Acta Crystallographica Section F
Structural Biology Communications
ISSN 2053-230X

# Purification, crystallization and preliminary X-ray crystallographic analysis of glycosyltransferase-1 from *Bacillus cereus*

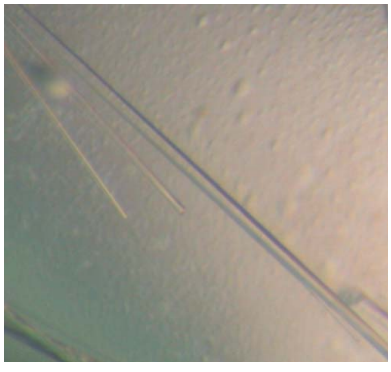
**Yin-Cheng Hsieh,[a]‡ Hsi-Ho Chiu,[b]‡ Yen-Chieh Huang,[a] Hoong-Kun Fun,[c,d] Chia-Yu Lu,[b] Yaw-Kuen Li[b]* and Chun-Jung Chen[a,c,e,f]***

[a]Life Science Group, Scientific Research Division, National Synchrotron Radiation Research Center, Hsinchu 30076, Taiwan, [b]Department of Applied Chemistry, National Chiao Tung University, Hsinchu 30010, Taiwan, [c]Department of Pharmaceutical Chemistry, College of Pharmacy, King Saud University, Riyadh 11451, Saudi Arabia, [d]X-ray Crystallography Unit, School of Physics, Universiti Sains Malaysia, 11800 USM, Malaysia, [e]Department of Physics, National Tsing Hua University, Hsinchu 30013, Taiwan, and [f]Institute of Biotechnology and the Center for Bioscience and Biotechnology, National Cheng Kung University, Tainan City 701, Taiwan

‡ These authors contributed equally to this work.

Correspondence e-mail: ykl@faculty.nctu.edu.tw, cjchen@nsrrc.org.tw

Received 11 April 2014
Accepted 20 June 2014





Glycosyltransferases (GTs), which are distributed widely in various organisms, including bacteria, fungi, plants and animals, play a role in synthesizing biological compounds. Glycosyltransferase-1 from *Bacillus cereus* (*Bc*GT-1), which is capable of transferring glucose to small molecules such as kaempferol and quercetin, has been identified as a member of the family 1 glycosyltransferases which utilize uridine diphosphate glucose (UDP-glucose) as the sugar donor. *Bc*GT-1 (molecular mass 45.5 kDa) has been overexpressed, purified and crystallized using the hanging-drop vapour-diffusion method. According to X-ray diffraction of *Bc*GT-1 crystals to 2.10 Å resolution, the crystal belonged to space group *P*1, with unit-cell parameters $a$ = 54.56, $b$ = 84.81, $c$ = 100.12 Å, $\alpha$ = 78.36, $\beta$ = 84.66, $\gamma$ = 84.84°. Preliminary analysis indicates the presence of four *Bc*GT-1 molecules in the asymmetric unit with a solvent content of 50.27%.

## 1. Introduction

Glycosyltransferases (GTs) have been found in a broad range of organisms such as microbes, plants and animals (Campbell *et al.*, 1997; Paquette *et al.*, 2003; Bowles *et al.*, 2006). These enzymes, which contain two binding sites to carry out their functions by transferring the sugar from a donor to an acceptor, are involved in several important antibiotic biosynthesis pathways, such as those of chloroeremomycin, vancomycin and balhimycin (Mulichak *et al.*, 2001; Kahne *et al.*, 2005; Pelzer *et al.*, 1999). In contrast to the highly diverse primary structures of GTs, which are presently classified into 94 families (CAZY database), their tertiary structural folds are divided into two major groups: GT-A and GT-B (Breton *et al.*, 2006; Lairson *et al.*, 2008). The GT-A fold contains one major globular structure that associates with the closely abutting binding sites of the sugar donor and acceptor, whereas the GT-B fold comprises two separated globular structures that face each other to bind the sugar donor and acceptor. Although GT-A and GT-B exhibit distinct overall structural folds, they share a general nucleotide-binding domain: the $\beta/\alpha/\beta$ Rossmann fold. The most common donors for GTs are nucleoside-diphosphate sugars such as UDP-galactose and GDP-mannose, whereas their diverse acceptors include proteins, lipids, steroids and small molecules (Hashimoto *et al.*, 2009). In addition, the sugar may be conjugated onto the carbon, nitrogen, sulfur or oxygen of their acceptors. Although many proteins have been classified into the GT family, the mechanism of glycoside-bond formation in the glycosyltransferases remains unclear.

