crystallization papers

Crystallization and preliminary X-ray analysis of the glycogen synthase from *Pyrococcus abyssi*

**Cristina Horcajada,[a,b] Emili Cid,[a]
Joan J. Guinovart,[a,b] Nuria
Verdaguer[c] and Juan C. Ferrer[a]***

[a]Departament de Bioquímica i Biologia
Molecular, Facultat de Química, Universitat de
Barcelona, Martí i Franquès 1,
E-08028 Barcelona, Spain, [b]Institut de Recerca
Biomèdica de Barcelona, Parc Científic de
Barcelona, Universitat de Barcelona, Josep
Samitier 4-5, E-08028 Barcelona, Spain, and
[c]Institut de Biologia Molecular de Barcelona
CSIC, Jordi Girona 18-26, E-08034 Barcelona,
Spain

Correspondence e-mail: ferrer@bq.ub.es

Glycogen synthase catalyzes the transfer of glucosyl residues from ADP- or UDP-glucose to the non-reducing end of a growing α-1,4-glucan chain. To date, no crystallographic structure of an animal/fungal glycogen synthase (family 3 of the glycosyl transferases) or a bacterial/plant glycogen/starch synthase (family 5) has been reported. This paper describes the recombinant expression, crystallization and preliminary X-ray analysis of the glycogen synthase from the hyperthermophilic archaeon *Pyrococcus abyssi*, the smallest enzyme of the members of families 3 and 5 of the glycosyl transferases. Crystals from this protein and from its selenomethionyl variant were grown in 100 mM sodium citrate pH 5.6 containing 20% PEG and 20% dioxane by the hanging-drop vapour-diffusion method at 293 K. The crystals, which grew as thin needles, diffracted to 3.5 Å resolution and belong to space group $C2$, with unit-cell parameters $a = 202$, $b = 73$, $c = 149$ Å, $\beta = 131°$. The crystallographic and biochemical data are consistent with either a dimer or a tetramer in the crystal asymmetric unit and a volume solvent content of 70 or 39%, respectively.

1. Introduction

The enzymatic formation of glycosidic bonds is by far the most frequent biochemical transformation on earth. In addition, the complex carbohydrates arising from these reactions are crucial for a large number of biological functions, which range from structural and energy-reserve roles to those related to signalling and molecular recognition. Despite its great relevance, little structural and mechanistic information is available for the enzymes that catalyze such reactions, glycosyl transferases (GTs). GTs catalyze the transfer of glycosyl residues from activated donors, such as di- and polysaccharides, sugar phosphates or, more frequently, glycosyl diphosphonucleotides, to oligosaccharides, proteins, lipids, DNA or several small molecules which act as specific acceptors. Essentially according to sequence similarities, GTs have been grouped into more than 60 families (Campbell *et al.*, 1997; Coutinho & Henrissat, 1999). GTs are also classified as retaining or inverting on the basis of the relative anomeric stereochemistry of the substrate and the product in the reaction catalyzed.

Compared with glycosyl hydrolases, the structural information available for GTs is recent and rather limited. Until 1999, only the structure of an inverting β-glucosyl transferase from the T4 phage was known (Vrielink *et al.*, 1994). Since then, a further six structures of GTs that also operate with inversion of the anomeric configuration of the transferred sugar have been solved (Gastinel *et al.*, 1999; Charnock & Davies, 1999; Ha *et al.*, 2000; Pedersen *et al.*, 2000; Ünligil *et al.*, 2000; Mulichak *et al.*, 2001). The structural knowledge of retaining GTs is restricted to six examples and has been obtained in the last 2 y. Two of these enzymes, LgtC from the bacterium *Neisseria meningitidis* (Persson *et al.*, 2001) and rabbit-muscle glycogenin (Gibbons *et al.*, 2002), belong to family 8 of the GTs. A bovine α-1,3-galactosyltransferase (Gastinel *et al.*, 2001) and the human blood-group GTs GTA and GTB (Patenaude *et al.*, 2002) are classified in family 6. Finally, OtsA from *Mycobacterium tuberculosis*, the structure of which has been solved recently (Gibson *et al.*, 2002), is classified in family 20. Preliminary crystallographic studies have also been reported on the glycogen synthase from *Agrobacterium tumefaciens* (Guerin *et al.*, 2003), an enzyme that belongs to family 5 of the GTs.

