Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Crystallization and preliminary crystallographic characterization of GumK, a membrane-associated glucuronosyltransferase from *Xanthomonas campestris* required for xanthan polysaccharide synthesis

Máximo Barreras,[a] Mario A. Bianchet[b] and Luis Ielpi[a]*

[a]Fundación Instituto Leloir, CONICET, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, C1405BWE Buenos Aires, Argentina, and [b]Department of Biophysics and Biophysical Chemistry, Johns Hopkins University, Baltimore, Maryland 21205, USA

Correspondence e-mail: lielpi@leloir.org.ar

Received 6 July 2006
Accepted 24 July 2006

GumK is a membrane-associated inverting glucuronosyltransferase that is part of the biosynthetic route of xanthan, an industrially important exopolysaccharide produced by *Xanthomonas campestris*. The enzyme catalyzes the fourth glycosylation step in the pentasaccharide-P-P-polyisoprenyl assembly, an oligosaccharide diphosphate lipid intermediate in xanthan biosynthesis. GumK has marginal homology to other glycosyltransferases (GTs). It belongs to the CAZy family GT 70, for which no structure is currently available, and indirect biochemical evidence suggests that it also belongs to the GT-B structural superfamily. Crystals of recombinant GumK from *X. campestris* have been grown that diffract to 1.9 Å resolution. Knowledge of the crystal structure of GumK will help in understanding xanthan biosynthesis and its regulation and will also allow a subsequent rational approach to enzyme design and engineering. The multiwavelength anomalous diffraction approach will be used to solve the phase problem.

1. Introduction

Extracellular polysaccharides are major secreted products in many bacteria. There is considerable interest in the molecular aspects of these compounds as they are involved in human and plant pathogenicity. Additionally, many of them display unique physical properties that are useful for industrial applications (Sutherland & Tait, 1992). Most exopolysaccharides consist of polymerized oligosaccharide repeating units. Each repeating unit is assembled as glycoside-P-P-polyisoprenyl in a sequential series of reactions performed by specific glycosyltransferases (GTs).

The plant pathogen *Xanthomonas campestris* produces an exopolysaccharide termed xanthan gum. Xanthan is widely used as thickener or viscosifier in both the food and non-food industries, among many other applications (Sutherland, 1994; Becker *et al.*, 1998). The structure of xanthan consists of a β-1,4-linked D-glucose backbone with trisaccharide side chains composed of mannose-β-1,4-glucuronic acid-β-1,2-mannose attached to alternate backbone glucoses by α-1,3 linkages (Jansson *et al.*, 1975). The biosynthesis of xanthan consists of the stepwise assembly of pentasaccharide units attached to a polyisoprenyl phosphate carrier, which are subsequently polymerized and exported (Ielpi *et al.*, 1993). The transfer of each sugar residue to build up the pentasaccharide diphosphate polyisoprenyl is catalyzed by specific GTs (Katzen *et al.*, 1998). In particular, the transfer of the glucuronic acid (GlcA) residue is carried out by the glucuronosyltransferase GumK (Barreras *et al.*, 2004).

The molecular structure of GumK is not known; therefore, functional knowledge has been derived from biochemical characterization. Previous results from our laboratory showed that GumK (MW 44.4 kDa) is an inverting glucuronosyltransferase that catalyses the

fourth glycosylation step in the assembly of the pentasaccharide diphosphate lipid. The enzyme transfers a glucuronic acid residue from uridine diphosphoglucuronic acid (UDP-GlcA) to mannose-α-1,3-glucose-β-1,4-glucose-P-P-polyisoprenyl to form glucuronic acid-β-1,2-mannose-α-1,3-glucose-β-1,4-glucose-P-P-polyisoprenyl. GumK was unable to use the trisaccharide acceptor freed from the pyrophosphate lipid moiety. Nonetheless, replacement of the natural lipid moiety by phytanyl renders an acceptor substrate recognized by the enzyme (Barreras *et al.*, 2004), a feature that has also been described for other GTs (Chen *et al.*, 2002). Finally, it was also shown by differential centrifugation and immunodetection that this protein is located in the membrane fraction as a peripheral protein in *X. campestris* as well as in the expression host *Escherichia coli* (Barreras *et al.*, 2004).

To our knowledge no further data on crystallographic characterization of xanthan specific GTs is available. The determination of the molecular structure of GumK will help us understand the catalytic mechanism of this and related enzymes and how the attachment of the membrane-bound glycolipid acceptor and soluble nucleotide-sugar donor substrates proceeds. In addition, it will shed light on how the catalytic efficiency might be improved and/or how to engineer the enzyme to use other substrates.