Based on sequence alignment, glycosyltransferase-1 (*Bc*GT-1) from *Bacillus cereus* is classified into the uridine diphosphate (UDP)-containing family 1 glycosyltransferases (UGT), which utilize UDP as a cofactor to transfer sugars to acceptors. Proteins in the UGT family commonly utilize their C-terminal regions to bind to sugar donors such as UDP-glucose, UDP-galactose and UDP-rhammose, whereas the N-terminal region binds to sugar acceptors such as alkaloids,

flavonoids, antibiotics and plant hormones. Four *Bcgt* genes have previously been identified by genome analysis, among which *Bcgt-1* has been cloned and studied (Ko *et al.*, 2006). *Bc*GT-1 employs glucose as a major sugar to conjugate onto the O-linking sites of products. Several flavonoids have been identified to activate *Bc*GT-1, among which kaempferol and quercetin generate the greatest activity of the enzyme. *Bc*GT-1 is capable of transferring three sugar molecules to three distinct hydroxyl groups of kaempferol (Ko *et al.*, 2006). Mutations of some residues on the protein surface have been proposed to convert the product *O*-triglucoside kaempferol to *O*-monoglucoside kaempferol. The product transformation was also related to alteration of the secondary structure of the protein (Jung *et al.*, 2010).

The products of GTs have been reported in therapies for several diseases. For example, flavonoids, which scavenge free radicals, have served in the treatment of congestive heart failure and the improvement of blood circulation (Middleton *et al.*, 2000; Perez-Vizcaino & Duarte, 2010). Etoposide, which inhibits DNA topoisomerase, is a participant in anticancer therapy (Pommier *et al.*, 2012). Anthracyclines, which inhibit DNA synthesis by intercalation, are common drugs for the treatment of many cancers such as leukaemias, lymphomas and breast, uterine, lung and ovarian cancers (Minotti *et al.*, 2004). GT is hence an important enzyme in drug development.

To elucidate the structure and function of *Bc*GT-1, we have obtained pure protein in a large quantity for crystallization. Here, we report the overexpression, purification, crystallization and preliminary characterization by X-ray diffraction of glycosyltransferase-1 from *B. cereus*.

## 2. Materials and methods

### 2.1. Protein expression and purification

The synthetic gene of *Bc*GT-1 (GenBank accession No. AAS41089.1) was inserted into the expression vector pRSET A (Invitrogen) between *Nde*I and *Hin*dIII restriction sites and was expressed in *Escherichia coli* ECOS21 (Yeastern, Taiwan). The bacteria were cultured in Luria–Bertani (LB) medium with ampicillin (1 m*M*) at 289 K and induced by 1 m*M* isopropyl $\beta$-D-1-thiogalactopyranoside (IPTG). After induction for 16–18 h, the cells were harvested by centrifugation (8000 rev min$^{-1}$); the cell pellet was washed with water and resuspended in 25 m*M* sodium phosphate buffer pH 7.0 containing 5 m*M* dithiothreitol (DTT) and 1 m*M* phenylmethanesulfonyl fluoride (PMSF). After disrupting the cells with a French press, the cell debris was removed by centrifugation (18 000 rev min$^{-1}$) at 277 K for 30 min. The supernatant was loaded onto an ion-exchange DEAE column (Toyopearl 650M) which had been pre-equilibrated with 25 m*M* sodium phosphate buffer pH 7.0 containing 5 m*M* DTT. After elution with 100 m*M* NaCl in pre-equilibration buffer, the fractions that showed positive results with the Bradford assay were collected. These fractions were further purified with a size-exclusion column (Superdex 200, GE Healthcare) and eluted with 25 m*M* Tris buffer pH 7.0 containing 5 m*M* DTT. A 12.5% SDS–PAGE analysis showed a single band near 45.5 kDa, which was further identified to be *Bc*GT-1 by in-gel trypsin digestion and peptide mapping with MALDI–TOF–MS (Bruker Autoflex III). For the purposes of crystallization, the concentration of purified *Bc*GT-1 was further enhanced to 12 mg ml$^{-1}$ by ultrafiltration (Vivaspin, 10 000 MWCO, GE Healthcare).

