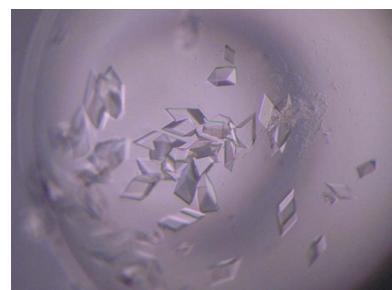


Tsutomu Nakamura,^a Yasuhiro Kashima,^b Shouhei Mine,^a Takashi Oku^b and Koichi Uegaki^{a*}

^aNational Institute of Advanced Industrial Science and Technology, Ikeda, Osaka 563-8577, Japan, and ^bThermostable Enzyme Laboratory Co. Ltd, Kobe, Hyogo 650-0047, Japan

Correspondence e-mail: k-uegaki@aist.go.jp

Received 5 September 2011
Accepted 6 October 2011



© 2011 International Union of Crystallography
All rights reserved

Crystallization and preliminary crystallographic analysis of a putative glucokinase/hexokinase from *Thermus thermophilus*

Glucokinase/hexokinase catalyzes the phosphorylation of glucose to glucose 6-phosphate, which is the first step of glycolysis. The open reading frame TTHA0299 of the extreme thermophile *Thermus thermophilus* encodes a putative glucokinase/hexokinase which contains the consensus sequence for proteins from the repressors, open reading frames and sugar kinases family. In this study, the glucokinase/hexokinase from *T. thermophilus* was purified and crystallized using polyethylene glycol 8000 as a precipitant. Diffraction data were collected and processed to 2.02 Å resolution. The crystal belonged to space group $P2_1$, with unit-cell parameters $a = 70.93$, $b = 138.14$, $c = 75.16$ Å, $\beta = 95.41^\circ$.

1. Introduction

Glycolysis is an important pathway for both energy metabolism and biosynthesis in cells. There are two major glycolytic pathways: the Embden–Meyerhof (EM) pathway and the Entner–Doudoroff (ED) pathway. In both pathways the first step of glycolysis is the phosphorylation of glucose to glucose 6-phosphate, which is catalyzed by either glucokinase (GK; EC 2.7.1.2) or hexokinase (HK; EC 2.7.1.1). These enzymes are homologous, yet functionally distinct: GK exhibits a narrow substrate specificity for glucose, whereas HK exhibits a broad substrate specificity for hexoses.

Microbial GKs/HKs have been classified into three groups. Group I includes ATP-dependent and ADP-dependent GKs (EC 2.7.1.147) from archaea such as *Pyrococcus furiosus* and *Thermococcus litoralis* (Sakuraba *et al.*, 2004) and a bifunctional ADP-dependent glucokinase/phosphofructokinase from *Methanococcus jannaschii* (Sakuraba *et al.*, 2002). Group II members include ATP-dependent GKs such as those from *Escherichia coli* (Lunin *et al.*, 2004). Group III includes ATP-dependent GKs and HKs (from both archaea and bacteria) that contain the ROK [repressors, open reading frames (ORFs) and sugar kinases] signature sequence [(L/I/V/M) X_2 G(L/I/V/M/F/C/T)GX(G/A)(L/I/V/M/F/A) X_8 GX₃₋₅(G/A/T/P)X₂G(R/K/H); consensus sequence 1; Hofmann *et al.*, 1999] and a CXCGX₂GCXE motif (consensus sequence 2; Mesak *et al.*, 2004). The ROK-type enzymes vary in their sugar specificity. The ROK GK from the hyperthermophilic bacterium *Thermotoga maritima* (Hansen & Schönheit, 2003) is highly specific for glucose, whereas the ROK enzymes from the hyperthermophilic archaea *Aeropyrum pernix* (Hansen *et al.*, 2002) and *Thermoproteus tenax* (Dörr *et al.*, 2003) and those from the thermophilic bacterium *Thermus caldophilus* (Bae *et al.*, 2005) have a broad substrate specificity. Crystallographic studies would help in understanding the factors that determine the substrate specificity of ROK-type GKs and HKs.

