Acta Crystallographica Section D
**Biological Crystallography**
ISSN 0907-4449

# Overexpression, purification, crystallization and preliminary crystallographic studies on a thermostable β-glycosidase from *Thermus nonproteolyticus* HG102

**Xiang-Yuan He,[a] Xin-Quan Wang,[b] Shou-Jun Yang,[a] Wen-Rui Chang[b] and Dong-Cai Liang[b]***

[a]Institute of Microbiology, Chinese Academy of Sciences, Beijing 100080, People's Republic of China, and [b]National Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, People's Republic of China

Correspondence e-mail: dcliang@sun5.ibp.ac.cn

Received 19 March 2001
Accepted 3 July 2001

A thermostable β-glycosidase (Tn-gly) from *Thermus nonproteolyticus* HG102 has been cloned and overexpressed in *Escherichia coli*. The recombinant enzyme, with a molecular mass of 48.9 kDa, was purified to homogeneity. It can hydrolyze a wide range of oligosaccharides and perform transglycosylation. Crystals of the recombinant enzyme were grown by the hanging-drop vapour-diffusion technique with MPD and NaCl as precipitants. They belong to the orthorhombic space group $P2_12_12_1$, with unit-cell parameters $a = 66.7$, $b = 94.6$, $c = 176.5$ Å.

## 1. Introduction

β-Glycosidases (E.C. 3.2.1.21) catalyze the hydrolytic cleavage of β-glycosyl linkages. They are ubiquitous enzymes which have been isolated and characterized from various organisms. These enzymes are characterized by a broad substrate specificity, which makes them potential tools for several applications (Ichikawa *et al.*, 1992). In this regard, β-glycosidases from thermophilic sources are particularly attractive because of their biotechnological advantages for many stabilized biocatalysts. Furthermore, they may contribute to the study of the structure–function relationship of thermophilic enzymes by comparison of their properties with those of mesophilic enzymes (Thomas & William, 1986).

The thermostable β-glycosidase (Tn-gly) was produced by the thermophilic eubacterium *T. nonproteolyticus* HG102 (Cai *et al.*, 1992), which was isolated from a hot spring in Guangdong Province, southern China. Tn-gly has been cloned, expressed in *E. coli* and characterized (He *et al.*, 2001). It is very thermostable, with an optimum temperature of 363 K and a half-life of 2.5 h at 363 K. It can hydrolyze a wide range of oligosaccharides and perform transglycosylation at high temperature.

Tn-gly belongs to the glycosyl hydrolase family 1 (Henrissat & Bairoch, 1993, 1996). Only a few three-dimensional structures of the glycosyl hydrolase family 1 have been reported (Barret *et al.*, 1995; Wiesmann *et al.*, 1995; Burmeister *et al.*, 1997; Aguilar *et al.*, 1997; Sanz-Aparicio *et al.*, 1998; Hakulinen *et al.*, 2000). However, no structures have been reported for β-glycosidases from thermophilic eubacterium. The percentage sequence identity between Tn-gly and the β-glycosidase from the thermophilic archaebacterium *Sulfolobus solfataricus* (PDB code 1gow; Aguilar *et al.*, 1997) is only 22.9%.

There have been several reviews on the molecular mechanism of the stability of thermoenzymes in recent years (Vieille & Zeikus, 1996, 2001; Danson *et al.*, 1996; Adams & Kelly, 1998; Zeikus *et al.*, 1998). The generally recognized mechanisms include ion pairs, hydrogen bonds, hydrophobic interactions, disulfide bridges, packing efficiency, decrease of the entropy of unfolding and intersubunit interactions.

Tn-gly has a high content of Ala (12.8%), Leu (10.9%), Arg (9.6%) and Pro (8.0%), but a low content of Cys (0.45%) and Met (0.91%). Ala and Leu make an important contribution to the hydrophobic interactions that are important for protein folding. Arg may protect the hydrocarbon chains from unfavourable contact with water, thus enhancing stability (Folcarelli *et al.*, 1996). Pro can improve protein stability by introducing a more fixed tertiary structure (Watanabe *et al.*, 1996). Several single and multiple mutants that introduce residues with high helix propensity (such as Ala) can increase the enzyme stability (Vieille & Zeikus, 1996). Therefore, Tn-gly may have a variety of structural features that contribute to its thermostability. The crystal structure analysis of Tn-gly will provide important structural information for understanding its biological function including hydrolysis and transglycosyl activity, and especially its high thermostability.

