Chongyun Cheng,^{a,b,†} Jianhua Gu,^{c,†} Jing Su,^b Wei Ding,^d Jie Yin,^b Wenguang Liang,^b Xiaoxia Yu,^b Jun Ma,^a Peng George Wang,^e Zhicheng Xiao^{a,*} and Zhi-Jie Liu^{b,*}

^aDepartment of Anatomy and Developmental Biology, Monash University, Clayton, Melbourne 3800, Australia, ^bNational Laboratory of Macrobiochemistry, Institute of Biophysics, Chinese Academy of Sciences, 15 Datun Road, Beijing 100101, People's Republic of China, ^cDepartment of General Surgery, Tianjin First Center Hospital, Tianjin 300192, People's Republic of China, ^dCenter for Biological Imaging, Institute of Biophysics, Chinese Academy of Sciences, 15 Datun Road, Beijing 100101, People's Republic of China, and ^eDepartments of Chemistry and Biochemistry, The Ohio State University, Columbus, OH 43210, USA

† These authors contributed equally to this work.

Correspondence e-mail: zhicheng.xiao@monash.edu, zjliu@ibp.ac.cn

Received 4 February 2014
Accepted 1 June 2014

© 2014 International Union of Crystallography
All rights reserved

Crystallization, preliminary X-ray crystallographic and cryo-electron microscopy analysis of a bifunctional enzyme fucokinase/L-fucose-1-P-guanylyltransferase from *Bacteroides fragilis*

Fucokinase/L-fucose-1-P-guanylyltransferase (FKP) is a bifunctional enzyme which converts L-fucose to Fuc-1-P and thence to GDP-L-fucose through a salvage pathway. The molecular weights of full-length FKP (F-FKP) and C-terminally truncated FKP (C-FKP, residues 300–949) are 105.7 and 71.7 kDa, respectively. In this study, both recombinant F-FKP and C-FKP were expressed and purified. Size-exclusion chromatography experiments and analytical ultracentrifugation results showed that both F-FKP and C-FKP are trimers. Native F-FKP protein was crystallized by the vapour-diffusion method and the crystals belonged to space group $P2_12_12_1$ and diffracted synchrotron X-rays to 3.7 Å resolution. The crystal unit-cell parameters are $a = 91.36$, $b = 172.03$, $c = 358.86$ Å, $\alpha = \beta = \gamma = 90.00^\circ$. The three-dimensional features of the F-FKP molecule were observed by cryo-EM (cryo-electron microscopy). The preliminary cryo-EM experiments showed the F-FKP molecules as two parallel disc-shaped objects stacking together. Combining all results together, it is assumed that there are six FKP molecules in one asymmetric unit, which corresponds to a calculated Matthews coefficient of $2.19 \text{ Å}^3 \text{ Da}^{-1}$ with 43.83% solvent content. These preliminary crystallographic and cryo-EM microscopy analyses provide basic structural information on FKP.

1. Introduction

Bacteroides fragilis fucokinase/L-fucose-1-P-guanylyltransferase (FKP) is a bifunctional enzyme which converts L-fucose to Fuc-1-P and thence to GDP-L-fucose through a salvage pathway. The salvage pathway is an alternative pathway (compared with a *de novo* pathway) for GDP-L-fucose biosynthesis which is enabled by two enzymes: fucose kinase and GDP-fucose pyrophosphorylase. This pathway was initially found only in mammalian cells (Becker & Lowe, 2003). FKP was the first bifunctional enzyme identified from *B. fragilis* by the Comstock group that can mimic the mammalian salvage pathway to synthesize GDP-L-fucose by using exogenously acquired L-fucose (Coyne *et al.*, 2005). This transformation is conserved in all *Bacteroides* species and the bifunctional enzyme FKP has been efficiently used to synthesize tailored and modified sugar substrates and derivatives *in vitro* (Yi *et al.*, 2009; Wang *et al.*, 2009).