Thermus thermophilus is a model organism for thermophilic bacteria and has been the subject of extensive genomic and proteomic studies. ORFs for all of the enzymes involved in the ED glycolytic pathway are found in the genome sequence of *T. thermophilus*. The ORF TTHA0299 encodes a ROK-type kinase (TthGK/HK; Fig. 1). Because this enzyme is highly homologous to HK from *T. caldophilus* (Bae *et al.*, 2005), with only one residue mismatch (Fig. 1), TthGK/HK can exhibit hexokinase activity. Although the tertiary structures of several group I thermophilic ADP-dependent GKs from thermophilic archaea (Ito *et al.*, 2001, 2003; Tsuge *et al.*, 2002) are available, the tertiary structures of thermophilic ROK-type GKs/HKs have not

been reported to date. In this paper, we describe the purification and crystallization of TthGK/HK. This is a pioneering study of an ATP-dependent GK/HK from a thermophilic organism.

2. Materials and methods

2.1. Expression and purification of TthGK/HK

The coding sequence of TthGK/HK was constructed from the amino-acid sequence (Fig. 1), with codon optimization of the *E. coli*

genes. The synthesized gene was cloned into the *Nde*I and *Eco*RI sites of plasmid vector pET-21a(+). *E. coli* BL21 (DE3) cells harbouring the expression plasmid were cultivated to an A_{600} of ~ 0.6 in LB medium supplemented with $100 \mu\text{g ml}^{-1}$ ampicillin at 310 K and expression of the recombinant protein was induced by the addition of $20 \mu\text{M}$ (final concentration) isopropyl β -D-thiogalactopyranoside and incubation for a further 22 h at 310 K.

The cells were disrupted by sonication in buffer A (20 mM Tris-HCl pH 8.0) and a soluble fraction was obtained by centrifugation at 31 000g for 30 min. The soluble materials were heat-treated at 348 K

