Expression, purification and preliminary crystallographic studies of the C-terminal SH3 domain of human Tks4

Yuxin Huang,[‡] Huolian Qian,[‡]
Xiaoying Wang, Zhong Cheng,
Jixia Ren, Weichen Zhao and
Yong Xie*

Institute of Medicinal Plant Development,
Chinese Academy of Medical Sciences and
Peking Union Medical College, 151 Malianwa
North Road, Haidian District, Beijing 100193,
People's Republic of China

[‡] These authors contributed equally to this study.

Correspondence e-mail: yxie@implad.ac.cn

Received 28 November 2013
Accepted 27 January 2014

The Src homology 3 (SH3) domain is a small, noncatalytic domain with a conserved sequence of about 60 amino-acid residues that interacts with proline-rich peptides to form a protein complex. In this study, the C-terminal SH3 domain of human Tks4 (residues 853–911) was expressed, purified and crystallized. X-ray diffraction data were collected to 2.3 Å resolution. The crystal belonged to the trigonal space group $P3_121$ (or $P3_221$), with unit-cell parameters $a = b = 83.87$, $c = 108.44$ Å, $\alpha = \beta = 90$, $\gamma = 120^\circ$. Calculating the self-rotation and the native Patterson function did not lead to the detection of any noncrystallographic translational symmetry. Six, seven or eight protein molecules are likely to be present in the asymmetric unit, resulting in a Matthews coefficient and approximate solvent content of 2.71 Å³ Da⁻¹ and 55%, 2.32 Å³ Da⁻¹ and 47%, and 2.03 Å³ Da⁻¹ and 39%, respectively. To solve the crystal structure of the C-terminal SH3 domain of human Tks4, the isomorphous replacement method is presently being utilized.

1. Introduction

The human gene *factor for adipocyte differentiation 49* (*fad49*) codes for a protein FAD49 which contains one polypeptide chain of 911 amino-acid residues (Hishida *et al.*, 2008). FAD49 contains one phospho (PX) domain and four Src homology 3 (SH3) domains. This protein is also referred to as PX and SH3 domain-containing protein 2B (NCBI reference sequence NP_001017995.1) or Tks4 (Buschman *et al.*, 2009). Tks4 are critical proteins involved in the regulation of many physiological functions. Lack of Tks4 resulted in incomplete podosome formation and inhibited extracellular matrix degradation (Buschman *et al.*, 2009). Tks4 has also been implicated in the production of reactive oxygen species by tumour cells (Gianni *et al.*, 2009, 2010, 2011) and in the differentiation of white adipose tissue (Hishida *et al.*, 2008). In humans, Tks4 deficiency is responsible for the development of Frank-ter Haar syndrome (Iqbal *et al.*, 2010). In response to epidermal growth-factor (EGF) treatment, Tks4 is tyrosine-phosphorylated and associated with the EGF receptor, providing a new mechanism of regulating cell migration. Tks4 can be targeted in the regulation of tumour-cell migration (Bögel *et al.*, 2012).

In quiescent cells, Tks4 is almost completely present in the cytoplasm. In response to growth-factor treatment, Tks4 is translocated to the plasma membrane through at least two independent sites: an Src-binding site and its lipid-binding PX domain. The PX domain may recognize lipid products of phosphoinositide 3 kinase (PI 3-kinase) in the membrane. PI 3-kinase catalyzes the production of phosphatidylinositol-3,4,5-trisphosphate in cell-survival pathways. Signalling pathways downstream of PI 3-kinase affect cell growth, cell survival and cell migration. Therefore, PI 3-kinase has been implicated in human diseases, including diabetes and cancer (Cantley, 2002). At the plasma membrane, Tks4 is phosphorylated by Src kinase in a tyrosine-kinase-dependent manner. The role of Tks4 tyrosine phosphorylation is yet unknown. Finally, Tks4 contributes to podosome formation and the actin cytoskeleton (Bögel *et al.*, 2012). SH3 domains are known as small noncatalytic domains with a conserved sequence of about 60 amino-acid residues that interacts with proline-rich peptides to form a protein complex. SH3 domains are involved in cell–cell communication and signal transduction from the cell surface

to the nucleus, and they also mediate protein–protein interactions (Cicchetti *et al.*, 1992; Ren *et al.*, 1993; Mayer, 2001).