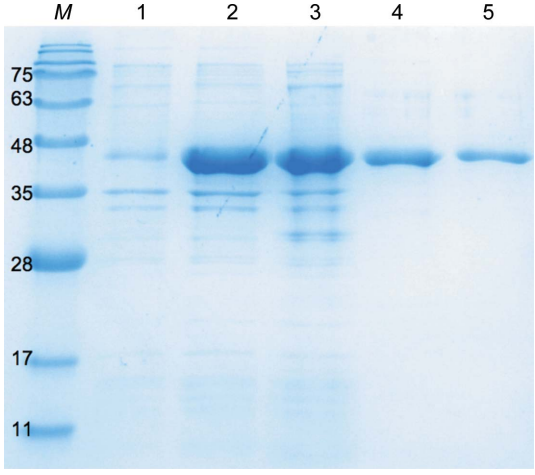



**Figure 1**
Analysis of molecular mass of *Bc*GT-1 by 12.5% SDS–PAGE. Lane *M*, protein marker (labelled in kDa). Lane 1, cellular proteins from whole cells before IPTG induction. Lane 2, cellular proteins from whole cells after IPTG induction. Lane 3, supernatant after cell disruption and centrifugation. Lane 4, purified enzyme from the ion-exchange (DEAE) column. Lane 5, purified enzyme from the size-exclusion (Superdex 200) column.

### 2.2. Crystallization

Crystallization was performed by the hanging-drop vapour-diffusion method at 291 K using 48-well VDX plates. Small crystals were observed from a condition consisting of 25%(*w*/*v*) PEG 3350, 200 m*M* sodium chloride, 100 m*M* Tris buffer pH 8.5 within 3 d of initial screening using the Crystal Screen kit (Hampton Research). This condition was further optimized to produce larger *Bc*GT-1 crystals using hanging drops (2 µl) consisting of 1 µl protein solution and 1 µl reservoir solution equilibrated against 150 µl reservoir solution consisting of 28%(*w*/*v*) PEG 3350, 200 m*M* sodium chloride, 100 m*M* Tris buffer pH 8.5. Crystals of diffraction quality were used to collect X-ray diffraction data.

### 2.3. X-ray data collection and processing

The protein crystals were initially screened and characterized with synchrotron radiation on the SPXF beamline BL13B1 equipped with a CCD detector (Q315r, ADSC) at the National Synchrotron Radiation Research Center (NSRRC), Taiwan. Data collection was

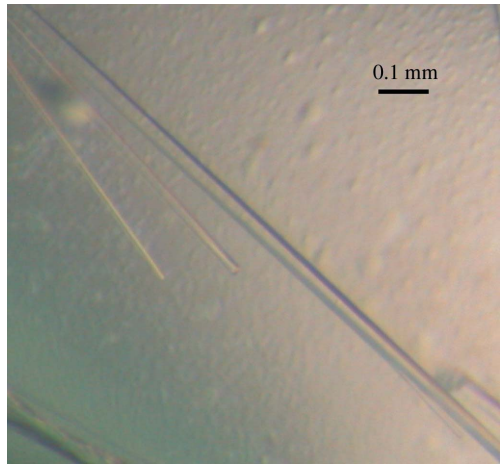



**Figure 2**
Single crystals of *Bc*GT-1 grown by the hanging-drop method.

**Table 1**
X-ray diffraction statistics of *Bc*GT-1 crystals.

Values in parentheses are for the highest resolution shell.

