Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Crystallization and preliminary X-ray data collection of the $L27_{PATJ}$–$(L27N,L27C)_{Pals1}$–$L27_{MALS}$ tripartite complex

Jinxiu Zhang,[a,b] Xue Yang,[a,b]* Yuequan Shen[a,b] and Jiafu Long[a,b]

[a]State Key Laboratory of Medicinal Chemical Biology, Nankai University, Tianjin 300071, People's Republic of China, and [b]Department of Biochemistry and Molecular Biology, College of Life Science, Nankai University, Tianjin 300071, People's Republic of China

Correspondence e-mail: xueyangnk@yahoo.com

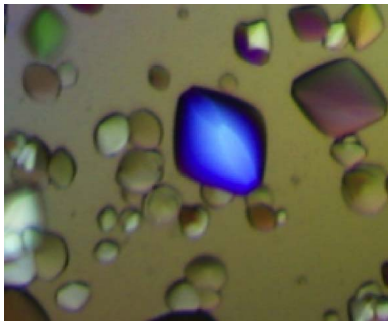
Received 11 August 2011
Accepted 12 September 2011

The L27 (LIN-2/LIN-7) domain is a protein–protein interaction module capable of assembling proteins into biologically important complexes. Pals1 contains two L27 domains: the first, L27N, interacts with PATJ, and the second, L27C, interacts with MALS, forming a tripartite complex that plays a crucial role in the establishment and maintenance of cell polarity. To provide a better understanding of the mechanism of assembly of this tripartite complex, four different $L27_{PATJ}$–$(L27N,L27C)_{Pals1}$–$L27_{MALS}$ constructs were cloned, expressed, purified and crystallized. Crystals of tripartite complex 1 of $L27_{PATJ}$–$(L27N,L27C)_{Pals1}$–$L27_{MALS}$ diffracted to 2.05 Å resolution. These crystals belonged to either space group $P6_122$ or $P6_522$, with unit-cell parameters $a = b = 145.2$, $c = 202.5$ Å. Assuming the presence of four molecules in the asymmetric unit, a Matthews coefficient of 2.69 Å^3 Da^{-1} was calculated, corresponding to a solvent content of 54.25%.

1. Introduction

The establishment and maintenance of epithelial cell polarity are crucial during epithelial differentiation, proliferation and morphogenesis in both unicellular and multicellular organisms. Disturbance of cell polarity may cause various diseases, including cancer, leprosy and atopic diseases (Iden & Collard, 2008; Kapsenberg *et al.*, 2000; Mitra *et al.*, 1999). The formation of epithelial cell polarity is mediated and regulated by the assembly of several evolutionarily conserved polarity complexes (Ohno, 2001; Pieczynski & Margolis, 2011). Two classes of scaffold-protein interactions mediate the assembly of these signalling complexes. One class is defined by specific interactions such as those between PDZ domains of scaffold proteins and short peptide motifs of nonscaffold molecules, including receptors and downstream effectors. The other class is characterized by scaffold molecules that form homoprotein or heteroprotein complexes with themselves or with other scaffold proteins *via* protein–protein interaction domains such as the L27 domain. The LIN-2–LIN-7–LIN-10 heterotrimeric complex, which was first identified in *Caenorhabditis elegans*, is a notable example of a multi-protein scaffold network. L27-domain-mediated LIN-2–LIN-7–LIN-10 complex formation plays a central role in targeting the receptor tyrosine kinase Let-23 signalosome to cell junctions during *C. elegans* vulval development (Kaech *et al.*, 1998).