Although the enzymes of families 6 and 8 present no significant sequence homology, they show some common characteristics, such as the presence of a divalent cation essential for catalysis in the active site and coordinated to the side-chain carboxylates of two aspartate residues of the DXD motif. They also exhibit great structural similarity in the subdomain responsible for binding the glycosyl donor and in the overall fold (Pedersen *et al.*, 2002; Ünligil & Rini, 2000). This protein fold has been

named GT-A (Bourne & Henrissat, 2001) and is prototypically represented by the structure of SpsA (Charnock & Davies, 1999). In contrast, OtsA displays the twin Rossmann fold characteristic of the 'fold family' GT-B, which is represented by the inverting β-glucosyl transferase from the T4 phage (Vrielink *et al.*, 1994).

Family 3 of the GTs, which comprises fungal and animal glycogen synthases (GSs), and family 5, which includes bacterial GSs and plant starch synthases (SSs), also operate with retention of configuration but possess characteristics that clearly differentiate them from families 6 and 8 and make them similar to family 20. Firstly, GSs and SSs do not require a divalent cation for activity and therefore do not possess the distinctive cation-binding DXD motif. Secondly, although no structural information is available for any member of these two families, structure-prediction studies have shown that members of families 3 and 5 probably possess a common catalytic mechanism and a similar overall fold (Cid *et al.*, 2000, 2002). This structural similarity has been extended to other GT families, which presumably also belong to the GT-B folding superfamily (Wrabl & Grishin, 2001; Bourne & Henrissat, 2001).

Here, we report the recombinant expression, crystallization and preliminary structural characterization of the GS from *Pyrococcus abyssi* (*Pa*GS). This archaeal protein, which has been classified into family 5 of the GTs, is the smallest member of families 3 and 5 (Cid *et al.*, 2002) and therefore represents the minimum catalytic unit of the α-1,4-glucan synthase superfamily.

2. Experimental results

2.1. Protein expression and characterization

The open reading frame of a hypothetical *Pa*GS (PA2292; accession No. NC_000868) was amplified by PCR from a sample of *P. abyssi* genomic DNA, a generous gift from Dr R. Guerrero. The PCR-amplified fragment was cloned into pGFPCR (Cormack & Somssich, 1997) and fully sequenced. The cDNA was then subcloned into pET28a(+) (Novagen), an expression vector that adds an amino-terminal hexa-His tag to the expressed protein. A culture of *Escherichia coli* BL21-CodonPlus (DE3)-RIL (Stratagene) transformed with the *Pa*GS-encoding plasmid was grown at 310 K in LB medium supplemented with 34 µg ml^{-1} chloramphenicol and 10 µg ml^{-1}

kanamycin to an A_{600} of 0.6, induced with 0.4 mM IPTG and then grown for an additional 24 h. Following harvesting, the cells were lysed by sonication and *Pa*GS was obtained mainly in the form of insoluble aggregates. The inclusion bodies were washed with a buffer containing 500 mM NaCl, 2% Triton X-100 in 20 mM Tris–HCl pH 8.0 and dissolved in 20 mM Tris–HCl pH 8.0 plus 500 mM NaCl, 5 mM imidazole, 6 M guanidinium hydrochloride and 5 mM 2-mercaptoethanol. The solubilized unfolded protein was loaded onto a HiTrap Ni-chelating column (Amersham Biosciences) and then washed with the same buffer in which 6 M guanidine hydrochloride was replaced by 6 M urea. The in-column refolding of the bound protein was performed by the use of a linear 6–0 M urea gradient. The refolded recombinant protein was finally eluted with a linear gradient of 5–500 mM imidazole in 20 mM Tris–HCl pH 8.0, 500 mM NaCl, 5 mM 2-mercaptoethanol. Fractions containing *Pa*GS were concentrated, loaded onto a Superdex 200 HR10/30 gel-filtration column (Amersham Biosciences) and eluted with 20 mM Tris–HCl pH 7.4 containing 500 mM NaCl and 1 mM DTT as a species with an apparent molecular weight of approximately 100 kDa. The purified protein ran on a denaturing SDS polyacrylamide gel as a single band of approximately 50 kDa (the calculated molecular weight of the His-tagged recombinant protein is 51.1 kDa). These observations indicate that *Pa*GS is a homodimer, as has been found for the GS isolated from *Thermococcus hydrothermalis*, another hyperthermophilic archaeon (Gruyer *et al.*, 2002).

The refolded recombinant protein was subjected to the classical assay of GS activity (Thomas *et al.*, 1968) and was shown to catalyze the transfer of glucosyl units from ADP-glucose or UDP-glucose to a growing chain of glycogen. Furthermore, this activity

Figure 1
Monoclinic crystals of the recombinant glycogen synthase from *P. abyssi*.