| | |
|---|---|
| Wavelength (Å) | 1.00 |
| Temperature (K) | 110 |
| Resolution range (Å) | 30.0–2.10 (2.18–2.10) |
| Space group | *P*1 |
| Unit-cell parameters (Å, °) | $a$ = 54.56, $b$ = 84.81, $c$ = 100.11, $\alpha$ = 78.36, $\beta$ = 84.66, $\gamma$ = 84.84 |
| Unique reflections | 96752 |
| Completeness (%) | 97.3 (96.8) |
| $\langle I/\sigma(I)\rangle$ | 11.2 (1.9) |
| Average multiplicity | 2.1 (2.1) |
| $R_{\rm merge}$† (%) | 5.6 |
| Mosaicity (°) | 0.71 |
| No. of molecules per asymmetric unit | 4 |
| Matthews coefficient (Å$^3$ Da$^{-1}$) | 2.47 |
| Solvent content (%) | 50.27 |

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the $i$th measurement and $\langle I(hkl)\rangle$ is the weighted mean of all measurements of $I(hkl)$.

completed on the Taiwan-contracted beamline BL12B2 for protein crystallography equipped with a CCD detector (Q210r, ADSC) at SPring-8, Japan. The crystal was transferred from a crystallization drop into 10 µl cryoprotectant solution consisting of 28%(*w*/*v*) PEG 3350, 200 m*M* sodium chloride, 20%(*v*/*v*) glycerol, 0.1 *M* Tris buffer pH 8.5 for a few seconds, mounted on a nylon loop (0.3–0.4 mm; Hampton Research) and subsequently flash-cooled in liquid nitrogen at 100 K. For complete data collection, 360° rotations with 1.0° oscillation were measured with an X-ray wavelength of 1.00 Å, an exposure duration of 15 s and a crystal-to-detector distance of 180 mm at 110 K in a nitrogen stream using a cryo-system (X-Stream, Rigaku/MSC Inc.). All data were indexed, integrated and scaled using *HKL*-2000 (Otwinowski & Minor, 1997).

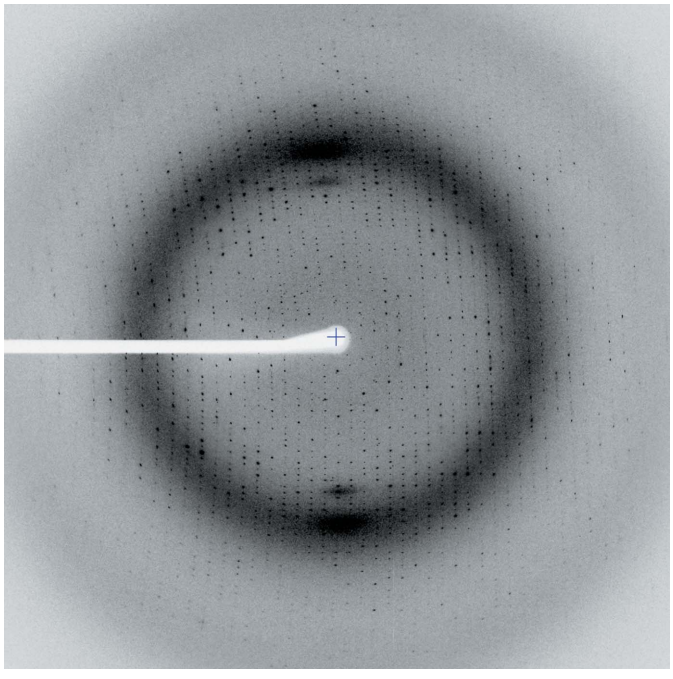


**Figure 3**
X-ray diffraction image of the *Bc*GT-1 crystal showing diffraction to 2.10 Å resolution.

## 3. Results and discussion

The *Bc*GT-1 gene was initially inserted into the vector pRSET A with *Xho*I and *Kpn*I cutting sites, which contained a 6×His tag before the N-terminus of *Bc*GT-1. However, this overexpression did not produce protein in a sufficient quantity for crystallization trials. Fortunately, excluding the 6×His tag by inserting the gene into the *Nde*I and *Hin*dIII sites produced *Bc*GT-1 in large quantities with an expression yield of ∼10 mg per 250 ml. Purification with both DEAE and Superdex-200 columns produced highly pure protein of molecular mass ∼45.5 kDa according to SDS–PAGE analysis (Fig. 1). The purified *Bc*GT-1 was further confirmed by peptide mapping with MALDI–TOF–MS.