F-FKP contains 949 amino acids with a molecular weight of 105.7 kDa. Sequence alignment shows that the FKP N-terminus (amino acids 1–430) shares 20% amino-acid identity with the human GDP-L-fucose pyrophosphorylase and the C-terminus (amino acids 584–949) is similar to mammalian L-fucokinases. The function of the connecting 153-amino-acid linker (431–583 amino acids) is unknown. The three-dimensional structure of FKP is not yet available; therefore we decided to study the three-dimensional structure to delineate the molecular catalytic mechanism of FKP.

2. Materials and methods

2.1. Gene cloning and protein expression and purification

2.1.1. Cloning. The *fkp* gene was amplified from pMCSG7-FKP (Yi *et al.*, 2009) and was cloned into a modified vector named pChis by

the ligation-independent cloning (LIC) method. The pChis vector was designed in-house for LIC use with an 8×His tag followed by a TEV protease cleavage site at the C-terminus.

The C-terminally truncated *fkp* gene (encoding amino-acid residues 300–949) was re-cloned into the LIC vector pMSGC7 (Stols *et al.*, 2002) with an N-terminal 6×His tag followed by a TEV protease cleavage site.

2.1.2. Expression. The expression plasmid was validated by sequencing and was then transformed into *Escherichia coli* BL21 (DE3) RIPL (Stratagene) strain for protein production. Cells were grown in Luria Broth (LB) medium with 100 µg ml⁻¹ ampicillin at 37°C until the optical density of the culture reached an OD_{600 nm} of 0.8; they were then induced by isopropyl β-D-1-thiogalactopyranoside (IPTG) with a final concentration of 0.2 mM at 16°C for a further 20 h. Cells were harvested by centrifugation at 4000 rev min⁻¹ for 30 min directly for use in purification or were frozen at ~-80°C.

2.1.3. Purification. Cells were lysed by sonication and then centrifuged at 16 000 rev min⁻¹ for 30 min. The clarified supernatant was subjected to a nickel-chelating affinity column (GE Healthcare). His-tagged protein was eluted using PBS buffer (50 mM Na₂HPO₄, 10 mM KH₂PO₄, 137 mM NaCl, 2.7 mM KCl pH 7.4) containing 300 mM imidazole, 10% glycerol. The eluted protein was exchanged to PBS buffer by centrifugation using Amicon Ultra-15 centrifugal filter units (Millipore) and then subjected to TEV protease treatment at 4°C overnight to remove the His tag. Uncut protein and TEV protease were removed by a second round of Ni-affinity chromatography. The tag-free protein was concentrated by centrifugation using Amicon Ultra-15 centrifugal filter units (Millipore). The concentrated protein was loaded onto a Superdex G200 gel-filtration column (GE Healthcare) previously equilibrated with 20 mM Tris-HCl pH 8.0, 200 mM NaCl, 10 mM dithiothreitol, 10% glycerol.

The size-exclusion chromatography profile showed a polymeric state for both the F-FKP and C-FKP proteins with target elution peaks at 52.27 ml (F-FKP) and 58.95 ml (C-FKP). Therefore, the oligomeric protein fractions were collected for further analytical ultracentrifugation, mass-spectrometric characterization and crystallographic studies.

2.2. Protein characterization

2.2.1. Analytical ultracentrifugation. Analytical sedimentation-velocity experiments were performed on a ProteomeLab XL-I protein characterization system (Beckman Coulter) at 20°C. Protein samples were diluted with buffer (20 mM Tris, 100 mM NaCl pH 7.5) to 400 µl at a concentration of about 0.5 mg ml⁻¹. Samples were loaded into a conventional double-sector quartz cell, mounted in a Beckman four-hole An-60 Ti rotor and centrifuged at 60 000 rev min⁻¹. Absorbance was read at a wavelength of 280 nm. Data were calculated and analyzed using the *SEDFIT* software (<http://www.analyticalultracentrifugation.com>).