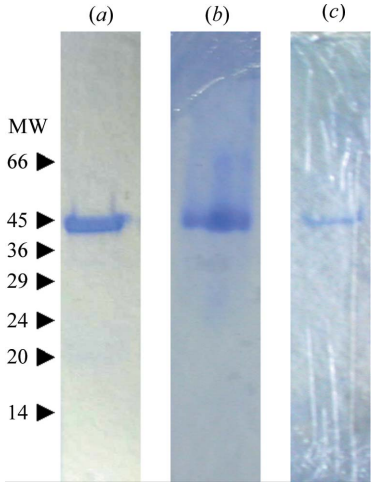

Figure 1
Purification of GumK. Coomassie blue-stained 10% SDS–PAGE gels showing (*a*) GumK, (*b*) selenomethionine-substituted GumK, (*c*) dissolved GumK crystal. MW indicates molecular-weight markers, the sizes of which are indicated in kDa

2. Results and discussion

2.1. Expression and purification

Recombinant *X. campestris* GumK containing a C-terminal His tag was overexpressed in *E. coli* BL21(DE3) cells transformed with pET22HisKC (Barreras *et al.*, 2004). *E. coli* cells were grown in 1500 ml LB medium supplemented with 200 µg ml^{-1} ampicillin. When the culture reached an A_{600} value of 0.8, GumK expression was induced by adding 0.5 m*M* isopropyl β-thiogalactopyranoside. After 15–18 h in a shaker at 200 rev min^{-1} and 294 K, cells were harvested, washed twice in 70 m*M* Tris–HCl pH 8.2, resuspended in the same buffer at 50 OD equivalents per millilitre and disrupted by two passages in a French pressure cell at 138 MPa and 277 K. Disrupted cells were diluted with 70 m*M* Tris–HCl pH 8.2, 1%(*v*/*v*) Triton X-100 buffer to a protein concentration of 5 mg ml^{-1}, incubated with gentle agitation at 277 K for 2 h and centrifuged at 100 000*g* for 1 h at 277 K. Up to 50 mg of supernatant total protein was applied onto a 5 ml Ni–NTA column (GE Healthcare) equilibrated with 0.5 *M* NaCl, 0.05%(*v*/*v*) Triton X-100, 20 m*M* Tris–HCl, 20 m*M* imidazole pH 8.0 (column buffer) at 0.5 ml min^{-1}. The column was washed with column buffer until no absorbance at 280 nm was detected and elution was performed with a linear gradient of 20–400 m*M* imidazole in 60 ml column buffer at 1.0 ml min^{-1}. Eluted protein was concentrated to 20 mg ml^{-1} by ultrafiltration at 2000*g* using a Centricon UF-15 (Millipore). 10 mg aliquots were applied onto a 10 000–600 000 range Superdex 200 size-exclusion column (GE Healthcare) and eluted with 0.4 *M* NaCl, 50 m*M* Tris–HCl, 0.05%(*v*/*v*) Triton X-100 pH 8.0 (buffer *A*). The enzyme was eluted as a single peak and displayed a single protein band in 10% SDS–PAGE (Fig. 1*a*). The yield of enzyme was approximately 20 mg per litre of bacterial culture.

2.2. Crystallization

Initial crystallogenesis conditions were searched for using the hanging-drop vapour-diffusion method (Hampel *et al.*, 1968) in Linbro plates using Hampton Research Crystal Screen kits (Hampton Research). Screening was carried out using 500 µl reservoir volumes. GumK was concentrated to 20–23 mg ml^{-1} in buffer *A* using a Centricon UF-15. Crystals were obtained by mixing 2 µl protein with 2 µl crystallization buffer consisting of 30%(*w*/*v*) PEG 4000, 0.1 *M* Tris–HCl pH 8.5 plus (i) 0.2 *M* Li_2SO_4, (ii) 0.2 *M* $MgCl_2$ or (iii) 0.2 *M* sodium acetate. The crystals obtained (Fig. 2*a*) showed a rhomboidal shape, growing to final dimensions of 0.2 × 0.1 mm in