Tth	-----MKVVGDLGG-TKIAAGVFDGKRLLSKVVPPTKEGGERVAEALAAAERAEREAGVRGEAIGLG-----	64
Tca	-----MKVVGDLGG-TKIAAGVFDGKRLLSKVVPPTKEGGERVAEALAAAERAEREAGVRGEAIGLG-----	64
Tma	-----MPKLIKGVLDGGTFSVGLVSEDGKILKKVTRDTLVEENGKEDVIR-RIAETILEVSDGEEAPYVGIG-----	67
Ape	-----MAEVVAVDVGATWVRMAIVRGGVIEAIKRERNPGTEEGLERVLQGLAEGLGI DRGRVEKVGAAISG-----	66
Tte	-----MI LAI DVGATWTRVLLVSRDGR-----VERREKFRTDESPVKAARLVEGWEFEAIGVG-----	54
Eco	-----MTKYALVGVDVGGTNARLALCDIASGEISQAKTYSGLDPSLEAVIRVYLEEHKVEVKDGCIAACPI-----	68
Pfu	-----MPTWEELYKNAIKEAIKSVPKVKGVLGYNTNI DAIKYLDSKDLEERII KAGKEEVIKYSEELPKINTVSQLLGSIWSIRRGKAELFVE-----	92
Pho	-----MITMTNWESLYEKALDKVEASIRKVRGVLLAYNTNI DAIKYLKREDLEKRIEKVGKEEVRLYSEELPKIEI TIPQLLGSIWSIKRGKAELLVV-----	95
Tli	MKESLKDRIRLWKRLYVNAFENALNAIPNVKGVLAYNTNI DAIKYLDADDLEKRVTEKGKEKFIEIENPPEKISSIEELLLGGILRSIKLGKAMEWFVE-----	100
[] [] []		
Tth	-----TPGPLDFRRGVIRFAPNIPGVQDFP--IRRILEEATGRPVFLENDAAAALAEHHLGAAQGEESLLYLTVSTG-----	135
Tca	-----TPGPLDFRRGVIRFAPNIPGVQDFP--IRRILEEATGRPVFLENDAAAALAEHHLGAAQGEESLLYLTVSTG-----	135
Tma	-----SPGSI DRENGIVRFSPNFPDWHNVP--LT DELAKRTGKVF LENDANAFLVGEKWFAGGRGHDHIVALT LGTG-----	138
Ape	-----PLDLRRGYIVGSPNIK--SHIVRLSTILKRLFPKSKVAI ANDAVAAAAGEYLLGRLAGTPDLGYITMSTG-----	134
Tte	-----SIGPMDLRRGVWTASPNSP---SKSFPLVEPLLKF KRPVYVANDCVAAAGEYVLGG-WGVDNLAYLT LSTG-----	122
Eco	-----GDWVAMTNHTWAFAISIEMKKNLGFSHLEI INDFTAWSMAI PMLKKEHILQFQGAEPVEGKPIAVYVGAG-----	136
Pfu	SCPVRFYMKRKGWNELRMGGQAGIMANLLGGVYGVPIVHVPQLSRLQANLFLDGPYVYPTLENGEVKLHPKFEFSGDEENCIHYIYEFPRGFRVFEFEA-----	192
Pho	SREVREYMRKRGWDELRMGGQVGIMANLLGGVYGPVIAHVPQLSEI QASLFLDGPYVYPTFERGELRLIHPREFRKGEEDCIHYIYEFPRNFKVLDFEA-----	195
Tli	SEEVRRYI LREWGWDLIRIGGGAGIMANLLGGVYRIPITVHVPQNPKI QASLFLDGPYVYPTFEGNKLKLVHPKDAIAEEEELIHYIYEFPRGFQVFDVQA-----	200
[] [] []		
Tth	-----IGGGVVLGGRVLR-----GERGQGGELGHLTLLPGGPACG-----CGLEGCLEALAAGRALERDATYAFQRP--VDT--RELFRLFQA 209	209
Tca	-----IGGGVVLGGRVLR-----GERGQGGELGHLTLLPGGPACG-----CGLEGCLEALAAGRALERDATYAFQCP--VDT--RELFRLFQA 209	209
Tma	-----IGGGVVTGHYLLT-----GRDGIGAELGHVVVEPNPGMCN-----CGTRGCLEAVASATAIRRFLREGYKYY--HSSLVYKLAGSPEKA 215	215
Ape	-----VGGGFVVGGRLLL-----GSRGNAHEVGHIVVDMGWEGRCG-----CGGTGHWEIAAGRWPRTSSVLARGWRGPETSLYRAALEGRVG 215	215
Tte	-----LGVAVVNGHIL-----GKEGNAHELGHAVLDVRGDVKCG-----CGGLGHWEAMAGGANIPSYFRVFAARLG-----LRAPEVR 193	193
Eco	-----TGLGVAHLVHVDKR-----WVSLPGEGGHVFDPNSEEAAIIL-----EILRAEIGHVSAERVLSGPGFLVNLRYRAIVKADNRLPENLKP-KD 217	217
Pfu	PRENRFIGSADDYNTTFLIREEFRESFSEVIKNVQLAI LSGLQALTKEN-----YKEPFEIVKSNEVLNREI PVHLEFAFT--PDEKVREEIILNVLMG 285	285
Pho	PRENRFIGAADDYNPILYVREEWIERFEEIAKRSELAI ISGLHPLTQEN-----HGKPIKLVREHLKILNDLGIRAHLEFAFT--PDEVVRLEIVKLLKH 288	288
Tli	PRENRFIANADDYNARVYMRREFREGFEEITRNVELAI ISGLQVLKEYYPDGTTYKDVLDRVESHILINLRYNVKSHFEFAYT--ANRRVREALVELLPK 298	298
[] [] []		
Tth	-----GDPKAERLVLQARYVGIGLASLVKA FDPG-----VVV LGGGVALNAPEG-YWEALLEAYRRLQGWEAP-----PLRRARLGAEAGLLGAALTALEV 297	297
Tca	-----GDPKAERLVLQARYVGIGLASLVKA FDPG-----VVV LGGGVALNAPEG-YWEALLEAYRRLQGWEAP-----PLRRARLGAEAGLLGAALTALEV 297	297
Tma	-----DAKHLF DAARQGDRFALMIRDRVVDALARAVAGYI HIFNP EIVI IGGGI SRAGE I LFGPLREKVVDYIMPSFGVYEVVVASPLVEDAGI LGAA SI KERI 315	315
Ape	SAREVF EAAVGDDFALHVI D YIARASAAGIASVKAAYDVDAVI IGGSVYLNRRMLRPLIERH LAAYAP-FSSRIEVVDA SF GDNEGVMGAYAIA YRNP 314	314
Tte	TAEDVF KLY REGDRVAQMFV D HWDV DVNAAGIAT I TAA YDPEV L VVGGSI ALNHWDV D LVEA A VDR LK Y TPL T P P-----ALEKAKFGDDEVAMGAA ALA I RPP 290	290
Eco	-----ITERALADSC TDC RRALSLF CVIMGRFGGNLALNLGT FGGVFIAGGIVPRFLEFFKASGRFRAAFEDKGRFKEYVHDIPVY LIVHDNPGLLGGSAHLRQTL 317	317
Pfu	FYSVGLNEVELASIMEI LGEKKLAKELLAHD PVDVPI A VTEAMLKLAKKTG-VKRIHFHTYGYYLALTEYKGEHVRDALLFAA LAAA A KMG NITS LEEI 384	384
Pho	FYSVGLNEVELASVVSVMGEKELAERI I SKDPADP I A VIEGLLKLK I KETG-VKRIHFHTYGYYLALTREKGEHVRDALLFSALAAATKAMG KNI EK LSDI 387	387
Tli	FTSVGLNEVELASIMEI I GDEELAKEVLEG-----HIFSVIDAMNVIMDET G-I ERIHFHTYGYYLALTQYRGEEV RDALLFAS LAAA A KAMG KNL RIEQI 394	394
[] [] []		
Tth	KDGSG-----	302
Tca	KDGSG-----	302
Tma	GG-----	317
Ape	EDLPIF-----	320
Tte	DTLKKFGYPRF-----	301
Eco	GHIL-----	321
Pfu	REATSPVNEKATQVEEKLRAEYGIKEG-I GEVEGYQIAFPTKIVAKPKSTVGI GDTI SSSA F I G E F S F T L- 455	455
Pho	REGLAVPIGEQGLEVEKILEKEFSLRDG-IGSIEDYQLTFIPTKVVKKPKSTVGI GDTI SSSA F VSE F S L H-- 457	457
Tli	RDALSVPTNERAIVL EEELEK F E F E N G L I D M V D R Q L A F V P T K I V A S P K S T V G I G D T I S S S A F V S E F G M R K R 467	467