The mammalian LIN-2 homologue Pals1 (protein associated with LIN-7 1, also known as MPP5) is a member of the MAGUK protein family and localizes to the tight junctions of epithelial cells. In addition to the characteristic MAGUK modules (PDZ, SH3 and GUK), Pals1 has two L27 domains (L27N and L27C) arranged in tandem (Fig. 1*a*). The L27N domain of Pals1 interacts with the L27 domain of the multi-PDZ-domain protein PATJ (Pals1-associated tight junction protein). The PDZ domains of human PATJ interact with CRB1, TJP3 and CLDN1, playing crucial roles in tight junction formation, protein targeting and cell polarity (Lemmers *et al.*, 2002; Roh, Liu *et al.*, 2002; Roh, Makarova *et al.*, 2002). Pals1 interacts with the L27 domain of MALS (mammalian LIN-7 protein) *via* its L27C



domain. MALS-1, MALS-2 and MALS-3 are single PDZ-containing proteins. The three MALS isoforms have different binding partners and they are tissue-dependently subcellularly localized (Jo *et al.*, 1999; Olsen *et al.*, 2005). The PDZ domain of LIN-7 binds to the epidermal growth-factor receptor in *C. elegans* (Kaech *et al.*, 1998; Li *et al.*, 2004). Recent work with mammalian cells indicated that the interaction between the L27 domains of MALS and MAGUK family members is essential for the stability of several polarity protein complexes and for mediating tight junction formation (Straight *et al.*, 2006). PATJ, Pals1 and MALS assemble in an L27-domain-mediated complex that functions as an organizing centre for large protein assemblies in the establishment and maintenance of epithelial tissue apical–basal polarity (Srinivasan *et al.*, 2008).

Structural studies of $L27_{PATJ}$–$L27N_{Pals1}$ and $L27_{SAP97}$–$L27N_{mLIN-2}$ heterodimers suggest a model in which two heterodimers form an assembly of symmetric dimers of dimers (Feng *et al.*, 2004; Li *et al.*, 2004). Furthermore, the structure of the human tripartite complex $L27_{Dlg1}$–$(L27N,L27C)_{MPP7}$–$L27_{MALS-3}$ showed an asymmetric and synergistic heterotrimer consisting of two pairs of heterodimeric L27 domains (Yang *et al.*, 2010). However, the L27 domains in the different proteins exhibit a high degree of diversity in terms of their location, copy number, primary sequence and binding preference (Li *et al.*, 2004). In the tripartite $L27_{PATJ}$–$(L27N,L27C)_{Pals1}$–$L27_{MALS}$ complex, two pairs of heterodimers assemble independently, in sharp contrast to the tripartite complex $L27_{Dlg1}$–$(L27N,L27C)_{MPP7}$–$L27_{MALS-3}$. Still unclear, however, is the mechanism of assembly of the $L27_{PATJ}$–$(L27N,L27C)_{Pals1}$–$L27_{MALS}$ tripartite complex. Here, we report the expression, purification and crystallization of different constructs of the $L27_{PATJ}$–$(L27N,L27C)_{Pals1}$–$L27_{MALS}$ tripartite complex for the purpose of obtaining high-resolution crystallographic data in order to clarify the molecular mechanism of tandem L27-mediated protein-complex assembly.

2. Materials and methods

2.1. Cloning and expression

2.1.1. Tripartite complex 1 (TC1). The DNA sequences encoding the L27 domain of PATJ (residues 1–68) from *Rattus norvegicus*, the tandem L27 domains of Pals1 (residues 119–232) from *Homo sapiens* and the L27 domain of MALS-2 (residues 3–66) from *Mus musculus* were amplified using the polymerase chain reaction (PCR). The three fragments were joined into a single open reading frame by two 3C protease-cleavable segments (LEVLFQGP). An additional 'GGG' cassette was inserted before the second 3C protease cleavage segment (Fig. 1*b*). The single fused DNA sequence was cloned into an in-house-modified version of the pET-32a vector (Novagen) in which the S-tag and the thrombin recognition site were replaced by a sequence encoding a TEV protease cleavage site (ENLYFQS).