Table 1
Data-collection statistics.

Values in parentheses are for the highest resolution shell.

	*Pa*GS	SeMet-*Pa*GS
Space group	C2	C2
Unit-cell parameters (Å, °)	a = 201.8, b = 73.5, c = 148.3, β = 131.2	a = 202.7, b = 73.2, c = 149.6, β = 131.3
Resolution (Å)	30–3.5 (3.6–3.5)	30–3.4 (3.5–3.4)
Unique reflections	14287	20188
Completeness (%)	68 (63)	88 (91)
Mosaicity (°)	0.7	0.5
R_{merge}† (%)	13.7 (47.1)	13.5 (49.0)
Multiplicity	2.0	2.9
$\langle I/\sigma(I)\rangle$	7.8 (1.9)	7.9 (1.8)

† $R_{merge} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl} \sum_i I_i(hkl)$.

was resistant to prolonged periods of incubation at high temperature (data not shown). Other GSs of archaeal origin use UDP-glucose (Cardona *et al.*, 2001) or ADP-glucose and UDP-glucose (Gruyer *et al.*, 2002) as glucosyl donors.

Selenomethionyl *Pa*GS was expressed in *E. coli* BL21-CodonPlus (DE3)-RIL in minimal media supplemented with 0.2% glucose, 2 µM MgSO$_4$, 0.1 µM CaCl$_2$ and a mix containing all essential amino acids at 40 mg $^{-1}$ except Gly, Ala, Pro, Asn, Cys and Met. When the culture reached an OD$_{600}$ of 0.6, selenomethionine (50 mg l^{-1}) was added and at the same time the synthesis of Met was repressed by the addition of 100 mg l^{-1} Phe, Thr and Lys and 50 mg l^{-1} Leu, Ile, Val and Pro (Van Duyne *et al.*, 1993). After an additional 15 min, expression of the recombinant protein was induced with 0.4 mM IPTG and the culture was grown for 24 h. Selenomethionyl *Pa*GS was purified following the same protocol as used for the native *Pa*GS.

2.2. Crystallization and X-ray analysis

Samples of the *Pa*GS protein (8 mg ml^{-1}) in 20 mM Tris–HCl buffer pH 7.4 containing 500 mM NaCl and 1 mM DTT were used for crystallization tests with the hanging-drop vapour-diffusion method, which were carried out at 293 K in multiwell plates utilizing commercial kits from Hampton Research. Typically, 1 µl protein solution was mixed with an equal volume of reservoir solution. Small needle-like crystals were obtained with 100 mM sodium citrate pH 5.6 containing 20% PEG 4000 and 20% dioxane. Microseeding with these crystals at 16% PEG 4000 and 20% dioxane produced larger needles (0.2 × 0.05 × 0.02 mm) (Fig. 1). The crystals were frozen in liquid

crystallization papers

Figure 2
Representation of the χ = 180° section of the self-rotation function of *Pa*GS crystals. The orthogonalization code follows the PDB convention. See text for interpretation.

nitrogen, adding 15% glycerol to the crystallization buffer as a cryoprotectant. Crystals of selenomethionyl *Pa*GS were obtained and processed following an identical protocol. Synchrotron radiation at the ESRF microfocus beamline ID13 (λ = 0.975) was used to obtain two data sets (Table 1). For technical reasons, data collection was not pursued to completeness and consequently the multiplicity in both data sets is low. A multiwavelength anomalous diffraction experiment with the selenomethionyl derivative was not attempted, since beamline ID13 is not easily tunable.

Images of both data sets were processed and scaled using the *DENZO/SCALE-PACK* package (Otwinowski & Minor, 1997) and *TRUNCATE* from the *CCP*4 program suite (Collaborative Computational Project, Number 4, 1994). Crystals proved to be monoclinic (*C*2), with a unit-cell volume of 1 655 000 Å^3. These data are compatible with the presence of two to four protein molecules in the asymmetric unit and a solvent content ranging from 70 to 39%, respectively (Matthews, 1968). A self-rotation function calculated with the program *MOLREP* (Vagin & Teplyakov, 1997) revealed an NCS peak perpendicular to the crystallographic twofold axis that could correspond to the local twofold axis of a crystallographically independent dimer in the crystal unit cell (Fig. 2).