Figure 1

Sequence alignment of TthGK/HK and homologous proteins. The white and black bars indicate consensus sequences 1 and 2, respectively, of TthGK/HK. Multiple alignments were performed by *ClustalW* (Higgins *et al.*, 1996). Tth, TthGK/HK (AAS82030); Tca, HK from *Thermus caldophilus* (AAV66920); Tma, GK from *Thermotoga maritima* (AAD36537); Ape, HK from *Aeropyrum pernix* (BA81102); Tte, HK from *Thermoproteus tenax* (CAD52839); Eco, GK from *E. coli* (ACI80495); Pfu, GK from *Pyrococcus furiosus* (AAL80436); Pho, GK from *P. horikoshii* (BAA29678); Tli, GK from *Thermococcus litoralis* (Q7M537). GenBank/Swiss-Prot accession numbers are given in parentheses.

Table 1

Crystallographic data-collection statistics.

Values in parentheses are for the highest resolution shell.

X-ray source	BL38B1, SPring-8
Wavelength (Å)	1.0000
Space group	$P2_1$
Unit-cell parameters (Å, °)	$a = 70.93, b = 138.14, c = 75.16, \beta = 95.41$
Resolution range (Å)	50.0–2.02 (2.09–2.02)
Z	8
V_M (Å ³ Da ⁻¹)	2.91
R_{merge}^\dagger (%)	7.1 (36.3)
Completeness (%)	99.7 (99.3)
Total reflections	415123
Unique reflections	93772
Multiplicity	4.4 (3.8)
$\langle I/\sigma(I) \rangle$	11.6 (3.22)
Wilson plot B factor (Å ²)	26.3

† $R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of the i th observation of reflection hkl and $\langle I(hkl) \rangle$ is the average of symmetry-related observations of a unique reflection.

for 60 min and then centrifuged. The supernatant was dialyzed against buffer *A* and then loaded onto a HiTrap Q column (GE Healthcare). The bound protein was eluted with a linear gradient of 0–1 *M* NaCl in buffer *A*. The fraction containing TthGK/HK was subjected to gel filtration on a Superdex 200 pg column (GE Healthcare) with buffer *A* containing 150 mM NaCl. The protein concentration was determined with a Protein Assay Kit (Bio-Rad) with bovine serum albumin as the standard (Bradford, 1976).