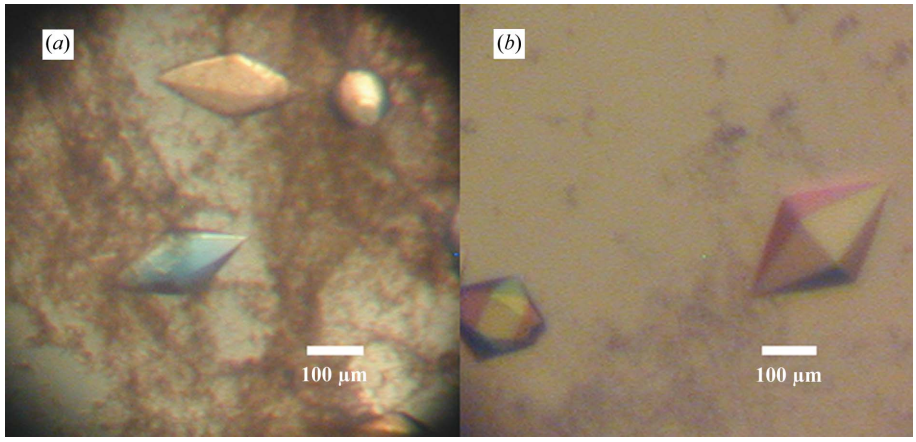

Figure 2
Crystals of *X. campestris* GumK obtained (*a*) from initial screening [30%(*w*/*v*) PEG 4000, 0.1 *M* Tris–HCl, 0.2 *M* Li_2SO_4 pH 8.5] and (*b*) from optimized conditions [35%(*w*/*v*)PEG 3350, 0.1 *M* Tris–HCl, 0.2 *M* Li_2SO_4, 0.1 *M* CsCl pH 8.2].

three weeks at 293 K. Optimization was achieved by fine screening around the initial conditions, varying the PEG or salt (Li_2SO_4, $MgCl_2$ or sodium acetate) concentrations. The pH was screened using Tris–HCl and phosphate buffers in the pH range 6–9. Further optimization was also achieved with the use of Hampton Research Additive and Detergent Screens 1, 2 and 3. The best shaped crystals (Fig. 2*b*) were obtained mixing 2 µl of protein with 1 µl 35%(*w*/*v*) PEG 3350, 0.1 *M* Tris–HCl, 0.2 *M* Li_2SO_4, 0.1 *M* CsCl pH 8.2 (optimized crystallization buffer). This produced crystals of 0.3 × 0.15 mm, which polarized light very strongly. Prior to data collection, crystals were transferred to 35%(*w*/*v*) PEG 3350, 0.1 *M* Tris–HCl, 0.2 *M* Li_2SO_4 pH 8.2 (cryoprotectant solution) for 1 min and flash-frozen in liquid nitrogen.

Table 1
Data-collection statistics for native crystal.

Values in parentheses are for the highest resolution shell.

Space group	$P6_522$
Crystal system	Primitive hexagonal
Unit-cell parameters (Å, °)	a = 123.63, b = 123.63, c = 174.30, α = 90, β = 90, γ = 120
Resolution range (Å)	35.5–1.9 (2.0–1.9)
Unique reflections	62401
R_{sym}† (%)	6.9 (32.3)
Completeness (%)	100 (100)
$\langle I/\sigma(I)\rangle$	9.9 (2.2)
Mosaicity (°)	0.18
Total observations	879919
Multiplicity	14.1

† $R_{sym} = \sum_{hkl}\sum_i |I_{hkl,i} - \langle I_{hkl}\rangle| / \sum_{hkl}\sum_i I_{hkl}$.

2.3. Preliminary crystallographic characterization

X-ray diffraction data sets were collected from single crystals (Figs. 3*a* and 3*b*) at 110 K at NSLS, Brookhaven National Laboratories, New York, USA on beamline X12C (λ = 1.100 Å) equipped with an ADSC Q210 detector. Crystals diffracted to 1.9 Å. Data were processed (Table 1) using the programs *MOSFLM*, *SCALA* and *TRUNCATE* from the *CCP*4 program suite (Collaborative Computational Project, Number 4, 1994). Crystals proved to have $P6_522$ symmetry, with a unit-cell volume of 2 307 219.25 Å^3. One molecule is present per asymmetric unit, resulting in a solvent content of 58.8%. No significant non-origin peaks were detected in the native Patterson map. Initial molecular-replacement calculations using low-homology models such as glycogen synthase (PDB code 1rzu; Buschiazzo *et al.*, 2004) or *N*-acetylglucosaminyl transferase MurG (PDB code 1f0k; Ha *et al.*, 2000) proved unsuccessful. Selenomethionine protein was produced in *E. coli* auxotrophic mutant B834(DE3). Cells were grown in 500 ml DLM medium (Coligan *et al.*, 2004) supplemented with 200 µg ml^{-1} ampicillin, 5 mg l^{-1} thiamine and 4 mg l^{-1} D-biotin. Protein expression and purification was carried essentially as described for native GumK, adding 5 m*M* tris (2-carboxyethyl)-phosphine hydrochloride to the column buffer. A single protein band was observed in 10% SDS–PAGE (Fig. 1*b*) and the yield of enzyme was approximately 12 mg per litre of bacterial culture. So far, all

