Structure of a carbohydrate esterase from Bacillus anthracis

ARTICLE *in* PROTEINS STRUCTURE FUNCTION AND BIOINFORMATICS · OCTOBER 2006

Impact Factor: 2.63 · DOI: 10.1002/prot.21217 · Source: PubMed

4 AUTHORS, INCLUDING:

Tracey M Gloster
University of St Andrews
49 PUBLICATIONS **1,685** CITATIONS

SEE PROFILE

PROTEINS: Structure, Function, and Bioinformatics 66:250–252 (2007)

STRUCTURE NOTE

Structure of a Carbohydrate Esterase from *Bacillus anthracis*

Leoni Oberbarnscheidt, Edward J. Taylor, Gideon J. Davies, and Tracey M. Gloster*
York Structural Biology Laboratory, University of York, York YO10 5YW, United Kingdom

Introduction. Family 4 carbohydrate esterases (CEs) catalyse the *N*- or *O*-deacetylation of substrates such as acetylated xylan, chitin, and peptidoglycan. CEs are classified into 14 families by sequence homology (see http://afmb.cnrs-mrs.fr/CAZY[1]). Family 4 is by far the largest of the CE families, with over 1000 open reading frames. The structure of CE4 enzymes from a number of bacterial species have been solved, including the peptidoglycan deacetylases from *Streptococcus pneumoniae*[2] and *Bacillus subtilis*,[3] acetyl xylan esterases from *Clostridium thermocellum* and *Streptomyces lividans*,[4] and an enzyme of unknown specificity from *Pseudomonas aeruginosa* (PDB code 1Z7A). CE4 enzymes contain a conserved NodB homology domain, and adopt a distorted $(\alpha/\beta)_8$ barrel fold. Most of the structures contain a divalent ion in the active site that is necessary for enzyme activity[2,4] and which is coordinated by highly conserved histidine and aspartate residues.

Here we report the structure of a putative peptidoglycan deacetylase CE4, ORF AAP24453 (hereafter *Ba*CE4), from *Bacillus anthracis*, the causative agent of anthrax. Peptidoglycan deacetylases catalyse the *N*-deacetylation of *N*-acetyl glucosamine and *N*-acetyl muramic acid moieties of bacterial peptidoglycan in the cell wall. This allows degradation and remodeling of the bacterial cell wall, which prevents recognition by mammalian host enzymes and is therefore an extremely useful defense mechanism for bacteria.[5] The closely related *B. subtilis* enzyme, PdaA, has been shown to deacetylate *N*-acetyl muramic acid residues from peptidoglycan in vitro,[6] and deletion in vivo causes spore germination to fail, as the muramic δ-lactam structure cannot be formed.[7] It is likely that *Ba*CE4 plays a similar role in vivo, and hence is important in the sporulation of *B. anthracis*. Examination of the genome sequence of *B. anthracis* has revealed the organism hosts 10 putative polysaccharide deacetylases, which are similarly likely to have roles in sporulation, spore germination, vegetative growth, and host–microbe interactions.[8] Inhibition of these enzymes is extremely attractive in the quest for preventing proliferation of these highly infectious organisms.

Materials and Methods. The open reading frame encoding the mature peptide of *Ba*CE4 was amplified from *Bacillus anthracis str. Ames* genomic DNA with ligation independent cloning compatible ends. The PCR product was annealed into a modified ligation-independent cloning pET28a vector. This plasmid was transformed into BL21 (DE3) *Escherichia coli* cells and cultured in 0.5 L autoinduction media[9] supplemented with 50 mg mL^{-1} kanamycin, at 37°C, for 8 h. Protein expression was induced overnight at 30°C. Cells were harvested and resuspended in 20 m*M* HEPES, pH 7, 150 m*M* NaCl, and subsequently sonicated to lyse the cells. The supernatant was applied to a 5 mL HisTrap column (GE Healthcare) where the protein was eluted with an imidazole gradient, and subsequently to a Superdex 200 16/60 gel filtration column. Pure protein was buffer exchanged into 20 m*M* HEPES, pH 7, and concentrated to 40 mg/mL for crystallization. *Ba*CE4 was crystallized from 0.1 *M* sodium cacodylate, pH 6.5, 0.2 *M* zinc acetate, and 18% (w/v) polyethylene glycol 8000 (Hampton screen I). Crystals were cryoprotected in the mother liquor solution with the addition of 25% glycerol and flash frozen.

