo-acting GH9s is projected strongly ‘inwards’ towards the active site compared with CbhA, thus occluding more of the enzyme binding pocket.

The active-site residues are nearly completely conserved between VC0615, PBPRA0520 and VP2484, with the only variation being the presence of Phe231 in PBPRA0520, compared with Tyr at the equivalent position in VC0615 and VP2484 (Tyr232 in VC0615). Although we attempted to obtain a ligand complex of VC0615 by soaking crystals with cellotriose or cellotetraose, these efforts did not produce a structure with interpretable active-site electron density. Thus, we examined the active-site interactions in VC0615 via comparison with a PBPRA0520–glucosamine ligand complex. Based on homology to PBPRA0520, Asp140 and Glu546 of VC0615 are postulated to act as the catalytic base and acid residues of VC0615, respectively, with Asp144 also likely to be important in coordinating the incoming water nucleophile. Asp140 and Glu546 are situated 8.4 Å apart, consistent with the typical separation for catalytic residues of an inverting glycosidase (Davies & Henrissat, 1995). Surprisingly, there appears to be no structural rationale within the VC0615 (or PBPRA0520) binding pocket for the observed preference for

![Figure 5](image)

\textit{Figure 5}
Structure of a VC0615 monomer. (a) Ribbon diagram of VC0615 (chain A), showing the Fn3 (pink) and \((\alpha/\alpha)_6\)-barrel (orange) domains. (b) Active-site residues of VC0615. Postulated acid and catalytic base residues are annotated.
chito-oligosaccharide over cellulosic substrates. A $-1$ subsite glucosamine would be expected to position its 2-amino group within hydrogen-bonding distance of Tyr148, Tyr232 and Trp220 of VC0615, none of which are obviously suited for discriminating the hydrogen-bonding pattern of an amino group over a hydroxyl and which are also incorrectly oriented (‘edge on’) to form a cation–π interaction with a charged NH$_3^+$.

It is possible that the observed preference of VC0615 for...
chito-oligosaccharides may arise from factors outside the $-1$ subsite binding pocket, such as interactions with the departing +1 sugar.

4. Conclusion

Here, we have reported the structure of the $V.~cholerae$ enzyme VC0615 at 3.17 Å resolution. The VC0615 crystal structure contains two relatively well ordered homodimers of VC0615 in each asymmetric unit (chains $A$ and $D$ and chains $B$ and $C$ in our structure), as well as a fifth highly disordered molecule of VC0615 (chain $E$) which appears to lack a dimeric partner. PISA analysis indicated that a third VC0615 homodimer (chains $E$ and $F$) cannot be easily accommodated by the VC0615 crystal lattice owing to steric clashes which would arise between symmetry-related copies of molecules of ‘chain
Given the intrinsic dimeric nature of VC0615, it is possible that ‘chain F’ molecules are present within the VC0615 crystal lattice, albeit not in a sufficiently ordered fashion to contribute to X-ray diffraction. Difficulty in packing molecules of ‘chain F’ within the VC0615 crystal lattice provides a clear explanation for the high degree of disorder observed in molecules of chain E, and may also be a substantial contributing factor to the poor diffraction quality of VC0615 crystals in general.

VC0615 forms part of an exo-acting glucosidase/glucosaminidase subgroup within the CAZy GH9 family, alongside previously reported structures of PRPBA0520 and VP2484. Structural comparison of the exo-acting GH9s with both an endo-acting GH9 cellulase and a GH9 cellobiohydrolase illustrate that the loop topologies around the catalytic (α/α)₉ domain are key for delineating substrate accessibility for all GH9 enzymes. In particular, loops from α-helices 1 to 2 and α-helices 9 to 10 cause steric blocks within the (α/α)₉ domain of exo-acting GH9s, which restrict substrate binding beyond the −1 subsite.