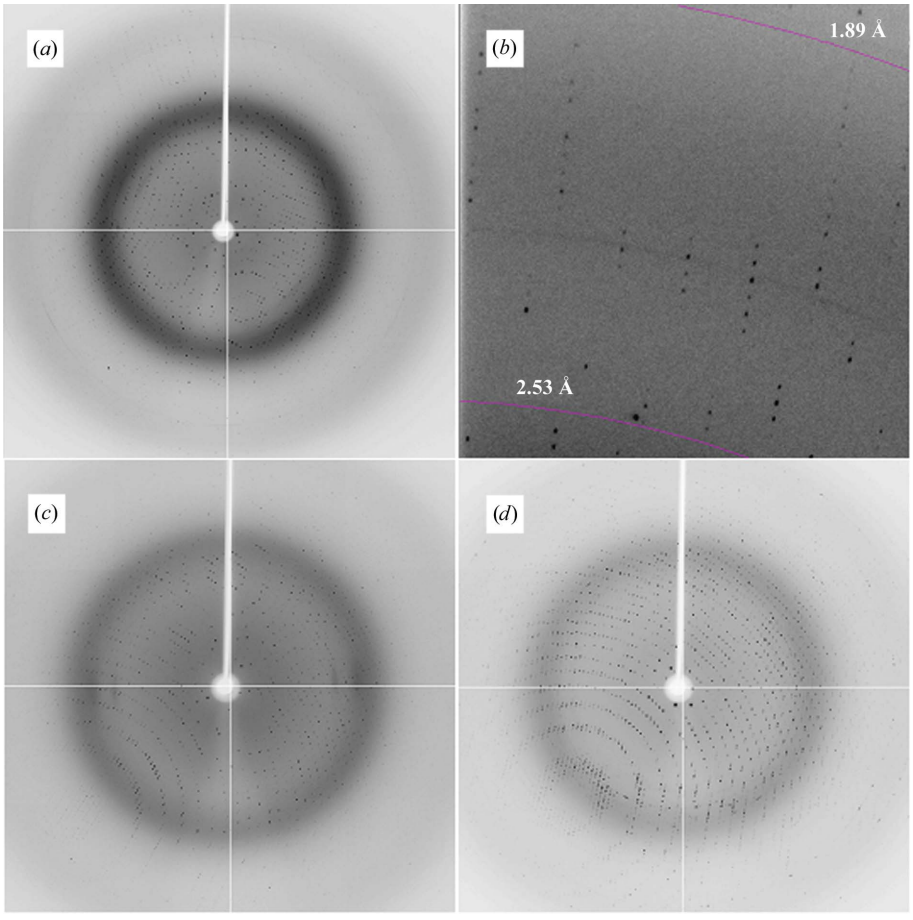

Figure 3
X-ray diffraction images of *X. campestris* GumK crystals. (*a*) Native, (*b*) detailed section of (*a*), (*c*) K_2PtCl_4 derivative, (*d*) $HgCl_2$ derivative.

Table 2
Data-collection statistics for heavy-atom derivatives.

Values in parentheses are for the highest resolution shell.

	K_2PtCl_4 derivative		$HgCl_2$ derivative	
Wavelength (Å)	1.0718	1.0722	1.009	1.0092
Resolution range (Å)	33.15–2.0 (2.1–2.0)	33.33–2.0 (2.1–2.0)	38.1–2.3 (2.4–2.3)	36.6–2.3 (2.4–2.3)
Unique reflections	51441	51635	33651	33504
R_{sym}† (%)	7.3 (38.9)	7.3 (38.9)	9.0 (36.4)	9.0 (36.5)
Completeness (%)	95.5 (75.1)	95.5 (74.8)	98.4 (97.3)	98.4 (97.3)
Anomalous completeness (%)	91.8 (63.9)	92.1 (74.8)	97.1 (87.7)	96.9 (87.4)
$\langle I/\sigma(I)\rangle$	17.3 (2.8)	17.4 (2.8)	25.6 (8.1)	24.9 (8.0)
Mosaicity (°)	0.21	0.21	0.34	0.34
Total observations	325071	330711	419416	408115
Multiplicity	6.3	6.3	12.5	12.5
Anomalous multiplicity	3.4	3.4	6.7	6.7
MAD phasing statistics				
Phasing power‡				
Isomorphous	—	0.543	—	0.338
Anomalous	2.094	1.267	1.010	0.867
Overall figure of merit	0.4811		0.3930	
R_{cullis}				
Isomorphous	—	0.683	—	0.840
Anomalous	0.726	0.739	0.883	0.901

† $R_{sym} = \sum_{hkl}\sum_i |I_{hkl,i} - \langle I_{hkl}\rangle| / \sum_{hkl}\sum_i I_{hkl}$. ‡ Phasing power is FH_c/E for the isomorphous case and $2\,FH''_c/E$ for the anomalous case, where FH_c is the calculated heavy-atom structure factor and E is the r.m.s. lack of closure.

crystallization attempts using this substituted protein have been unsuccessful.