2.2.2. Matrix-assisted laser desorption/ionization–time-of-flight mass spectrum (MALDI-TOF MS). The protein molecular weight was measured by MALDI-TOF MS (Shimadzu AximaCRF Plus). The protein sample was processed with Zip-Tip C18 pipette tips (Millipore) for desalting prior to MS analysis. The tip was washed with buffer consisting of 75% acetonitrile, 0.1% trifluoroacetic acid and then equilibrated with equilibration buffer consisting of 5% methanol, 0.1% trifluoroacetic acid. Approximately 5 µg protein was applied and rinsed with equilibration buffer. The protein was eluted into 5 µl elution buffer consisting of 75% acetonitrile and 0.1% trifluoroacetic acid. Sinapinic acid was used for MS analysis. The data

were observed using the *Kompect* software (<http://www.shimadzu.com>).

2.2.3. N-terminal sequencing. The protein was firstly fractionated by SDS-PAGE (10%) and then transferred electrophoretically to a polyvinylidene difluoride (PVDF) membrane in an ice bath using a Bio-Rad machine for 1 h (300 mA). The stable band after degradation was visualized by staining buffer (0.1% Coomassie Brilliant Blue R-250 in 1.0% acetic acid, 40% methanol) and cut out to dry manually. The PVDF membrane containing target protein band was digested and extracted. Finally the protein was sequenced on an automated protein sequencer (ABI Procise 491) to recognize the first four or five amino acids at the N-terminus.

2.3. Crystallization and optimization

The purified proteins were concentrated to 30 mg ml⁻¹, before setting up for crystallizations drops. Initial crystallization experiments were performed manually by the hanging-drop vapour-diffusion method at 4, 16 and 22°C. A total of 500 different conditions from commercially available sparse-matrix screens (Hampton Research and Emerald Bio) were used for screening at each temperature. Crystallization drops comprised 1 µl protein solution mixed with 1 µl reservoir solution and were equilibrated over 300 µl reservoir solution. The protein crystallized after 2 d at 22°C, one week at 16°C and one month at 4°C. The shape of the crystals was diverse at the different crystallization temperatures. A total of 16 conditions yielded native F-FKP protein crystals. Interestingly, all of the crystals grew using the same precipitant, 20% (w/v) PEG 3350, with slight differences in buffer pH and salt concentration, including (i) 0.2 M ammonium sulfate, 0.1 M bis-tris pH 6.5, (ii) 0.2 M ammonium sulfate, 0.1 M HEPES pH 7.5, (iii) 0.2 M lithium sulfate monohydrate, 0.1 M bis-tris pH 6.5, (iv) 0.2 M lithium sulfate monohydrate, 0.1 M HEPES pH 7.5, (v) 0.2 M lithium sulfate monohydrate, 0.1 M Tris pH 8.5, (vi) 0.2 M potassium sodium tartrate tetrahydrate, (vii) 0.2 M sodium citrate tribasic dihydrate pH 8.3, (viii) 0.2 M sodium malonate pH 6.0, (ix) 8% (v/v) Tacsimite pH 6.0, (x) 8% (v/v) Tacsimite pH 7.0, (xi) 8% (v/v) Tacsimite pH 8.0, (xii) 0.2 M succinic acid pH 7.0, (xiii) 0.2 M DL-malic acid pH 7.0, (xiv) 0.2 M ammonium tartrate dibasic

Figure 1 Identification of the stable fragment of the FKP protein. (a) SDS-PAGE characterization of autodegraded FKP; the sample was collected after incubation for 2 d at 25°C. The gel was then transferred electrophoretically to PVDF membrane for further N-terminal sequencing experiments and the results showed that the degraded fragment had the sequence ³⁰⁰K-V-K-P³⁰³. (b) MALDI-TOF MS measurement of recombinant C-FKP protein. The measured molecular weight of C-FKP was 72 064 Da. The calculated molecular weight is 72 065 Da (amino acids 300–949 plus three amino acids S-N-A encoded by the vector), indicating a deviation of only 1 Da.

Table 1

Data-collection and refinement statistics of full-length native FKP.