Some key questions remain regarding the interaction of exo-acting GH9s such as VC0615 with their substrates. In particular, no obvious basis for the discrimination of the −1 subsite for GlcN over Glc has yet been discerned in the structures of VC0615, PRPBA0520 or VP2484, implying the possibility of GlcN discrimination at the departing +1 subsite. This is supported by the observation by Honda and coworkers that whilst (GlcN)₂ is hydrolysed by PRPBA0520 with a $k_{cat}/K_m$ almost 15 times greater than for (Glc)₂, the $k_{cat}/K_m$ for the hydrolysis of the aryl substrate pNP-Glc is only approximately threefold greater than that for pNP-Glc (Honda et al., 2011). The hypothesis by Hunt and coworkers that VC0615 acts on a GlcN-GlcN-O6-P substrate in vivo (Hunt et al., 2008) also remains to be tested, especially with regard to the potential binding contributions made by an O6 phosphate at the +1 enzyme subsite. Improved structural and biochemical understanding of exo-acting GH9 enzymes will help to shed light on these questions, as well as broader questions regarding the molecular basis of chitin utilization by the Vibrionaceae.

Acknowledgements

We thank Diamond Light Source for access to beamline I03 (proposal mx-13587), which contributed to the results presented here.

Funding information

The following funding is acknowledged: European Research Council (grant No. ErC-2012-AdG329242 to Gideon J. Davies); the Royal Society (award to Gideon J. Davies).

References

Ali, M., Nelson, A. R., Lopez, A. L. & Sack, D. A. (2015). PLoS Negl. Trop. Dis. 9, e0003832.
Amako, K., Shimodori, S., Imoto, T., Miske, S. & Umeda, A. (1987). Appl. Environ. Microbiol. 53, 603–605.
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.
Davies, G. J., Wilson, K. S. & Henrissat, B. (1997). Biochem. J. 321, 557–559.
Emsley, P. & Cowtan, K. (2004). Acta Cryst. D60, 2126–2132.
Evans, P. R. & Murshudov, G. N. (2013). Acta Cryst. D69, 1204–1214.
Heidelberg, J. F., Heidelberg, K. B. & Colwell, R. R. (2002). Appl. Environ. Microbiol. 68, 5498–5507.
Honda, Y., Arai, S., Suzuki, K., Kataoka, M. & Fushinobu, S. (2016). Biochem. J. 473, 463–472.
Honda, Y., Shimaya, N., Iishiaki, K., Ebihara, M. & Taniguchi, H. (2011). Glycobiology, 21, 503–511.
Hu, T.-C., Korczyńska, J., Smith, D. K. & Brzuszkowski, A. M. (2008). Acta Cryst. D64, 957–963.
Hunt, D. E., Gevers, D., Vahora, N. M. & Polz, M. F. (2008). Appl. Environ. Microbiol. 74, 44–51.
Huq, A., Sack, R. B., Nizam, A., Longini, I. M., Nair, G. B., Ali, A., Morris, J. G. Jr, Khan, M. N., Siddique, A. K., Yunus, M., Albert, M. J., Sack, D. A. & Colwell, R. R. (2005). Appl. Environ. Microbiol. 71, 4645–4654.
Kabsch, W. (2010). Acta Cryst. D66, 125–132.
Krispinel, E. & Henrick, K. (2007). J. Mol. Biol. 372, 774–797.
McCoy, A. J. (2007). Acta Cryst. D63, 32–41.
McNicholas, S., Potterton, E., Wilson, K. S. & Noble, M. E. M. (2011). Acta Cryst. D67, 386–394.
Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). Acta Cryst. D67, 355–367.
Nalin, D. R., Daya, V., Reid, A., Levine, M. M. & Cisneros, L. (1979). Infect. Immun. 25, 768–770.
Nomenclature Committee of IUB and IUPAC–IUB Joint Commission on Biochemical Nomenclature (1988). Arch. Biochem. Biophys. 260, 851–854.
Park, J. K., Wang, L.-X., Patel, H. V. & Roseman, S. (2002). J. Biol. Chem. 277, 29555–29560.
Potterton, L. et al. (2018). Acta Cryst. D74, 68–84.
Sakon, J., Irwin, D., Wilson, D. B. & Karplus, P. A. (1997). Nature Struct. Biol. 4, 810–818.
Schubot, F. D., Kataeva, I. A., Chang, J., Shah, A. K., Ljungdahl, L. G., Rose, J. P. & Wang, B.-C. (2004). Biochemistry, 43, 1163–1170.
Sievers, F. & Higgins, D. G. (2018). Protein Sci. 27, 135–145.
The CAzYPDedia Consortium (2018). Glycobiology, 28, 3–8.
Vagin, A. & Teplyakov, A. (2010). Acta Cryst. D66, 22–25.

Acta Cryst. (2018). F74, 512–523

Wu & Davies • GH9 glucosidase/glucosaminidase