## 2. Materials and methods

### 2.1. Protein preparation

The Tn-gly gene with flanking sequences was originally cloned into plasmid pHY using the previously described methods (He *et al.*,

2001). After DNA sequencing (GENBANK accession number AF225213), PCR was used to amplify the active gene. The two primers were 5′-TAGAATTCCATATG-ACCGAGAACGCCGAAAA-3′ and 5′-ATAAGCTTGGATCAATGCGCGGACC-CCCC-3′. The fragment was digested with *Nde*I and *Hin*dIII and then inserted into the pET21a vector treated with the same enzymes.

BL21(DE3) *E. coli* was transformed with the resulting pETng plasmid, grown at 310 K, induced with 1 m*M* IPTG at a cell density of $6 \times 10^8$ cells ml$^{-1}$ and grown for another 5 h. Cells were harvested by centrifugation, resuspended in buffer *A* (50 m*M* Tris–HCl pH 8.0) and disrupted by sonication. The extract was centrifuged for 15 min at 10 000*g* and the supernatant was incubated at 348 K for 30 min. After centrifugation for 15 min at 10 000*g*, the supernatant was concentrated by adding ammonium sulfate to 60% final saturation. The precipitate was dissolved in buffer *A* and the resulting solution was purified on the Äkta Prime system using a Q-Sepharose Fast Flow column (Pharmacia) pre-equilibrated with buffer *A*, then eluted with a linear gradient of 0.05–0.2 *M* NaCl in buffer *A*. The enzyme was pooled and purified further using a Resource-ISO column on the Äkta Prime system, with a linear gradient of 3.0–0 *M* $(NH_4)_2SO_4$ in buffer *A*. The enzyme was pooled and purified by MonoQ on FPLC (Pharmacia) with an linear elution gradient of 0–0.2 *M* NaCl in buffer *B* (50 m*M* Tris–HCl pH 8.5). The enzyme fractions were collected and dialyzed against 50 m*M* $NH_4HCO_3$, then lyophilized and stored at 253 K. The Tn-gly activities were assayed using previously reported methods (He *et al.*, 2001).

### 2.2. Crystallization

Crystallization experiments were performed using the hanging-drop vapour-diffusion method. Crystallization conditions were screened using sparse-matrix sampling (Jancarik & Kim, 1991). Crystals were obtained in a system using MPD and NaCl as precipitants. After optimization of the crystallization conditions, crystals were obtained by mixing 4 µl protein solution (4 mg ml$^{-1}$, 5 m*M* HEPES pH 7.5) with 2 µl reservoir solution at 291 K. The reservoir solution was 20% MPD, 0.2 *M* NaCl, 0.1 *M* sodium acetate buffer pH 4.6.

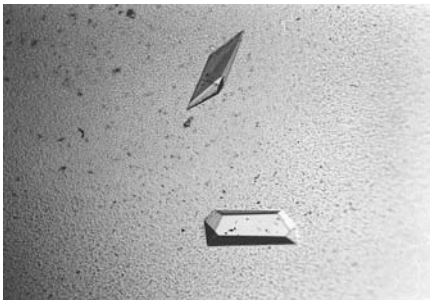

**Figure 1**
Photograph of crystals of Tn-gly grown in 20% MPD, 0.2 *M* NaCl and 100 m*M* NaAc buffer pH 4.6. Dimensions are approximately 0.25 × 0.15 × 0.1 mm.

**Table 1**
Data-processing statistics.

Values in parentheses refer to the outer resolution shell.

| | |
|---|---|
| Space group | $P2_12_12_1$ |
| Unit-cell parameters (Å) | |
| *a* | 66.7 |
| *b* | 94.6 |
| *c* | 176.5 |
| Resolution range (Å) | 50.0–2.70 (2.76–2.70) |
| No. of unique reflections | 31054 |
| Multiplicity | 4.2 |
| Completeness (%) | 98.6 (98.8) |
| $I/\sigma(I)$ | 9.8 (3.2) |
| $R_{\rm merge}$† (%) | 12.7 (36.1) |

† $R_{\rm merge} = \sum_h \sum_i |I(h,i) - \langle I(h)\rangle| / \sum_h \sum_i I(h,i)$, where $I(h,i)$ is the intensity of the $i$th measurement of reflection $h$ and $\langle I(h)\rangle$ is the mean value of $I(h,i)$ for all $i$ measurements.

### 2.3. Data collection

The diffraction ability of a small crystal (0.1 × 0.05 × 0.05 mm) was measured with a Weissenberg camera (Sakabe, 1991) at the Photon Factory (Tsukuba, Japan). The crystal initially diffracted to 2.6 Å, but no longer diffracted after the collection of several images. For data collection on our MAR Research MAR345 image-plate system, a single crystal (0.3 × 0.15 × 0.1 mm) was transferred into a cryoprotecant solution containing 25% MPD and flash-frozen to 110 K. A total of 120 images with an oscillation angle of 1° per image were collected at 110 K and processed with the *DENZO* and *SCALEPACK* programs (Otwinowski & Minor, 1997).

