Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization of GcnA, an *N*-acetyl-*β*-D-glucosaminidase, from *Streptococcus gordonii*

David B. Langley,[a] Derek W. S. Harty,[b] Stephen C. Graham,[a] J. Mitchell Guss,[a] Neil Hunter[b] and Charles Collyer[a]*

[a]School of Molecular and Microbial Biosciences, University of Sydney, NSW 2006, Australia, and [b]Institute of Dental Research, University of Sydney, NSW 2031, Australia

Correspondence e-mail: c.collyer@staff.usyd.edu.au

© 2004 International Union of Crystallography
Printed in Denmark – all rights reserved

Received 19 July 2004
Accepted 9 August 2004

Streptococcus gordonii is a primary colonizer of the surface of human teeth. The *gcnA* gene is one of a number of genes involved in glycoside metabolism. GcnA has *N*-acetyl-*β*-D-glucosaminidase (EC 3.2.1.52) activity; it has been overexpressed, purified and crystallized. Diffraction has been observed to beyond 1.5 Å resolution and synchrotron data have been recorded to 1.55 Å resolution. The crystals belong to the orthorhombic space group $P2_12_12$, with unit-cell parameters a = 112.5, b = 104.0, c = 110.0 Å. The crystals contain either a monomer or a dimer in the asymmetric unit.

1. Introduction

Streptococcus gordonii, a primary coloniser of the surface of human teeth, is generally associated with oral health (Nyvad & Kilian, 1990). *S. gordonii* is also one of the major causative organisms of the disease infective endocarditis (IE) where, in susceptible individuals, the bacterium is able to enter the bloodstream from the oral cavity and colonize damaged heart valves. It is not known how *S. gordonii* obtains its nutritional requirements when infecting the heart valve, but like other invasive streptococci it possesses a range of glycosidases which may be able to degrade mammalian glycoproteins as a source of carbohydrate for growth (Ford & Douglas, 1997; Gong *et al.*, 1998; Herzberg *et al.*, 1992; Herzberg, 1996).

Using an *in vivo* expression technology *S. gordonii* library, 13 genes involved in *β*-glycoside metabolism were shown by Kilic *et al.* (2004) to be expressed *in vivo* on infected heart valves in the rabbit model of IE, while none of these genes were expressed *in vitro* under laboratory conditions. One gene, *bhsA*, described as part of the *gom* locus, was identified from *S. gordonii* genomic data (http://www.tigr.org) as an *N*-acetyl-*β*-D-hexosaminidase encoding a 72 kDa protein (Kilic *et al.*, 2004). We have also cloned, sequenced and characterized the same gene from *S. gordonii* FSS2 (designated by us as *gcnA*) which codes for an *N*-acetyl-*β*-D-glucosaminidase (EC 3.2.1.52; Harty *et al.*, 2004). The protein, GcnA, is composed of 72 120 Da polypeptides of 627 amino acids and is homodimeric in solution as revealed by gel-filtration chromatography and SDS–PAGE analysis. In this report, we describe the crystallization and preliminary analysis of GcnA.

2. Expression and purification

GcnA was overexpressed in *Escherichia coli* from the pUC19-based construct pHAR101, as described previously (Harty *et al.*, 2004). Briefly, in order to reduce the proteolytic digestion of the expressed protein, the overexpressing host was cotransformed with the plasmid pREP4 (Qiagen) which conferred tighter control of expression owing to the overproduction of *Lac*I repressor. GcnA expression was induced *via* addition of isopropyl-*β*-D-thiogalactopyranoside (IPTG). A cocktail of protease inhibitors was added during induction, growth, harvest and also during purification by anion-exchange and gel-filtration chromatography to limit proteolytic digestion of the expressed protein. The purity of the expressed protein was monitored by SDS–PAGE and a sample of approximately 95% purity was used for crystallization.

3. Crystallization

Crystals were grown in hanging drops by vapour diffusion at room temperature. The 3 µl drops contained equal volumes of protein (10 mg ml^{-1} in 20 m*M* Tris pH 7.4) and well solution [200 m*M* ammonium acetate, 100 m*M* sodium citrate pH 6.0, 24%(*w*/*v*) PEG 4000]; the condition was optimized from Hampton Screen I condition No. 9 [200 m*M* ammonium acetate, 100 m*M* sodium citrate pH 5.6, 30%(*w*/*v*) PEG 4000]. The crystals, which grew over the course of several days, resemble craggy keys in morphology (Fig. 1), but despite their appearance give X-ray diffraction patterns indicative of single crystals. Crystals were cryoprotected prior to flash-freezing by brief exposure (~10 s) to well solution with

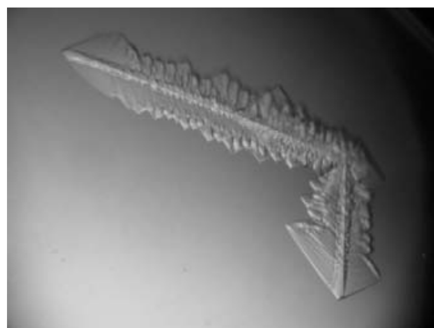

Figure 1
Crystals of GcnA. The picture shows three single crystals. Two lie at right angles and a third smaller crystal is lying on the top of the larger crystal near the 'elbow'. The largest of the three crystals has approximate dimensions of 1 × 0.2 × 0.03 mm.

2-methyl-2,4-pentanediol added to a final concentration of 15%(*v*/*v*).