The recombinant plasmid was transformed into *Escherichia coli* BL21 CodonPlus (DE3) and the *E. coli* cells were cultured at 310 K until the OD_{600} reached about 0.6, whereupon isopropyl β-D-1-thio-galactopyranoside was added to a final concentration of 0.3 m*M* to induce protein expression. After further incubation for 16 h at 289 K, the cells were harvested by centrifugation at 5000*g* for 15 min and resuspended in 30 ml buffer *A* (50 m*M* Tris–HCl pH 7.9, 500 m*M*

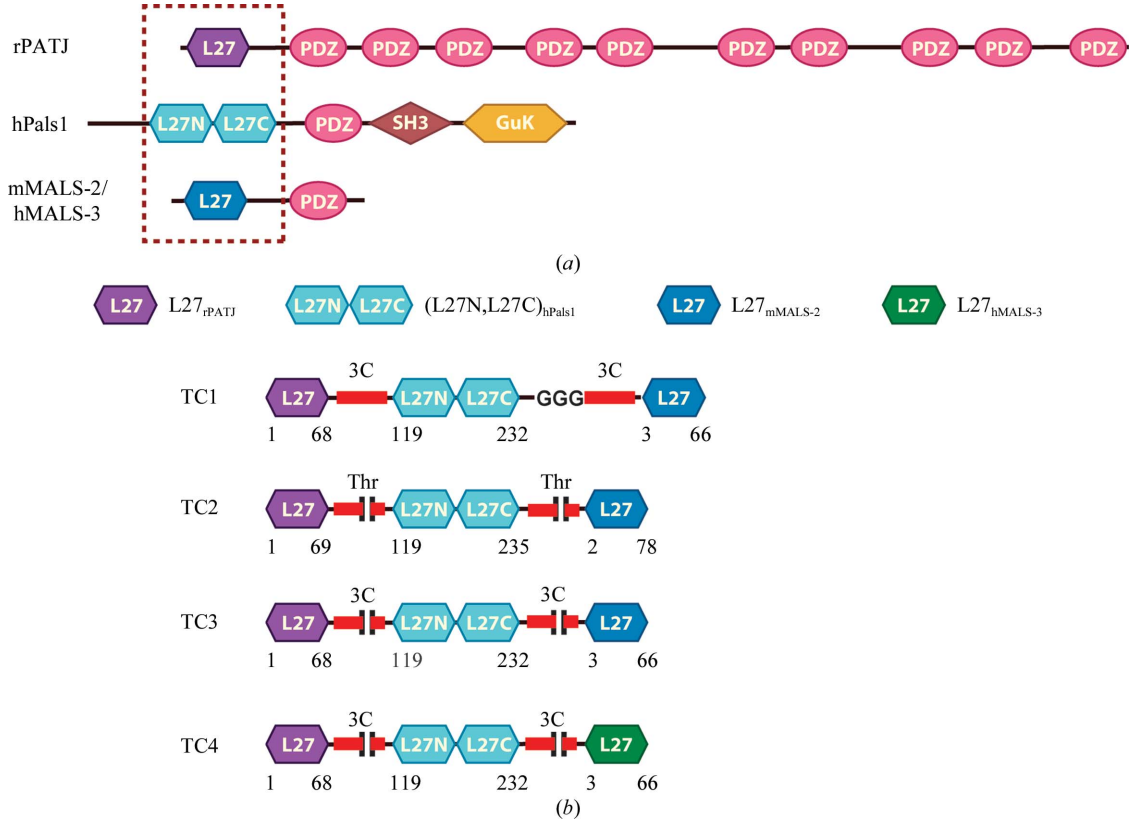


Figure 1
(*a*) Schematic representation of the domain organization of rat PATJ (rPATJ), human Pals1 (hPals1), mouse MALS-2 (mMALS-2) and human MALS-3 (hMALS-3). A brown dashed box shows the heterotrimer consisting of four L27 domains from the tripartite complex rPATJ–hPals1–mMALS-2 or rPATJ–hPals1–hMALS-3. The L27 domains of rPATJ, hPals1, mMALS-2 and huMALS-3 are coloured purple, cyan, blue and green, respectively. (*b*) Schematic representation of the constructs described in this paper. Thr and 3C represent thrombin and 3C protease cleavage sites, respectively. In TC1 the three proteins were linked into a single chain, whereas the three proteins were separated using thrombin digestion of TC2 or 3C digestion of TC3 and TC4.