3. Conclusions

The open reading frame PA2292, which resulted from the complete sequencing of the *P. abyssi* genome and was annotated as a hypothetical GS, encodes a thermostable enzyme capable of catalyzing the transfer of glucosyl units from ADP- or UDP-glucose to a growing glycogen chain. The recombinant expression of *Pa*GS in *E. coli* yielded insoluble aggregated protein, which was purified, refolded and crystallized. The crystals diffracted to 3.5 Å using synchroton radiation. Further improvement of the quality of the crystals and complete MAD collection should allow the solution of the crystal structure of *Pa*GS. These experiments are under way.

References

Bourne, Y. & Henrissat, B. (2001). *Curr. Opin. Struct. Biol.* **11**, 593–600.
Campbell, J. A., Davies, G. J., Bulone, V. & Henrissat, B. (1997). *Biochem. J.* **326**, 929–939.
Cardona, S., Remonsellez, F., Guiliani, N. & Jerez, C. A. (2001). *Appl. Environ. Microbiol.* **67**, 4773–4780.
Charnock, S. J. & Davies, G. J. (1999). *Biochemistry*, **38**, 6380–6385.
Cid, E., Geremia, R. A., Guinovart, J. J. & Ferrer, J. C. (2002). *FEBS Lett.* **528**, 5–11.
Cid, E., Gomis, R. R., Geremia, R. A., Guinovart, J. J. & Ferrer, J. C. (2000). *J. Biol. Chem.* **275**, 33614–33621.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Cormack, R. S. & Somssich, I. E. (1997). *Tech. Tips Online*, **1**, T01107.
Coutinho, P. M. & Henrissat, B. (1999). *Carbohydrate-Active Enzymes Server.* http://afmb.cnrs-mrs.fr/CAZY/index.html.
Gastinel, L. N., Bignon, C., Misra, A. K., Hindsgaul, O., Shaper, J. H. & Joziasse, D. H. (2001). *EMBO J.* **20**, 638–649.
Gastinel, L. N., Cambillau, C. & Bourne, Y. (1999). *EMBO J.* **18**, 3546–3557.
Gibbons, B. J., Roach, P. J. & Hurley, T. D. (2002). *J. Mol. Biol.* **319**, 463–477.
Gibson, R. P., Turkenburg, J. P., Charnock, S. J., Lloyd, R. & Davies, G. J. (2002). *Chem. Biol.* **12**, 1337–1346.
Gruyer, S., Legin, E., Bliard, C., Ball, S. & Duchiron, F. (2002). *Curr. Microbiol.* **44**, 206–211.
Guerin, M. E., Buschiazzo, A., Ugalde, J. E., Ugalde, R. A. & Alzari, P. M. (2003). *Acta Cryst.* D**59**, 526–528.
Ha, S., Walker, D., Shi, Y. & Walker, S. (2000). *Protein Sci.* **9**, 1045–1052.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Mulichak, A. M., Losey, H. C., Walsh, C. T. & Garavito, R. M. (2001). *Structure*, **9**, 547–557.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Patenaude, S. I., Seto, N. O., Borisova, S. N., Szpacenko, A., Marcus, S. L., Palcic, M. M. & Evans, S. V. (2002). *Nature Struct. Biol.* **9**, 685–690.
Pedersen, L. C., Darden, T. A. & Negishi, M. (2002). *J. Biol. Chem.* **277**, 21869–21873.
Pedersen, L. C., Tsuchida, K., Kitagawa, H., Sugahara, K., Darden, T. A. & Negishi, M. (2000). *J. Biol. Chem.* **275**, 34580–34585.
Persson, K., Ly, H. D., Dieckelmann, M., Wakarchuk, W. W., Withers, S. G. & Strynadka, N. C. J. (2001). *Nature Struct. Biol.* **8**, 166–175.
Thomas, J. A., Schlender, K. K. & Larner, J. A. (1968). *Anal. Biochem.* **25**, 486–499.
Ünligil, U. M. & Rini, J. M. (2000). *Curr. Opin. Struct. Biol.* **10**, 510–517.
Ünligil, U. M., Zhou, S., Yuwaraj, S., Sarkar, M., Schachter, H. & Rini, J. M. (2000). *EMBO J.* **19**, 5269–5280.
Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.
Van Duyne, G. D., Standaert, R. F., Karplus, P. A., Schreiber, S. L. & Clardy, J. (1993). *J. Mol. Biol.* **229**, 105–124.
Vrielink, A., Ruger, W., Driessen, H. P. & Freemont, P. S. (1994). *EMBO J.* **13**, 3413–3422.
Wrabl, J. O. & Grishin, N. V. (2001). *J. Mol. Biol.* **314**, 365–374.