An initial crystal screening trial was performed with *Bc*GT-1 in phosphate buffer, which easily generated flat and rectangular salt crystals in the crystallization drops. After buffer exchange from phosphate buffer to 25 m*M* Tris buffer pH 7.0, crystals of *Bc*GT-1 could be obtained in several independent conditions after initial screening, but only the conditions reported here produced X-ray diffraction data of high quality with a resolution of up to 2.10 Å. Although the protein purity was fairly good (>95%) after the first DEAE column, crystals were only obtained after another size-exclusion column (Superdex 200) was applied to improve the purity in the final purification step. Protein crystals with a needle shape appeared after 3 d and continued to grow to a final size of 0.06 × 0.06 × 0.8 mm within one week in an incubator at 291 K (Fig. 2). A long needle crystal was cut and shortened to a smaller size of about 0.06 × 0.06 × 0.3 mm before being mounted on a loop for data collection. The protein crystals were sensitive to variation in precipitant concentration while being transferred to the cryoprotectant solution containing glycerol (20%).

Selected crystals with mosaicity <1° that diffracted to higher resolution (>2.5 Å) were collected. Radiation damage that caused a decrease in $I/\sigma(I)$ and an increase in $R_{\rm merge}$ was observed on protracted exposure during data collection. Analysis of the diffraction pattern indicates that the crystals exhibit a triclinic symmetry and belong to space group *P*1 (Fig. 3). Assuming the presence of four

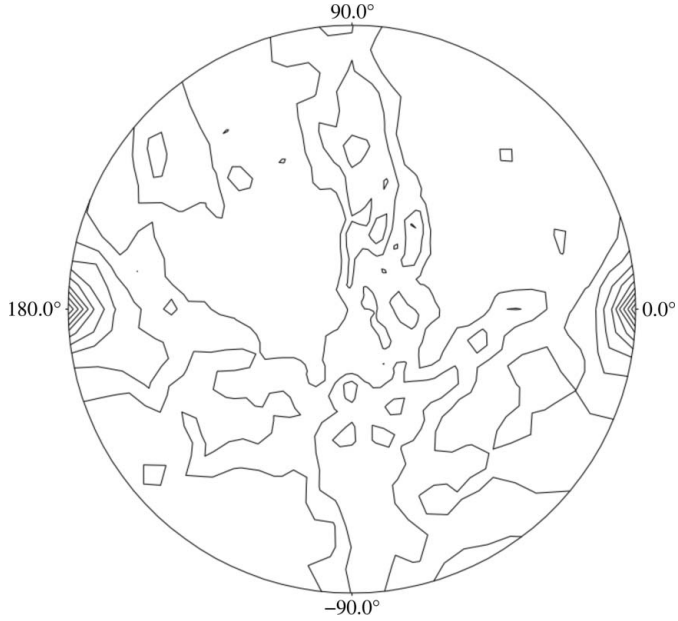



**Figure 4**
The self-rotation function with $\kappa$ = 180° shows noncrystallographic twofold symmetry.

molecules of *Bc*GT-1 per asymmetric unit, the Matthews coefficient is estimated to be 2.47 Å$^3$ Da$^{-1}$ with a solvent content of 50.27% (Matthews, 1968), which is within the normal range for protein crystals. The self-rotation function calculated using the *Self Rotation Function Server* (http://www.doe-mbi.ucla.edu/~sawaya/selfrot/) with data in the resolution range 15–4 Å shows noncrystallographic twofold symmetry (Fig. 4), suggesting that two sets of homologous dimers exist with the twofold axes parallel to each other in an asymmetric unit. Details of the data statistics of X-ray diffraction are summarized in Table 1.

An initial attempt to solve the crystal structure of *Bc*GT-1 with molecular replacement was performed using the monomer structures of glycosyltransferase from *Micromonospora echinospora* (sequence identity 29%; PDB entry 3rsc; Chang *et al.*, 2011) and *Streptomyces antibioticus* (sequence identity 27%; PDB entry 2iyf; Bolam *et al.*, 2007), but all the calculations failed to yield an obvious solution. From the low sequence identities and several gaps based on sequence alignments, the structure of *Bc*GT-1 is expected to differ markedly from the available structures. A search for useful heavy-atom derivatives and preparation of selenomethionine derivatives for SAD or MAD phasing for structure determination is in progress.