The amino- and carboxy-termini of the SH3 domain are located in close proximity to one another, suggesting that this domain can extend from the surface of a protein without perturbing the adjacent structure (Yu *et al.*, 1992). In human Tks4, the PX domain is located at the N-terminus and the fourth SH3 domain (residues 853–911) is located at the C-terminus. The PX domain binds to the third SH3 domain so that the fourth SH3 domain located at the C-terminus of Tks4 is kept distant from the PX domain and the first to third SH3 domains and extends from the surface of human Tks4 (Hishida *et al.*, 2008). Usually, SH3 domains have a characteristic β -barrel fold which consists of five (or six) β -strands arranged as two tightly packed antiparallel β -sheets (Whisstock & Lesk, 1999; Martin-Garcia *et al.*, 2012; Xiao *et al.*, 2013). However, a few SH3 domains have distinctive structures. For example, SH3 domain A of human p47phox could form a fourth and fifth β -strand swapped dimer structure (Yuzawa *et al.*, 2004; Groemping *et al.*, 2003). The C-terminal SH3 domain of human Tks4 shows 43% amino-acid sequence identity to the SH3 domain A of human p47phox (including residues 159–213). Whether the C-terminal SH3 domain of human Tks4 forms a dimer like the SH3 domain A of human p47phox is still unknown. In this study, crystallographic studies were carried out in order to understand the relationship between the structure and function of the C-terminal SH3 domain of human Tks4. The results of crystallization and preliminary crystallographic studies are reported in this paper.

2. Materials and methods

2.1. Protein expression and purification

The gene encoding the C-terminal SH3 domain of human Tks4 (NCBI reference sequence NP_001017995.1), including residues 853–911, was amplified by PCR using full-length human complementary DNA, which was cloned using reverse transcription of the messenger RNA extracted from HeLa cells using TRIzol reagent (Invitrogen) as a template. The forward primer sequence 5'-CGGGATCCTTGTATGTGGCCGTGGCCGAC-3' contains a *Bam*HI restriction site (bold). The reverse primer sequence 5'-GGAATTCCTACGGCTTC-TTTCTGAGATAGTTGG-3' contains an *Eco*RI restriction site (bold) and a stop codon (italics). The cDNA of the human Tks4 (including residues 853–911) was cloned and inserted into the pGEX-

4T-2 expression vector (Amersham, England) between the *Bam*HI restriction site and the *Eco*RI restriction site. The C-terminal SH3 domain of human Tks4 was expressed in *Escherichia coli* strain BL21(DE3) Codon Plus RIL cells (Stratagene, USA) as a GST-fusion protein. The *E. coli* cells were harvested by centrifugation, resuspended in PBS (140 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄, 1.8 mM KH₂PO₄, pH 7.3) and lysed by sonication.

After the cell debris had been removed by centrifugation at 35 000g for 20 min, the supernatant was applied onto a glutathione-Sepharose 4B column. The column was washed using PBS to remove proteins that were weakly bound to the column. Ten units of thrombin (GE Healthcare) were then added to cleave off the GST tag at 293 K overnight. While the GST tag remained bound to the column, the C-terminal SH3 domain of human Tks4 plus one glycine residue at the N-terminus were eluted in the PBS solution. Further purification of the C-terminal SH3 domain of human Tks4 was carried out at 277 K using an ÄKTAexplorer system (GE Healthcare). The SH3 domain was re-buffered into a solution consisting of 20 mM Tris–HCl pH 8.5, 10 mM NaCl, 2 mM β -mercaptoethanol using dialysis and purified by anion-exchange chromatography using a HiTrap Q column (5 ml). Further purification was performed by gel-filtration chromatography using a HiLoad 16/60 Superdex 75 prep-grade column. The column was equilibrated and the C-terminal SH3 domain of human Tks4 was eluted with a solution consisting of 20 mM Tris–HCl pH 8.0, 150 mM NaCl, 2 mM dithiothreitol.

2.2. Crystallization

The purified C-terminal SH3 domain of human Tks4 was concentrated to approximately 10 mg ml^{−1} using an Amicon Ultra-15 (MWCO 3000) filter unit (Millipore). The protein solvent used in the crystallization experiments is the same as that used in the gel-filtration chromatography. Crystallization was carried out using the sitting-drop vapour-diffusion method at 293 K. Each sitting drop consisted of 1 μ l protein solution and 1 μ l reservoir solution. The Crystal Screen HT, PEG/Ion and Index kits (Hampton Research) were used to establish initial crystallization conditions. Plate-shaped crystal clusters of the C-terminal SH3 domain of human Tks4 that included several single crystals with approximate dimensions of 0.15 $\times$ 0.1 $\times$ 0.05 mm were obtained using condition No. 32 [200 mM MgSO₄, 20% (v/v) PEG 3350] of the PEG/Ion kit within one week (Fig. 1).