2.2. Crystallization

TthGK/HK was crystallized using the hanging-drop vapour-diffusion method with reservoir solution comprising 0.1 *M* Tris–HCl pH 8.5 and 8% (*w/v*) polyethylene glycol 8000. 1.5 µl protein sample (20 mg ml⁻¹ in 20 mM Tris–HCl pH 7.5) was mixed with an equal volume of reservoir solution and left to equilibrate against 400 µl reservoir solution at 293 K.

2.3. Data collection and processing

A single crystal was dipped into reservoir solution supplemented with 25% glycerol for a few seconds and cooled to 100 K in a stream of nitrogen gas. Diffraction data were collected on BL38B1, SPring-8, Hyogo, Japan using an ADSC Quantum 315 detector. The crystal-to-detector distance was 290 mm. The oscillation range was 1° per frame.

**Figure 2**

Crystals of TthGK/HK.

240 frames of diffraction data were integrated and scaled using the HKL-2000 suite (Otwinowski & Minor, 1997). The self-rotation function was calculated using the MOLREP program (Vagin & Teplyakov, 2010) from the CCP4 suite (Winn *et al.*, 2011).

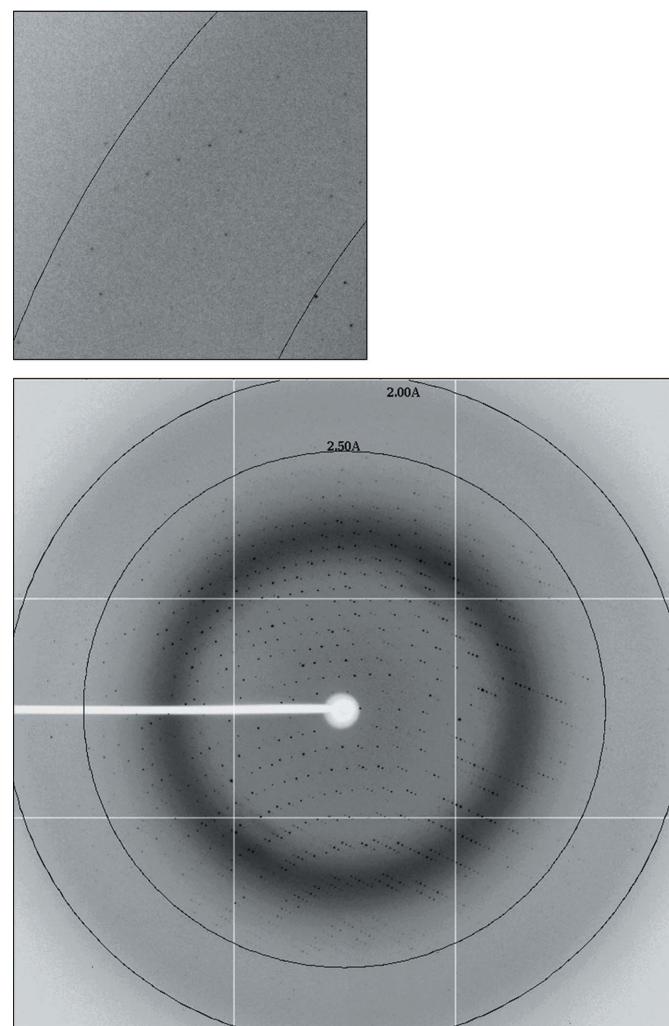
2.4. Gel filtration

Analytical gel-filtration chromatography was performed using a Superdex 200 GL (10/30) column (GE Healthcare) with a buffer consisting of 20 mM Tris–HCl pH 8.0 and 150 mM NaCl. The flow rate was 0.5 ml min⁻¹.