We are grateful to our colleagues, Dr Yuch-Cheng Jean and supporting staff for technical assistance with and discussion of the use of the synchrotron-radiation X-ray facility for data collection at BL13B1 of the NSRRC, Taiwan, Masato Yoshimura for technical assistance at BL12B2 and Osaka beamline BL44XU of SPring-8, Japan under proposal Nos. 201A4009 and 2012A6760, and Ms Ko-Jen Su and Dr Yu-Kuo Wang at NCTU for help in protein preparation and MALDI–TOF–MS peptide-mapping analysis. This work was supported in part by National Science Council (NSC) grants 101-2628-B-213-001-MY4, 102-2627-M-213-001-MY3 and National Synchrotron Radiation Center (NSRRC) grants 1013RSB02 and 1023RSB02 to C-JC and NSC 101-2113-M-009-012 -MY3 to Y-KL. The authors extend their appreciation to the Deanship of Scientific Research at King Saud University for funding the work through research group project No. RGP-VPP-207.

## References

Bolam, D. N., Roberts, S. M., Proctor, M. R., Turkenburg, J. P., Dodson, E. J., Martinez-Fleites, C., Yang, M., Davis, B. G., Davies, G. J. & Gilbert, H. J. (2007). *Proc. Natl Acad. Sci. USA*, **104**, 5336–5341.
Bowles, D., Lim, E. K., Poppenberger, B. & Vaistij, F. E. (2006). *Annu. Rev. Plant Biol.* **57**, 567–597.
Breton, C., Šnajdrová, L., Jeanneau, C., Koča, J. & Imberty, A. (2006). *Glycobiology*, **16**, 29R–37R.
Campbell, J. A., Davies, G. J., Bulone, V. & Henrissat, B. (1997). *Biochem. J.* **15**, 929–939.
Chang, A., Singh, S., Helmich, K. E., Goff, R. D., Bingman, C. A., Thorson, J. S. & Phillips, G. N. (2011). *Proc. Natl Acad. Sci. USA*, **108**, 17649–17654.
Hashimoto, K., Tokimatsu, T., Kawano, S., Yoshizawa, A. C., Okuda, S., Goto, S. & Kanehisa, M. (2009). *Carbohydr. Res.* **344**, 881–887.
Jung, N. R., Joe, E. J., Kim, B.-G., Ahn, B. C., Park, J. C., Chong, Y. & Ahn, J.-H. (2010). *J. Microbiol. Biotechnol.* **20**, 1393–1396.
Kahne, D., Leimkuhler, C., Lu, W. & Walsh, C. (2005). *Chem. Rev.* **105**, 425–448.
Ko, J. H., Kim, B. G. & Ahn, J.-H. (2006). *FEMS Microbiol. Lett.* **258**, 263–268.
Lairson, L. L., Henrissat, B., Davies, G. J. & Withers, S. G. (2008). *Annu. Rev. Biochem.* **77**, 521–555.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Middleton, E. Jr, Kandaswami, C. & Theoharides, T. C. (2000). *Pharmacol. Rev.* **52**, 673–751.
Minotti, G., Menna, P., Salvatorelli, E., Cairo, G. & Gianni, L. (2004). *Pharmacol. Rev.* **56**, 185–229.
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.
Paquette, S., Møller, B. L. & Bak, S. (2003). *Phytochemistry*, **62**, 399–413.
Pelzer, S., Sübmuth, R., Heckmann, D., Recktenwald, Z., Walsh, C. T. & Garavito, R. M. (1999). *Antimicrob. Agents Chemother.* **43**, 1565–1573.
Perez-Vizcaino, F. & Duarte, J. (2010). *Mol. Aspects Med.* **31**, 478–494.
Pommier, Y., Leo, E., Zhang, H. & Marchand, C. (2012). *Chem. Biol.* **17**, 421–433.