Data were collected from a single crystal at 100 K on ID14-3 at the European Synchrotron Radiation Facility, Grenoble, to 1.7 Å resolution. Data were processed with HKL2000,[10] and all subsequent computing used the CCP4 suite of programs.[11] The *Ba*CE4 structure was solved by molecular replacement using AMoRe[12] with the protein atoms of a monomer from PDB entry 1W1B. Data were used between 10 and 4 Å, with an integration sphere of 25 Å. There is a single molecule in the asymmetric unit with a packing density of 2.6 $Å^3$ Da^{-1} and solvent content of ~52%. The structure was refined using iterative cycles of REFMAC[13] and model building/solvent addition with COOT[14] (statistics are shown in Table I).

Grant sponsor: Biotechnology and Biological Sciences Research Council (BBSRC).

*Correspondence to: Tracey M. Gloster, York Structural Biology Laboratory, University of York, York YO10 5YW, United Kingdom. E-mail: gloster@ysbl.york.ac.uk

Received 4 July 2006; Accepted 3 August 2006

Published online 24 October 2006 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.21217

TABLE I. Data Processing and Refinement Statistics for *Ba*CE4

	*Ba*CE4
Data collection beamline (ESRF)	ID14-3
Resolution (outer shell) (Å)	30–1.70 (1.76–1.70)
Space group	$P3_221$
R_{merge} (outer shell)	0.068 (0.322)
Mean $I/\sigma I$ (outer shell)	25.6 (4.2)
Completeness (outer shell) (%)	99.2 (95.7)
Multiplicity (outer shell)	5.0 (3.2)
Number of unique reflections	33502
R_{cryst}	0.19
R_{free}	0.23
RMSD bonds (Å)	0.014
RMSD angles (°)	1.54
RMSD chiral volume (Å^3)	0.19
Average main chain *B*-factor (Å^2)	31
Average side chain *B*-factor (Å^2)	33
Average water *B*-factor (Å^2)	45
PDB code	2J13

Results and Discussion. Data for *Ba*CE4 crystals were collected to 1.7 Å resolution, and the structure subsequently solved using molecular replacement methods with the *B. subtilis* enzyme PdaA, which has 56% sequence identity, as the search model. *Ba*CE4 is shown to adopt a distorted $(\alpha/\beta)_8$ barrel fold [Fig. 1(A)], as observed with other CE4 enzymes, with the active site lying in a groove. The majority of the 237-residue chain could be traced, but has loops missing between residues 13 and 22, and 174 and 187, and residues 236 and 237 at the *C*-terminus can also not be observed. Examination of the active site reveals a zinc ion coordinated to an acetate ion and a cacodylate ion [Fig. 1(B)], all of which have been sequestered from the crystallization mother liquor. Three other zinc ions and another acetate ion can also be observed in the electron density, which appear at crystal contacts and aid packing. The active site zinc ion coordinates five electronegative atoms with tetrahedral geometry. There is a bidentate, but asymmetric, interaction with both oxygen atoms of the acetate molecule; one of these interactions is at a distance of 2.5 Å, which is longer than usually observed with zinc ion coordination.[18] The nitrogen atoms of His103 and His107 coordinate the zinc ion at distances of 2.0 and 2.1 Å, respectively, an oxygen atom of the cacodylate ion at a distance of 1.9 Å and it is 2.2 Å away from the other oxygen atom of the acetate ion.

*Ba*CE4 lacks an otherwise conserved aspartate residue, which coordinates the metal ion in other CE4 structures. In its place, *Ba*CE4 possesses an asparagine residue which points away from the active site metal ion into the core of the protein. The bound acetate ion fills