4. Data collection

Crystals were flash-frozen in a stream of nitrogen gas at 100 K (Oxford Cryosystems) in Sydney prior to shipment to Lund. Data were recorded at beamline 7-11 at MAX-Lab, Lund, Sweden on a MAR 165 mm CCD detector (MAR Research) using radiation of wavelength 1.008 Å (Fig. 2). The diffraction data were recorded in two passes in order to resolve both the closely spaced intense low-resolution reflections as well as the weaker data at higher angles. The data from the low-resolution and high-resolution passes were integrated with *AUTOMAR* (Bartels & Klein, 2003) and *MOSFLM* (Leslie, 1992), respectively. The integrated data were subsequently averaged and merged with *SCALA* (Evans, 1993) (Table 1).

Systematic absences of axial reflections uniquely assign the crystals to the orthorhombic space group $P2_12_12$, with unit-cell parameters a = 112.5, b = 104.0, c = 110.0 Å. The calculated Matthews coefficients (Matthews, 1968) for one and two molecules per asymmetric unit are 4.56 and 2.28 Å^3 Da^{-1}, corresponding to solvent contents of 73 and 46%, respectively. While it is most likely that the asymmetric unit contains a dimer of GcnA, a self-rotation function calculated with the data to 4 Å resolution has no strong peaks consistent with the presence of a non-crystallographic twofold axis. If monomers are related by a non-crystallographic twofold axis of symmetry, it is likely that it lies parallel to one of the crystallographic axes. Since there are no known structural homologues of GcnA, the structure solution will be attempted by conventional isomorphous replacement methods.

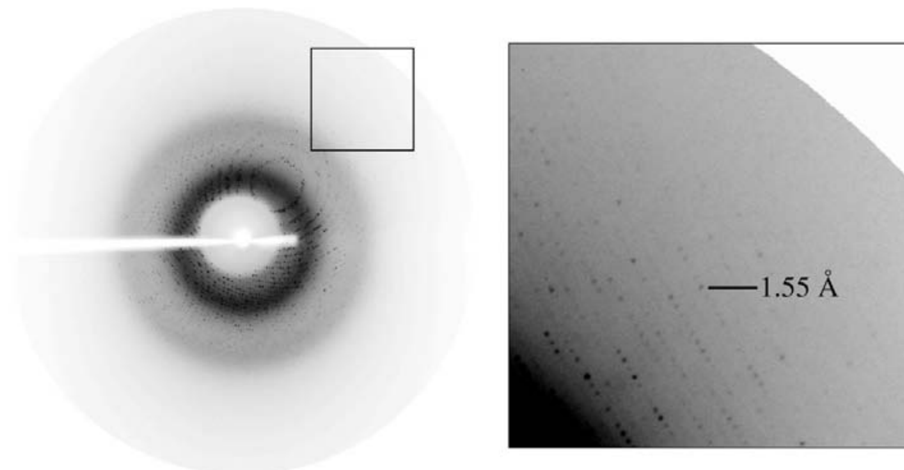

Figure 2
0.25° oscillation image of GcnA. The inset shows diffraction beyond 1.55 Å resolution. Owing to radiation-induced crystal damage, data were integrated to the 1.55 Å resolution limit.

Table 1
Data-collection statistics for GcnA.

Values in parentheses are for data in the highest resolution shell.

Space group	$P2_12_12$		
Unit-cell parameters (Å)	a = 112.5, b = 104.0, c = 110.0		
Temperature (K)	100		
Wavelength (Å)	1.008		
Oscillation angle (°)	0.5	0.25	
Crystal-to-detector distance (mm)	250	70	
Resolution limits (Å)	35.0–3.525	4.5–1.55	35.0–1.55
Exposure time per image (s)	7	20	
Total No. observations	62223	596028	187477
Unique reflections	15249	172228	179604
Completeness (%)	93.1 (95.9)	95.0 (100.0)	97.0 (100.0)
$I/\sigma(I)$	16.3 (11.9)	3.5 (1.6)	4.9
$R_{\rm merge}$†	0.041 (0.059)	0.093 (0.433)	0.103

† $R_{\rm merge} = \sum_{hkl}\sum_j |I(hkl)_j - \langle I(hkl)\rangle| / \sum_{hkl}\sum_j I(hkl)_j$.

We wish to thank Dr Janet Deane for transporting the crystals to Sweden.

References

Bartels, K. & Klein, C. (2003). *The AUTOMAR User's Guide*, v.1.04. MAR Research GmbH.

Evans, P. R. (1993). *Proceedings of the CCP4 Study Weekend. Data Collection and Processing*, edited by L. Sawyer, N. Isaacs & S. Bailey, pp. 114–122. Warrington: Daresbury Laboratory.

Ford, I. & Douglas, C. W. I. (1997). *Platelets*, **8**, 285–294.

Gong, K., Ouyang, T. & Herzberg, M. C. (1998). *Infect. Immun.* **66**, 5388–5392.

Harty, D. W., Chen, Y., Simpson, C. L., Berg, T., Cook, S. L., Mayo, J. A., Hunter, N. & Jacques, N. A. (2004). *Biochem. Biophys. Res. Commun.* **319**, 439–447.

Herzberg, M. C. (1996). *Crit. Rev. Oral Biol. Med.* **7**, 222–236.

Herzberg, M. C., MacFarlane, G. D., Gong, K., Armstrong, N. N., Witt, A. R., Erikson, P. R. & Meyer, M. W. (1992). *Infect. Immun.* **60**, 4809–4818.

Kilic, A. O., Tao, L., Zhang, Y., Lei, Y., Khammanivong, A. & Herzberg, M. C. (2004). *J. Bacteriol.* **186**, 4246–4253.

Leslie, A. G. W. (1992). *Jnt CCP4/ESF–EAMCB Newsl. Protein Crystallogr.* **26**.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

Nyvad, B. & Kilian, M. (1990). *Caries Res.* **24**, 267–272.