NaCl, 5 m*M* imidazole) supplemented with 1 m*M* phenylmethylsulfonyl fluoride, 1 µ*M* leupeptin and 1 µ*M* antipain. The cells were lysed by sonication. After sonication, the debris was removed by centrifugation at 20 000*g* for 30 min. The crude cell supernatant was passed through a 0.22 µm filter and loaded onto a 5 ml nickel-chelating column (GE Healthcare, Piscataway, New Jersey, USA) which was equilibrated with buffer *A*. The column was washed with ten column volumes of buffer *A* to remove the unbound protein. The bound protein was then eluted with a 40-column-volume linear gradient performed by changing the concentration of imidazole from 5 m*M* to 1 *M* in 60 min at a flow rate of 3 ml min^{-1}. Fractions containing the fusion protein were identified by SDS–PAGE. Fractions with the highest purity (>95%) were pooled, concentrated to 15 ml, loaded onto a HiLoad Superdex 200 26/60 column (GE Healthcare, Piscataway, New Jersey, USA) and eluted with buffer *B* (50 m*M* Tris–HCl pH 7.5, 50 m*M* NaCl, 1 m*M* EDTA, 1 m*M* DTT) at a flow rate of 2.5 ml min^{-1}. After gel filtration, the protein was digested with TEV protease. The final target protein was loaded again onto a HiLoad Superdex 200 26/60 column to remove the His_6 tag and eluted with buffer *B*. The protein was concentrated to 90 mg ml^{-1} using a 10 kDa cutoff membrane and immediately subjected to crystallization trials.

2.1.2. Tripartite complex 2 (TC2). The DNA sequences encoding the L27 domain of PATJ (residues 1–69) from *R. norvegicus*, the tandem L27 domains of Pals1 (residues 119–235) from *H. sapiens* and the L27 domain of MALS-2 (residues 2–78) from *M. musculus* were amplified and connected by a human thrombin-cleavable segment (LVPRGS) following a short linker (SG). The single open reading frame was cloned into the in-house-modified pET-32a plasmid described above. The target protein was produced by following the same protocol as that for TC1, with the only exception being that the thrombin recognition linkers were cleaved by thrombin after the first gel filtration to produce three separate proteins (Fig. 1*b*).

2.1.3. Tripartite complex 3 (TC3). The cloning, expression and purification of TC3 was performed by following the same protocol as that for TC1, with the exception that the 'GGG' cassette was not inserted before the second 3C cleavage site and the 3C recognition linkers were cleaved by rhinovirus 3C protease after the first gel filtration (Fig. 1*b*).

2.1.4. Tripartite complex 4 (TC4). The cloning, expression and purification of TC4 was performed by following the same protocol as that for TC3, with the exception that the L27 domain of MALS-2 (residues 3–66) from *M. musculus* was replaced by the L27 domain (residues 3–66) of MALS-3 from *H. sapiens* (Fig. 1*b*).