X-ray source	22-ID, APS
Crystal-to-detector distance (mm)	400
No. of images	417
Oscillation width (°)	0.5
Wavelength (Å)	0.9794
Space group	$P2_12_12_1$
Unit-cell parameters (Å, °)	$a = 91.36, b = 172.03, c = 358.86,$ $\alpha = \beta = \gamma = 90.00$
Mosaicity (°)	0.91
Matthews calculations	
Matthews coefficient (Å ³ Da ⁻¹)	2.19
Molecules in asymmetric unit	6
Solvent content (%)	43.83
Probability (3.74 Å)	0.39
Probability (total)	0.45
Resolution range (Å)	48.22–3.74 (3.94–3.74)
R_{merge} (%)	12.6 (53.9)
Mean $I/\sigma(I)$	20.37 (4.73)
Completeness (%)	96.6 (96.3)
Multiplicity	(6.4)

pH 7.0, (xv) 0.1 M sodium citrate tribasic dihydrate pH 5.6 and (xvi) 1% (w/v) tryptone, 0.05 M HEPES sodium pH 7.0. C-FKP protein crystallized under similar conditions.

2.4. X-ray diffraction data collection

Initial crystal diffraction-quality screening was performed at the home laboratory X-ray facility (Rigaku MicroMax CCD007). Crystals were transferred into mother-liquor solution supplemented with 20% (v/v) glycerol for 15 s, placed into a loop and then directly mounted on the goniometer pre-chilled in a liquid nitrogen-gas stream at -180°C . Although F-FKP and C-FKP native protein crystals were obtained in quite a few conditions, their diffraction was weak, with high resolution ranging from 4.0 to 9.0 Å. The maximum resolution of 3.7 Å was obtained from an F-FKP native protein crystal obtained using 20% (w/v) PEG 3350, 200 mM trisodium citrate

Figure 2

Size-exclusion chromatography profile of F-FKP and C-FKP. Proteins were run on a HiLoad Superdex 200 gel-filtration column equilibrated with 20 mM Tris-HCl pH 8.0, 200 mM NaCl, 10 mM dithiothreitol, 10% glycerol. The results revealed that the F-FKP and C-FKP proteins exist as polymers and the molecular weights of the target peak fall between 440 and 158 kDa. (a) F-FKP; the first peak eluted in the void volume and the homogeneous protein peak eluted at 52.27 ml. (b) C-FKP; the protein peak eluted at 58.95 ml.

pH 8.3. The data were collected on beamline 22-ID of the Advanced Photon Source (APS), Argonne National Laboratory, USA. A total of 417 raw diffraction images with an oscillation range of 0.5° were indexed and scaled using the *HKL-2000* software package (Otwinowski & Minor, 1997). The crystal belonged to space group $P2_12_12_1$, with unit-cell parameters $a = 91.36, b = 172.03, c = 358.86$ Å, $\alpha = \beta = \gamma = 90.00^\circ$. The Matthews coefficient was calculated using *MATTHEWS_COEF* from the *CCP4* program suite (Matthews, 1968; Kantardjieff & Rupp, 2003; Winn *et al.*, 2011). Even though five, six and seven molecules in one asymmetric unit yield reasonable Matthews coefficient values, when taking into account the information from size-exclusion chromatography experiments and subsequent cryo-EM (cryo-electron microscopy) analysis, we assume that there are six FKP molecules in one asymmetric unit, which corresponds to a calculated Matthews coefficient of $2.19 \text{ Å}^3 \text{ Da}^{-1}$ with 43.83% solvent content. The diffraction data set has a resolution range of 48.2–3.7 Å with 96.6% completeness as shown in Table 1.