2.3. Diffraction data collection

A single crystal with approximate dimensions of 0.15 $\times$ 0.1 $\times$ 0.05 mm was separated from the crystal clusters (Fig. 1) using a micro-tool (Hampton Research, catalogue No. HR4-817) and was swished about to remove the reservoir solution using a cryoprotectant solution consisting of 50% Paratone and 50% paraffin. The crystal was mounted on a rayon loop in a liquid-nitrogen gas stream at 100 K. X-ray diffraction data from the crystal were collected on beamline BL17U at the Shanghai Synchrotron Radiation Facility (SSRF), Shanghai, People's Republic of China. The wavelength, camera distance, oscillation range and exposure time were 1.000 Å, 200 mm, 1.0° and 8 s, respectively. The diffraction patterns were recorded on an ADSC Quantum 315r CCD camera. A complete data set was collected from 200 images, covering 200° in total. Diffraction intensity data were processed and scaled using the *HKL*-2000 program (Otwinowski & Minor, 1997).

Figure 1
Several single crystals of the C-terminal SH3 domain of human Tks4 with approximate dimensions of 0.15 $\times$ 0.1 $\times$ 0.05 mm formed a cluster of plate-shaped crystals.

3. Results and discussion

X-ray diffraction data were collected to 2.30 Å resolution. The crystal of the C-terminal SH3 domain belonged to the trigonal crystal system, with unit-cell parameters $a = b = 83.87$, $c = 108.44$ Å, $\alpha = \beta = 90$, $\gamma = 120^\circ$. Processing the X-ray diffraction data in space groups $P3_121$ or $P3_12$ resulted in an overall R_{merge} of 7.2 or 46.5%, respectively. The $0kl$ X-ray diffraction images displayed using the *phenix.data_viewer* program in the *PHENIX* suite (Adams *et al.*, 2010) revealed that the

Figure 2

Plot of the self-rotation function calculated using the X-ray diffraction data in the resolution range 20–4 Å in the $\kappa = 120^\circ$ section (a) and the $\kappa = 180^\circ$ section (b). The map is contoured from the 1.0σ level in steps of 1.0σ .

Table 1

Crystal parameters and X-ray diffraction data-collection statistics.

Values in parentheses are for the outermost resolution shell.

Space group	$P3_121$ (or $P3_12$)
Unit-cell parameters (Å, °)	$a = 83.87$, $b = 83.87$, $c = 108.44$ $\alpha = \beta = 90$, $\gamma = 120$
No. of protein molecules per asymmetric unit	6 (or 7 or 8)
V_M (Å ³ Da ⁻¹)	2.71 (or 2.32 or 2.03)
Solvent content (%)	55 (or 47 or 39)
Resolution range (Å)	50–2.30 (2.34–2.30)
Total reflections	162331
Total unique reflections	20187 (974)
Completeness (%)	100 (100)
$R_{\text{merge}}^\dagger$ (%)	7.2 (35.4)
Average multiplicity	8.0 (8.0)
$\langle I/\sigma(I) \rangle$	31.8 (4.3)

$^\dagger R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is the observed integrated intensity and $\langle I(hkl) \rangle$ is the average integrated intensity obtained from multiple measurements.

allowed reflections are 0, 0, $l = 3n$. Therefore, the space group of the C-terminal SH3 domain of human Tks4 crystal is $P3_121$ (or $P3_12$). With the glycine residue attached at the N-terminus, the C-terminal SH3 domain of human Tks4 has a calculated molecular weight of 6775.49 Da. The values of the Matthews coefficient and solvent content of this protein crystal (Matthews, 1968) calculated using the *MATTHEWS_COEF* program from the *CCP4* program suite (Winn *et al.*, 2011) suggest that six, seven or eight protein molecules are likely to be present in the asymmetric unit. The self-rotation function calculated using *MOLREP* (Vagin & Teplyakov, 2010) revealed no additional twofold and threefold symmetry peaks besides those due to crystallographic symmetry (Fig. 2). The *phenix.xtriage* program in the *PHENIX* suite (Adams *et al.*, 2010) shows that the largest off-origin peak in the native Patterson function is 4.6% of the height of the origin peak. These results suggest that no noncrystallographic translational symmetry can be detected. Therefore, six, seven or eight C-terminal SH3 domains are likely to be present in the asymmetric unit, resulting in a Matthews coefficient and approximate solvent content of 2.71 Å³ Da⁻¹ and 55%, 2.32 Å³ Da⁻¹ and 47%, 2.03 Å³ Da⁻¹ and 39%, respectively. The crystal parameters and X-ray diffraction data-collection statistics are shown in Table 1.