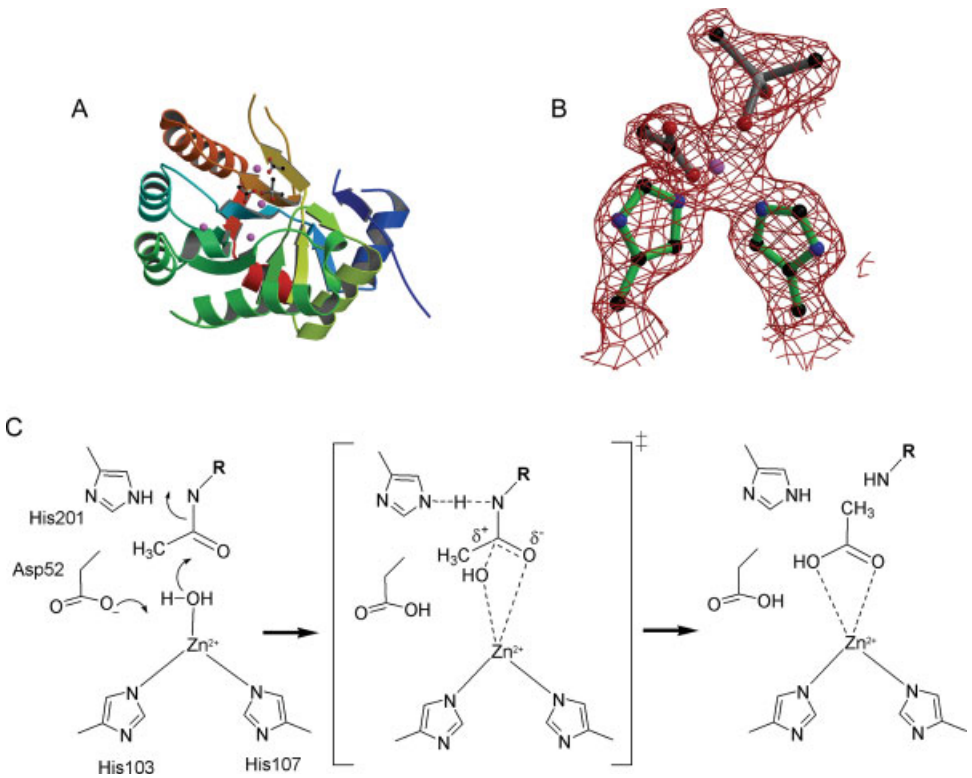

Fig. 1. (**A**) Ribbon representation of *Ba*CE4, color-ramped from *N*- (blue) to *C*- (red) terminus. Acetate and cacodylate ions are shown in ball-and-stick representation, and zinc ions are shown as spheres. (**B**) Active site of *Ba*CE4, which shows a zinc ion (sphere) coordinating a cacodylate ion and an acetate ion that were all sequestered from the crystallization mother liquor, and His103 and His107. Observed electron density for the maximum likelihood weighted $2F_{obs} - F_{calc}$ map is contoured at 1.5σ (0.36e Å^{-3}). (**C**) Proposed mechanism for CE4 polysaccharide deacetylation. Asp52 acts as a base by activating the nucleophilic water residue and His 201 acts as an acid to aid leaving group departure. The reaction passes through a tetrahedral transition state, which is mimicked here by the cacodylate ion. Figures for (A) and (B) were drawn using MOLSCRIPT[15] and BOBSCRIPT[16] and rendered using RASTER3D.[17]

the void left by this aspartate residue; the acetate overlaps with the carboxylate group of the aspartate residue observed in other structures and makes similar interactions with the metal ion. An asparagine residue is also seen in this sequence position in the closely related PdaA structure,[3] and is also observed to point away from the active site. Alignments with a number of CE4 sequences indicate that although aspartate is common at this position, asparagine, as well as hydrophobic residues such as alanine and valine are also observed, and thus is not strictly necessary for a "conserved metal ion binding triad" as had been suggested by other alignments.[2,19] The structural studies of CE4 enzymes with different residues at this position may provide information about substrate specificity (e.g., chitin vs. acetylated xylan vs. peptidoglycan). It is highly likely that *Ba*CE4 deacetylates *N*-acetyl muramic acid and hence may utilize the extra space provided by the lack of an aspartate residue to bind the bulky group at the O3 position of the muramic acid. It has indeed been suggested that the carboxylate on the *O*-lactoyl group of the substrate may interact with the metal ion.[2,3]

This study supports the proposed acid/base catalytic mechanism for CE4 enzymes [Fig. 1(C)], which has been deduced by biochemical, structural, and mutagenesis data.[2,4,20] Asp52 would act as a base to activate a water molecule, which would be stabilized by the zinc ion. The resulting hydroxide ion would act as a nucleophile to attack the acetate. This would be concomitant with leaving group departure, which may receive general acid assistance from His201. The reaction would pass through a tetrahedral geometry transition state. The binding of the cacodylate ion is positioned where the acetate attached to the sugar would bind during catalysis, but its tetrahedral geometry mimics that of the reaction transition state. Previous complexes with a sulphate ion (with similar tetrahedral geometry)[2] and acetate have also been observed in this position on other systems.[4] It is significant that the cacodylate ion rather than an acetate ion is bound here, and confirms that the transition state mimicking tetrahedral geometry has a higher affinity than the product-like acetate ion. It is interesting to note that in most CE4 structures, including those with acetate bound, the metal ion has octahedral coordination,[2,4] despite tetrahedral coordination being more common for catalytic enzymes.[20] However, it is interesting to note that in the sulphate complex with the *S. pneumoniae* enzyme,[2] and in the cacodylate complex described here, which both mimic the geometry at the transition state, the metal ion has tetrahedral coordination. These differences in metal ion coordination may reflect changes that occur to the coordination during catalysis.