2.5. Cryo-EM data collection

Before the cryo-EM experiment the cooled, purified F-FKP protein was exchanged into 20 mM Tris pH 7.5, 100 mM NaCl buffer using a Superdex 200 size-exclusion column (GE Healthcare). The gel-filtration profile showed an earlier eluting peak compared with the fresh protein, which may indicate a change to a higher oligomeric order during cooling. The concentration of a 1 ml fraction direct from the elution peak was measured as 2.7 mg ml^{-1} using a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific); this fraction was subjected to cryo-EM. Cryo-EM grids were prepared with 3 µl

Figure 3

Analytical ultracentrifugation analysis of F-FKP and C-FKP. Sedimentation-velocity experiments performed using analytical ultracentrifugation suggested that both F-FKP and C-FKP are trimers in solution. (a) F-FKP; average molecular weight of 315.25 kDa with a standard deviation of 15.96 kDa. (b) C-FKP; average molecular weight of 205.90 kDa with a standard deviation of 14.90 kDa.

Figure 4

(a) Typical F-FKP native protein crystals from the crystallization condition 20% (w/v) PEG 3350, 200 mM trisodium citrate pH 8.3. The approximate dimensions of each plate-like crystal are about 150–300 μm in length and 40–50 μm in width. (b) A typical diffraction pattern of the native F-FKP crystal. The exposure time was 5 s, the crystal-to-detector distance was 400 mm and the oscillation range per frame was 0.5°. The diffraction images were collected on APS beamline 22-ID. (c) A typical cryo-EM image of the F-FKP native protein. The size of the F-FKP particle is about 10 nm and consists of two parallel objects.

sample per grid. The grid was flash-cooled by plunging it into liquid ethane. Cryo-electron micrographs were collected using an FEI Titan Krios cryo-transmission electron microscope (300 kV FEG). A Gatan Ultrascan 895 4k × 4k CCD camera was used for recording images with a magnification of 96 000×. An electron dose of 20–30 e Å⁻² was used and the defocus values of the micrographs ranged from 1.5 to 3.0 μm. Particle images were selected from micrographs using *EMAN* (Ludtke *et al.*, 1999). Unfortunately, the number of particles was not sufficient for three-dimensional reconstruction.

3. Results and discussion

The initial expression construct was pMCSG7-F-FKP, which encodes F-FKP with an N-terminal 6×His tag. However, the His tag could not be removed by Ni-affinity chromatography when subjected to TEV protease treatment. In order to obtain tag-free protein for use in crystallization, a series of vectors were screened (based on an *E. coli* expression system). Finally, an in-house-modified LIC vector pChis proved to be suitable for this study. The new construct pChis-F-FKP led to a higher protein production yield and a significantly improved purity. The His tag on the F-FKP protein can be removed completely after two rounds of purification by Ni-affinity chromatography.

Preliminary stability tests showed that F-FKP tends to degrade at room temperature to a more stable fragment, as shown in Fig. 1(a). Subsequent N-terminal sequencing of autodegraded FKP identified the C-terminal stable fragment of the FKP protein as ³⁰⁰K-V-K-P³⁰³. C-FKP covers the whole fucokinase domain, the linker and part of the pyrophosphorylase domain. Therefore, pMCSG7-C-FKP was constructed for expression, purification and crystallization. The subsequent experiments confirmed soluble protein expression of C-FKP. Molecular-mass measurement experiments by MALDI-TOF MS showed a molecular weight of 72.06 kDa (Fig. 1b).

The sequence-alignment result shows FKP has a kinase domain and a pyrophosphorylase domain with a 153-amino-acid linker. Therefore, some domain-based truncations with different lengths were designed according to the alignment information. However, initial protein-expression screening showed that none of the domain truncations were soluble, which probably indicated that the kinase domain and a pyrophosphorylase domain of FKP may interact closely, and that the linker between these two domains may be important for its structural stability.

Gel-filtration chromatography showed that both F-FKP and C-FKP appear as oligomers in solution (Fig. 2). The size of target peaks fall into a molecular-weight window from 440 to 158 kDa when compared with the gel-filtration calibration proteins ferritin (440 kDa, peak volume covers 50–60 ml) and aldolase (158 kDa, peak volume covers 60–70 ml).