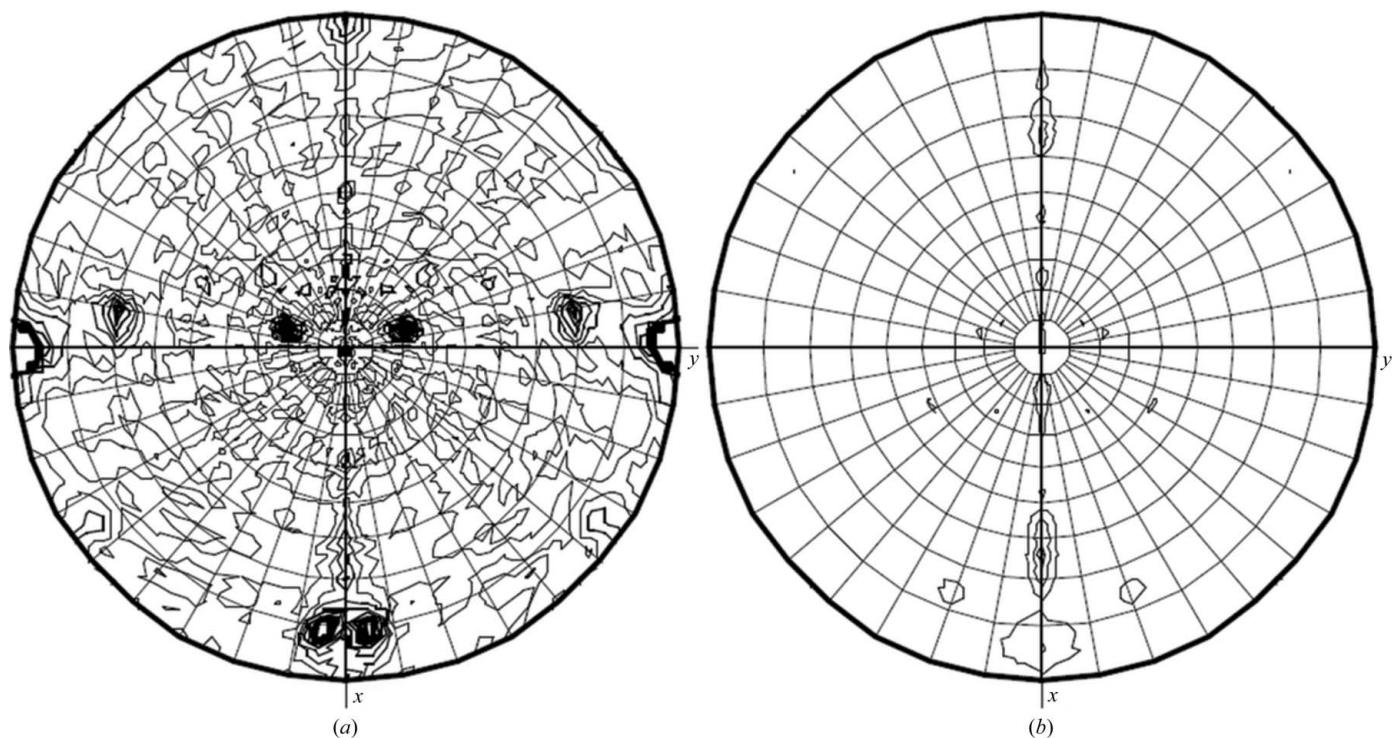
3. Results and discussion

100 mg of TthGK/HK was obtained from 1 l of culture. Crystals of dimensions 0.1 × 0.1 × 0.1 mm were obtained by the vapour-diffusion method within a week (Fig. 2). An X-ray diffraction data set was collected to 2.02 Å resolution with an overall completeness of 99.7% (Fig. 3). The space group of the crystals was $P2_1$. The detailed conditions and measurements are summarized in Table 1.

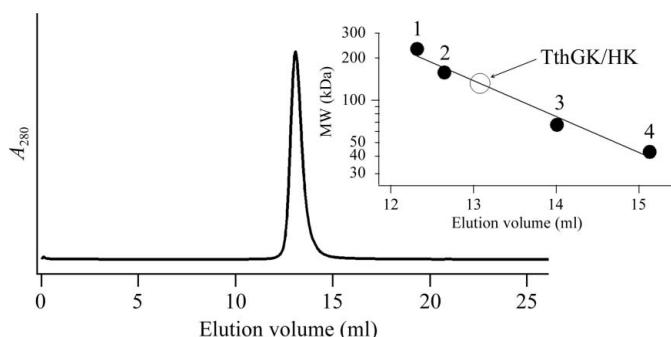
Assuming the presence of four polypeptides in the asymmetric unit, the crystal volume per unit protein weight (V_M) and solvent

**Figure 3**

Diffraction image of TthGK/HK. Circles corresponding to 2.0 and 2.5 Å resolution are shown.