The structure of *Ba*CE4 highlights both similarities and important differences between related CE4 enzymes, which need further investigation with biochemical, mutagenesis, and structural studies in order to ascertain the requirements for substrate specificity. It seems apparent that the residues vital for catalysis, such as Asp52 and His201 in *Ba*CE4, are highly conserved, whereas those that bind the metal ion can tolerate a number of different residues other than the "His-His-Asp" triad that was once predicted. These enzymes play a vital role in bacteria cell wall maintenance, and the lack of such enzymes in humans could be exploited by the drug design process to inhibit their proliferation.

REFERENCES

1. Coutinho PM, Henrissat B. Carbohydrate-active enzymes: an integrated database approach. In: Gilbert HJ, Davies GJ, Henrissat B, Svensson B, editors. Recent advances in carbohydrate bioengineering. Cambridge: Royal Society of Chemistry; 1999. pp 3–12.
2. Blair DE, Schüttelkopf AW, MacRae JI, van Aalten DMF. Structure and metal-dependent mechanism of peptidoglycan deacetylase, a streptococcal virulence factor. Proc Natl Acad Sci USA 2005;102:15429–15434.
3. Blair DE, van Aalten DMF. Structures of *Bacillus subtilis* PdaA, a family 4 carbohydrate esterase, and a complex with *N*-acetyl-glucosamine. FEBS Lett 2004;570:13–19.
4. Taylor EJ, Gloster TM, Turkenburg JP, Vincent F, Brzozowski AM, Dupont C, Shareck F, Centeno MS, Prates JA, Puchart V, Ferreira LM, Fontes CM, Biely P, Davies GJ. Structure and activity of two metal ion-dependent acetylxylan esterases involved in plant cell wall degradation reveals a close similarity to peptidoglycan deacetylases. J Biol Chem 2006;281:10968–10975.
5. Boneca IG. The role of peptidoglycan in pathogenesis. Curr Opin Microbiol 2005;8:46–53.
6. Fukushima T, Kitajima T, Sekiguchi J. A polysaccharide deacetylase homologue, PdaA, in *Bacillus subtilis* acts as an *N*-acetylmuramic acid deacetylase in vitro. J Bacteriol 2005;187:1287–1292.
7. Fukushima T, Yamamoto H, Atrih A, Foster SJ, Sekiguchi J. A polysaccharide deacetylase gene (*pdaA*) is required for germination and for production of muramic δ-lactam residues in the spore cortex of *Bacillus subtilis*. J Bacteriol 2002;184:6007–6015.
8. Psylinakis E, Boneca IG, Mavromatis K, Deli A, Hayhurst E, Foster SJ, Vårum KM, Bouriotis V. Peptidoglycan *N*-acetylglucosamine deacetylases from *Bacillus cereus*, highly conserved proteins in *Bacillus anthracis*. J Biol Chem 2005;280:30856–30863.
9. Studier FW. Protein production by auto-induction in high density shaking cultures. Protein Expr Purif 2005;41:207–234.
10. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol 1997;276:307–326.
11. Collaborative Computational Project Number 4. The CCP4 suite: programs for protein crystallography. Acta Crystallogr D Biol Crystallogr 1994;50(Part 5):760–763.
12. Navaza J. *AMoRe*: An automated package for molecular replacement. Acta Crystallogr A Found Crystallogr 1994;50:157–163.
13. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr 1997;53(Part 3):240–255.
14. Emsley P, Cowtan K. *Coot*: Model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr 2004;60:2126–2132.
15. Kraulis PJ. MOLSCRIPT A program to produce both detailed and schematic plots of protein structures. J Appl Cryst 1991;24: 946–950.
16. Esnouf RM. An extensively modified version of MolScript that includes greatly enhanced coloring capabilities. J Mol Graph Model 1997;15:132–134.
17. Merritt EA, Murphy MEP. Raster3D version 20 A program for photorealistic molecular graphics. Acta Crystallogr D Biol Crystallogr 1994;50:869–873.
18. Harding MM. The geometry of metal-ligand interactions relevant to proteins. Acta Crystallogr D Biol Crystallogr 1999;55: 1432–1443.
19. Puchart V, Gariépy M-C, Shareck F, Dupont C. Identification of catalytically important amino acid residues of *Streptomyces lividans* acetylxylan esterase A from carbohydrate esterase family 4. Biochim Biophys Acta 2006;1764:263–274.
20. Hernick M, Fierke CA. Zinc hydrolases: the mechanisms of zinc-dependent deacetylases. Arch Biochem Biophys 2005;433: 71–84.