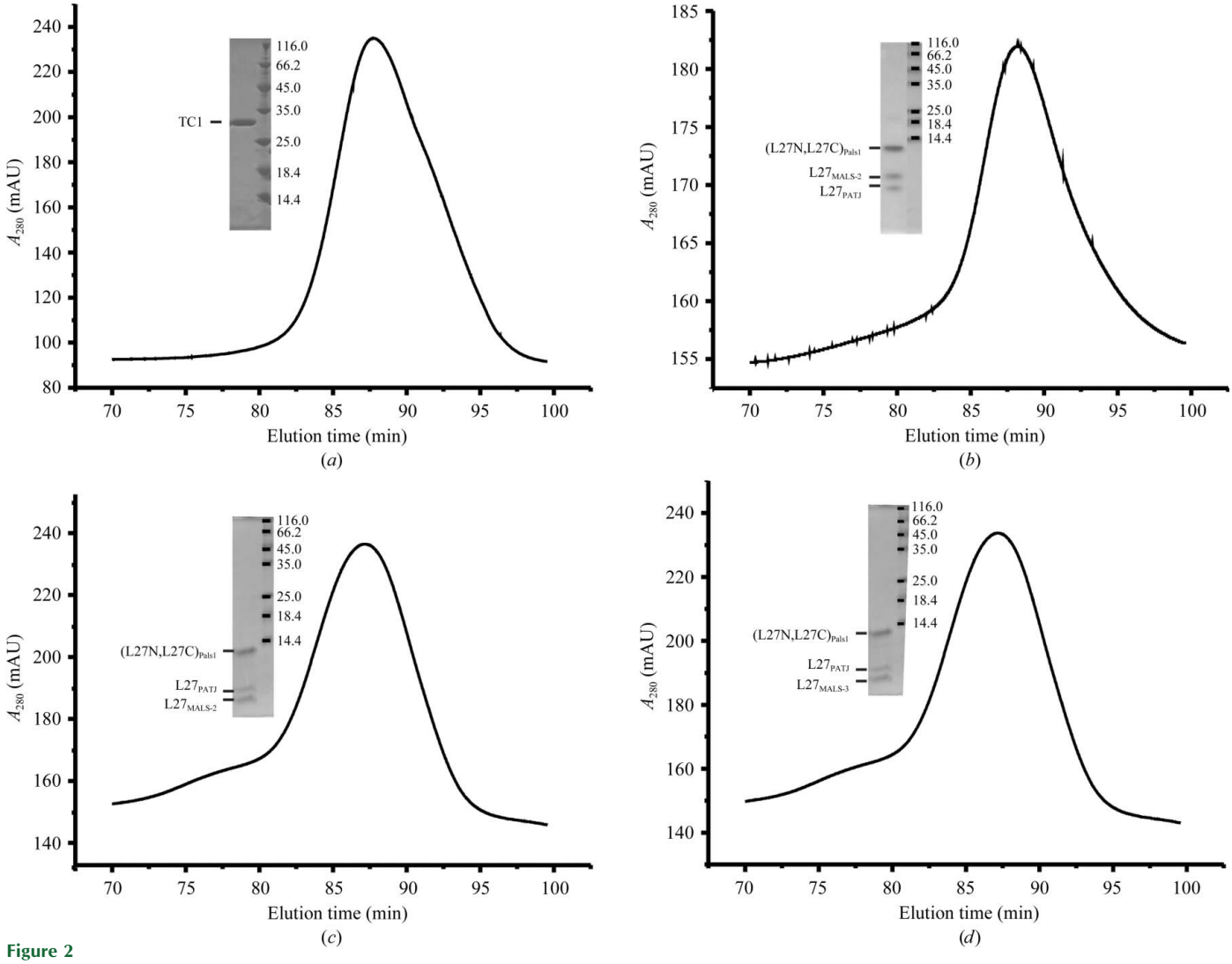


Figure 2
Elution profiles of (*a*) TC1, (*b*) TC2, (*c*) TC3 and (*d*) TC4 using a Superdex 200 26/60 Hi-Load gel-filtration column. SDS–PAGE analyses are shown as insets (the right lanes contain molecular-weight markers labelled in kDa).

2.2. Crystallization

Sitting-drop crystallization trials were performed at 293 K by mixing 1 µl protein solution (90 mg ml^{-1} in buffer *B*) with 1 µl precipitant solution. Up to 16 different screens from Hampton Research (California, USA) and Emerald BioSystems (Bainbridge Island, USA) kits were applied to each protein. Screening for crystallization conditions and optimizations were both performed in 48-well sitting-drop plates. Crystals of TC1 were recovered in 0.2–0.3 mm nylon loops (Hampton Research) and immediately flash-cooled in liquid nitrogen.

2.3. Data collection and processing

X-ray data were collected on beamline BL-17U1 at the Shanghai Synchrotron Radiation Facility. Diffraction experiments were conducted at 100 K and the images were recorded on a 225 mm MAR CCD camera (MAR Research, Norderstedt, Germany). Data were indexed, integrated and scaled using *HKL*-2000 (Otwinowski & Minor, 1997).