SDS-PAGE and MALDI-TOF MS results showed that F-FKP and C-FKP have molecular weights of 105.99 kDa (data not shown) and 72.06 kDa, respectively. Analytical ultracentrifugation experiment results showed the average molecular weights of F-FKP and C-FKP are about 315.25 and 205.90 kDa, respectively, which indicated trimeric FKP (Fig. 3) and were in agreement with the results of the gel-filtration chromatography experiments.

We tried to crystallize F-FKP and C-FKP proteins at three different temperatures of 4, 16 and 25°C. Crystals formed at lower temperatures showed a better shape and a slower growth rate. However, the diffraction quality was not improved (around 4.0–9.0 Å) and only shows slight differences at different temperatures.

In order to improve crystal diffraction quality, several salvage methods were tried including crystal dehydration (Heras & Martin,

2005), *in situ* limited proteolysis (Dong *et al.*, 2007) and macro/microseeding (Zhu *et al.*, 2005). Unfortunately, none of them led to significant improvement. In summary, the best resolution was 3.7 Å for data collected from a native F-FKP protein crystal using synchrotron X-rays, shown in Fig. 4. In addition, crystallization of selenomethionine-labelled FKP and chemically modified FKP (Shaw *et al.*, 2007) was also attempted. Although most of the above protein samples could crystallize, none of them diffracted beyond 5.0 Å resolution.

In summary, the gel-filtration chromatography and analytical ultracentrifugation experiments indicated that both F-FKP and C-FKP are trimeric molecules. The cryo-EM study on F-FKP particles showed two parallel stacked objects with a diameter of 10 nm from the preliminary size and shape measurements. The preliminary X-ray crystallographic results showed six F-FKP molecules in one asymmetric unit, indicating that the massive F-FKP oligomers were involved in crystal packing during the crystallization process. Combined with the above results, we propose that F-FKP and C-FKP are trimeric in solution and that the structure of the higher oligomer of F-FKP may be constructed by two trimers of FKP.

This study was supported by a National Health and Medical Research Council of the Australian Government Early Career Fellowship (NHMRC-ECF-APP1054172). This work was also funded by the Ministry of Sciences and Technology of China (grants

2014CB910400 and 2013ZX10004602). We thank Professor Fei-Sun and Dr Gang-Ji at the Institute of Biophysics, Chinese Academy of Sciences for cryo-EM image collection and analysis. We thank Ms Megan Cross for editing the manuscript.

References

- Becker, D. J. & Lowe, J. B. (2003). *Glycobiology*, **13**, 41R–53R.
- Coyne, M. J., Reinap, B., Lee, M. M. & Comstock, L. E. (2005). *Science*, **307**, 1778–1781.
- Dong, A. *et al.* (2007). *Nature Methods*, **4**, 1019–1021.
- Heras, B. & Martin, J. L. (2005). *Acta Cryst.* **D61**, 1173–1180.
- Kantardjieff, K. A. & Rupp, B. (2003). *Protein Sci.* **12**, 1865–1871.
- Ludtke, S. J., Baldwin, P. R. & Chiu, W. (1999). *J. Struct. Biol.* **128**, 82–97.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Shaw, N., Cheng, C. & Liu, Z.-J. (2007). *Protoc. Exchange*, doi: 10.1038/nprot.2007.287.
- Stols, L., Gu, M., Dieckman, L., Raffin, R., Collart, F. R. & Donnelly, M. I. (2002). *Protein. Expr. Purif.* **25**, 8–15.
- Wang, W., Hu, T., Frantom, P. A., Zheng, T., Gerwe, B., Del Amo, D. S., Garret, S., Seidel III, R. D. & Wu, P. (2009). *Proc. Natl Acad. Sci. USA*, **106**, 16096–16101.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.
- Yi, W., Liu, X., Li, Y., Li, J., Xia, C., Zhou, G., Zhang, W., Zhao, W., Chen, X. & Wang, P. G. (2009). *Proc. Natl Acad. Sci. USA*, **106**, 4207–4212.
- Zhu, D.-Y., Zhu, Y.-Q., Xiang, Y. & Wang, D.-C. (2005). *Acta Cryst.* **D61**, 772–775.