## 3. Results and discussion

The sequence of the Tn-gly gene (1131 bp) produced by PCR was identified by nucleotide sequencing. It codes for 436 amino acids with a molecular mass of 48.9 kDa. The expression level of plasmid pETng was much higher than that of plasmid pHY and the recombinant enzyme was easier to purify.

The purified recombinant enzyme had the same biochemical characteristics as native Tn-gly. The enzyme shows optimum activity at pH 5.6 and 363 K. It catalyzes the hydrolysis of β-D-glucoside, β-D-galactoside, β-D-fucoside and β-D-mannoside. The enzyme was extremely thermostable, with a half-life of 2.5 h at 363 K and was stable over a wide pH range. It has transglycosidic activity at high temperature.

After optimization of the crystallization conditions, crystals were obtained using MPD and NaCl as precipitants after four weeks (Fig. 1). The diffraction data was consistent with the orthorhombic space group $P2_12_12_1$, with unit-cell parameters $a = 66.7$, $b = 94.6$, $c = 176.5$ Å. Other statistics are listed in Table 1. Assuming there to be two molecules per asymmetric unit, the Matthews coefficient (Matthews, 1968) was calculated to be 2.9 Å$^3$ Da$^{-1}$, corresponding to a solvent content of 50%. Structure determination is in progress.

## References


Adams, M. W. W. & Kelly, R. M. (1998). *Trends Biotechnol.* **16**, 329–332.
Aguilar, C. F., Sanderson, I., Moracci, M., Ciaramella, M., Nucci, R., Rossi, M. & Pearl, L. H. (1997). *J. Mol. Biol.* **271**, 789–802.
Barret, T., Suresh, C. G., Tolley, S. P., Dodson, E. J. & Hughes, M. A. (1995). *Structure*, **3**, 951–960.
Burmeister, W. P., Cottaz, S., Driguez, H., Iori, R., Palmieri, S. & Henrissat, B. (1997). *Structure*, **5**, 663–675.
Cai, M. Y., Yang, S. J., Liu, J. F. & Wei, J. (1992). *Acta Microbiol. Sin.* **32**, 233–237.
Danson, M. J., Hough, D. W., Russell, R. J. M., Taylor, G. L. & Pearl, L. (1996). *Protein Eng.* **9**, 629–630.
Folcarelli, S., Battistoni, A., Carri, M. T., Polticelli, F., Falconi, M., Nicolini, L., Stella, L., Rosato, N., Rotilio, G. & Desideri, A. (1996). *Protein Eng.* **9**, 323–325.
He, X.-Y., Zhang, S.-Z. & Yang, S.-J. (2001). In the press.
Henrissat, B. & Bairoch, A. (1993). *Biochem. J.* **293**, 781–788.
Henrissat, B. & Bairoch, A. (1996). *Biochem. J.* **316**, 695–696.
Hakulinen, N., Paavilainen, S., Korpela, T. & Rorvinen, J. (2000). *J. Struct. Biol.* **129**, 69–79.
Ichikawa, Y., Look, G. C. & Wong, C. H. (1992). *Anal. Biochem.* **202**, 215–238.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **273**, 307–326.
Sakabe, N. (1991). *Nucl. Instrum. Methods Phys. Res. A*, **303**, 9728–9735.
Sanz-Aparicio, J., Hermoso, J. A., Martinez-Ripoll, M., Lequerica, J. L. & Polaina, J. (1998). *J. Mol. Biol.* **275**, 491–502.
Thomas, K. N. & William, R. K. (1986). *Thermophiles*, edited by Thomas D. Brock, pp. 197–216. New York: Wiley-Interscience.
Vieille, C. & Zeikus, J. G. (1996). *Trends Biotechnol.* **14**, 183–190.
Vieille, C. & Zeikus, J. G. (2001). *Microbiol. Mol. Biol. Rev.* **65**, 1–43.
Watanabe, K., Kitamura, K. & Suzuki, Y. (1996). *Appl. Environm. Microbiol.* **62**, 2066–2073.
Wiesmann, C., Beste, G., Hengstenberg, W. & Schulz, G. E. (1995). *Structure*, **3**, 961–968.
Zeikus, J. G., Vieille, C. & Savchenko, A. (1998). *Extremophiles*, **2**, 179–183.