2.4. Heavy-atom derivatization

Isomorphous crystal derivatives were obtained by soaking crystals in optimized crystallization buffer supplemented with 10 m*M* K_2PtCl_4 for 2 h or 1 m*M* $HgCl_2$ for 6 h. These crystals diffracted to 2.0 and 2.3 Å, respectively (Figs. 3*c* and 3*d*). Data sets for multiwavelength anomalous diffraction experiments for the platinum derivative (λ = 1.0718 and 1.0722 Å) and for the mercurial derivative (λ = 1.009 and 1.0092 Å) were collected on beamline X12C, NSLS at 110 K. Data were processed using the programs *MOSFLM*, *SCALA* and *TRUNCATE* from the *CCP*4 program suite (Collaborative Computational Project, Number 4, 1994) (Table 2). Crystals proved to have $P6_522$ symmetry, with identical unit-cell parameters to the native crystal. Anomalous phasing power (Table 2), calculated using *SHARP* (Bricogne *et al.*, 2003), for the platinum derivative was around 6.0 for the lower resolution shells, decreasing to 1.0 at 2.5 Å. For the mercurial derivative, the anomalous phasing power was 3.5 for the lower resolution shells and 1.1 at 3.9 Å. Progress is being made towards solving the phase problem using multiwavelength anomalous diffraction.

We thank Gastón Mayol (Fundación Instituto Leloir) for help with protein purification, Anand Saxena (X12C, NSLS) for help with data collection and Eleonora Freire, Juan Pablo Acierno and Sebastian Klinke (Fundación Instituto Leloir) for helpful discussions. This work was supported by grants PICT 11-703 from Agencia Nacional de Promoción Científica y Tecnológica and UBACyT X-193 from Universidad de Buenos Aires (Argentina). Máximo Barreras is recipient of a doctoral fellowship from Consejo Nacional de Investigaciones Cientificas y Técnicas, Argentina.

References

Barreras, M., Abdian, P. L. & Ielpi, L. (2004). *Glycobiology*, **14**, 233–241.
Becker, A., Katzen, F., Puhler, A. & Ielpi, L. (1998). *Appl. Microbiol. Biotechnol.* **50**, 145–152.
Bricogne, G., Vonrhein, C., Flensburg, C., Schiltz, M. & Paciorek, W. (2003). *Acta Cryst.* D**59**, 2023–2030.
Buschiazzo, A., Ugalde, J. E., Guerin, M. E., Shepard, W., Ugalde, R. A. & Alzari, P. M. (2004). *EMBO J.* **23**, 3196–3205.
Chen, L., Men, H., Ha, S., Ye, X. Y., Brunner, L., Hu, Y. & Walker, S. (2002). *Biochemistry*, **41**, 6824–6833.
Coligan, J. E., Dunn, B. M., Ploegh, H. L., Speicher, D. W. & Wingfield, P. T. (2004). Editors. *Current Protocols in Protein Science*. New York: John Wiley & Sons.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Ha, S., Walker, D., Shi, Y. & Walker, S. (2000). *Protein Sci.* **9**, 1045–1052.
Hampel, A., Labanauskas, M., Connors, P. G., Kirkegard, L., RajBhandary, U. L., Sigler, P. B. & Bock, R. M. (1968). *Science*, **162**, 1384–1387.
Ielpi, L., Couso, R. O. & Dankert, M. A. (1993). *J. Bacteriol.* **175**, 2490–2500.
Jansson, P. E., Kenne, L. & Lindberg, B. (1975). *Carbohydr. Res.* **45**, 275–282.
Katzen, F., Ferreiro, D. U., Oddo, C. G., Ielmini, M. V., Becker, A., Pühler, A. & Ielpi, L. (1998). *J. Bacteriol.* **180**, 1607–1617.
Sutherland, I. W. (1994). *Biotechnol. Adv.* **12**, 393–448.
Sutherland, I. W. & Tait, M. I. (1992). *Encyclopedia of Microbiology*, edited by J. Lederburg, pp. 339–349. San Diego, CA, USA: Academic Press.