Molecular-replacement (MR) searches were performed using the *MrBUMP* program (Keegan & Winn, 2007) to solve the crystal structure of the C-terminal SH3 domain of human Tks4. *MOLREP* (Vagin & Teplyakov, 2010) and *Phaser* (McCoy *et al.*, 2007) and the crystal structures of SH3 domain A of human p47phox (residues 159–213; PDB entry 1ov3; Groemping *et al.*, 2003) were used to search for a solution; however, a correct solution could not be found. The following crystal structures of SH3 domains with more than 30% sequence identity to the C-terminal SH3 domain of human Tks4 were also used as MR search models: the SH3 domain of human tyrosine-protein kinase HCK (PDB entry 3reb; residues 79–138; 35.8% sequence identity; Breuer *et al.*, 2011), the SH3 domain of human proto-oncogene tyrosine-protein kinase Fyn (PDB entry 3h0i; residues 73–142; 34.0% sequence identity; A. Lugari, L. Ponchon, F. Hoh, C. Ktori, J. E. Ladbury, M.-P. Strub, Y. Collette, X. Morelli & S. T. Arold, unpublished work), the SH3 domain of human tyrosine-protein kinase SRC (PDB entry 2src; residues 84–140; 32.7% sequence identity; Xu *et al.*, 1999) and the solution NMR structure of the fifth SH3 domain of human KIAA0418 protein (PDB entry 2egc; residues 1072–1133; 56% sequence identity; RIKEN Structural Genomics/Proteomics Initiative, unpublished work). The crystal packing of each MR solution using each mentioned structure as a search model revealed that multiple molecules clash. As no correct MR solutions were found with these models, the isomorphous

replacement method is presently being utilized to solve the crystal structure of the C-terminal SH3 domain of human Tks4.

The authors are grateful to the staff for the use of beamline BL17U at the SSRF. This study was supported by the National Natural Science Foundation of China (No. 81273432), the Program for New Century Excellent Talents in University, the Fundamental Research Funds for the Central Universities and PUMC Youth Fund.

References

- Adams, P. D. *et al.* (2010). *Acta Cryst.* **D66**, 213–221.
- Bögel, G., Gujdár, A., Geiszt, M., Lányi, Á., Fekete, A., Sipeki, S., Downward, J. & Buday, L. (2012). *J. Biol. Chem.* **287**, 31321–31329.
- Breuer, S., Schievink, S. I., Schulte, A., Blankenfeldt, W., Fackler, O. T. & Geyer, M. (2011). *PLoS One*, **6**, e20033.
- Buschman, M. D., Bromann, P. A., Cejudo-Martin, P., Wen, F., Pass, I. & Courtneidge, S. A. (2009). *Mol. Biol. Cell*, **20**, 1302–1311.
- Cantley, L. C. (2002). *Science*, **296**, 1655–1657.
- Cicchetti, P., Mayer, B. J., Thiel, G. & Baltimore, D. (1992). *Science*, **257**, 803–806.
- Gianni, D., DerMardirossian, C. & Bokoch, G. M. (2011). *Eur. J. Cell Biol.* **90**, 164–171.
- Gianni, D., Diaz, B., Taulet, N., Fowler, B., Courtneidge, S. A. & Bokoch, G. M. (2009). *Sci. Signal.* **2**, ra54.
- Gianni, D., Taulet, N., DerMardirossian, C. & Bokoch, G. M. (2010). *Mol. Biol. Cell*, **21**, 4287–4298.
- Groemping, Y., Lapouge, K., Smerdon, S. J. & Rittinger, K. (2003). *Cell*, **113**, 343–355.
- Hishida, T., Eguchi, T., Osada, S., Nishizuka, M. & Imagawa, M. (2008). *FEBS J.* **275**, 5576–5588.
- Iqbal, Z. *et al.* (2010). *Am. J. Hum. Genet.* **86**, 254–261.
- Keegan, R. M. & Winn, M. D. (2007). *Acta Cryst.* **D63**, 447–457.
- Martin-Garcia, J. M., Luque, I., Ruiz-Sanz, J. & Camara-Artigas, A. (2012). *Acta Cryst.* **D68**, 1030–1040.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- Mayer, B. J. (2001). *J. Cell Sci.* **114**, 1253–1263.
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Ren, R., Mayer, B. J., Cicchetti, P. & Baltimore, D. (1993). *Science*, **259**, 1157–1161.
- Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* **D66**, 22–25.
- Whisstock, J. C. & Lesk, A. M. (1999). *Trends Biochem. Sci.* **24**, 132–133.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.
- Xiao, R., Xi, X.-D., Chen, Z., Chen, S.-J. & Meng, G. (2013). *Blood*, **121**, 700–706.
- Xu, W., Doshi, A., Lei, M., Eck, M. J. & Harrison, S. C. (1999). *Mol. Cell*, **3**, 629–638.
- Yu, H., Rosen, M. K., Shin, T. B., Seidel-Dugan, C., Brugge, J. S. & Schreiber, S. L. (1992). *Science*, **258**, 1665–1668.
- Yuzawa, S., Suzuki, N. N., Fujioka, Y., Ogura, K., Sumimoto, H. & Inagaki, F. (2004). *Genes Cells*, **9**, 443–456.