**Figure 4**

Self-rotation function calculated in the (a) $\chi = 180^\circ$ and (b) $\chi = 90^\circ$ sections.

**Figure 5**

Gel-filtration chromatogram of TthGK/HK and calibration curve (inset). The standard proteins are (1) catalase, (2) aldolase, (3) albumin and (4) ovalbumin. The estimated molecular mass of TthGK/HK is 131.9 kDa, while the calculated monomer mass is 31 479 Da.

content were calculated to be $2.91 \text{ \AA}^3 \text{ Da}^{-1}$ and 57.8%, respectively. Fig. 4 shows the results of the self-rotation function in the $\chi = 180^\circ$ and $\chi = 90^\circ$ sections. While three twofold axes were clearly observed, no fourfold axis was found. These results support a tetrameric assembly of TthGK/HK molecules in the crystal; these molecules have 222 symmetry. To confirm the molecular assembly in solution, gel-filtration chromatography was performed (Fig. 5). The results showed a tetrameric assembly of TthGK/HK in solution (see legend to Fig. 5). However, structure solution requires phase determination and this is currently in progress.

The diffraction data were collected on BL38B1 at SPring-8, Hyogo, Japan with the approval of the Japan Synchrotron Radiation Research Institute (approval No. 2011A1986).

References

- Bae, J., Kim, D., Choi, Y., Koh, S., Park, J. E., Kim, J. S., Moon, S. H., Park, B.-H., Park, M., Song, H.-E., Hong, S.-I. & Lee, D.-S. (2005). *Biochem. Biophys. Res. Commun.* **334**, 754–763.
- Bradford, M. M. (1976). *Anal. Biochem.* **72**, 248–254.
- Dörr, C., Zaparty, M., Tjaden, B., Brinkmann, H. & Siebers, B. (2003). *J. Biol. Chem.* **278**, 18744–18753.
- Hansen, T., Reichstein, B., Schmid, R. & Schönheit, P. (2002). *J. Bacteriol.* **184**, 5955–5965.
- Hansen, T. & Schönheit, P. (2003). *FEMS Microbiol. Lett.* **226**, 405–411.
- Higgins, D. G., Thompson, J. D. & Gibson, T. J. (1996). *Methods Enzymol.* **266**, 383–402.
- Hofmann, K., Bucher, P., Falquet, L. & Bairoch, A. (1999). *Nucleic Acids Res.* **27**, 215–219.
- Ito, S., Fushinobu, S., Jeong, J.-J., Yoshioka, I., Koga, S., Shoun, H. & Wakagi, T. (2003). *J. Mol. Biol.* **331**, 871–883.
- Ito, S., Fushinobu, S., Yoshioka, I., Koga, S., Matsuzawa, H. & Wakagi, T. (2001). *Structure*, **9**, 205–214.
- Lunin, V. V., Li, Y., Schrag, J. D., Iannuzzi, P., Cygler, M. & Matte, A. (2004). *J. Bacteriol.* **186**, 6915–6927.
- Mesak, L. R., Mesak, F. M. & Dahl, M. K. (2004). *BMC Microbiol.* **4**, 6.
- Otwowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Sakuraba, H., Goda, S. & Ohshima, T. (2004). *Chem. Rec.* **3**, 281–287.
- Sakuraba, H., Yoshioka, I., Koga, S., Takahashi, M., Kitahama, Y., Satomura, T., Kawakami, R. & Ohshima, T. (2002). *J. Biol. Chem.* **277**, 12495–12498.
- Tsuge, H., Sakuraba, H., Kobe, T., Kujime, A., Katunuma, N. & Ohshima, T. (2002). *Protein Sci.* **11**, 2456–2463.
- Vagin, A. & Teplyakov, A. (2010). *Acta Cryst. D66*, 22–25.
- Winn, M. D. *et al.* (2011). *Acta Cryst. D67*, 235–242.