3. Results and discussion

3.1. Preparation and confirmation of protein complexes

All of the proteins were highly overproduced in *E. coli* BL21 CodonPlus (DE3) and were purified as described in §2. Gel-filtration experiments using the target proteins showed elution peaks corresponding to the expected molecular weight of approximately 30 kDa, confirming that the tripartite complexes had formed. Analysis of these peaks on reducing SDS–PAGE yielded a single band with a molecular weight of ∼30 kDa for TC1 (Fig. 2*a*). In contrast, TC2 (Fig. 2*b*), TC3 (Fig. 2*c*) and TC4 (Fig. 2*d*) exhibited three bands corresponding to the single components of the complexes. The L27 domains of homologous proteins from different species are usually conserved. The amino-acid sequence identity of $L27_{PATJ}$ from mouse, rat and human is very high (>90%) and the amino-acid sequences of mouse and human $L27_{Mals\text{-}2}$ are identical. High sequence homology suggests that L27 domains in the same protein from various species may have similar biological functions (Feng *et al.*, 2004).

3.2. Crystallization

Initial crystal hits were obtained for each protein construct. The crystallization conditions were optimized by varying the precipitant/salt concentration and the buffer pH and using the Additive Screen Kit, Detergent Screen Kit and Silver Bullet Kit (Hampton Research, California, USA).

Table 1
X-ray data-collection statistics.

Values in parentheses are for the highest resolution shell.

Space group	$P6_122$ or $P6_522$
Unit-cell parameters (Å)	$a = b = 145.2$, $c = 202.5$
Wavelength (Å)	0.9795
Resolution range (Å)	50–2.05 (2.12–2.05)
No. of unique reflections	79261
Multiplicity	20.2 (12.1)
R_{merge}† (%)	8.1 (58.6)
$\langle I/\sigma(I)\rangle$	36.3 (3.1)
Completeness (%)	100.0 (100.0)

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$.

Crystals of TC1 were obtained ∼30 d after crystallization setup in Index (Hampton Research, California, USA) condition B12 (2.8 *M* sodium acetate) at 293 K (Fig. 3*a*). Crystals suitable for X-ray analysis were flash-cooled directly without additional cryoprotectant using liquid nitrogen and diffracted to 2.05 Å resolution on beamline BL-17U1 at the Shanghai Synchrotron Radiation Facility.

TC2 appeared to form small crystals within a week after crystallization setup in SaltRx 2 (Hampton Research, California, USA) condition G11 (1.3 *M* ammonium tartrate dibasic, 0.1 *M* bis-tris propane pH 7.0; Fig. 3*b*). However, after extensive optimization experiments the crystals were still too small for diffraction.

Crystals of TC3 were obtained 1 d after crystallization setup in PEG/Ion 2 condition No. 27 (0.1 *M* sodium formate pH 7.0, 12% PEG 3350) at 277 K (Fig. 3*c*). Well shaped large crystals were transferred into several different cryoprotectant solutions; however, subsequent liquid-nitrogen flash-cooling yielded crystals with diffraction patterns that either showed no spots or that exhibited blurry low-resolution (8–10 Å) diffraction.

Crystals of TC4 were rounded and were obtained approximately one week after crystallization setup in Wizard I (Emerald BioSystems, Bainbridge Island, USA) condition No. 38 (1.0 *M* sodium/potassium tartrate, 0.2 *M* lithium sulfate, 0.1 *M* CHES pH 9.5) at 293 K. After optimization trials, the crystals could not be improved to give sharper edges or a regular shape (Fig. 3*d*).

Despite being homogenous and having high purity and solubility, proteins TC2, TC3 and TC4 failed to form crystals that were suitable for diffraction. This failure may be a consequence of the flexibility of each N-terminal and C-terminal L27 domain, resulting in multiple conformational states in the crystal packing. Previous studies have shown that the presence of the covalent linker between the dimeric or trimeric L27 domains has a minimal effect on the overall structure (Feng *et al.*, 2004; Yang *et al.*, 2010). The artificial linker between the

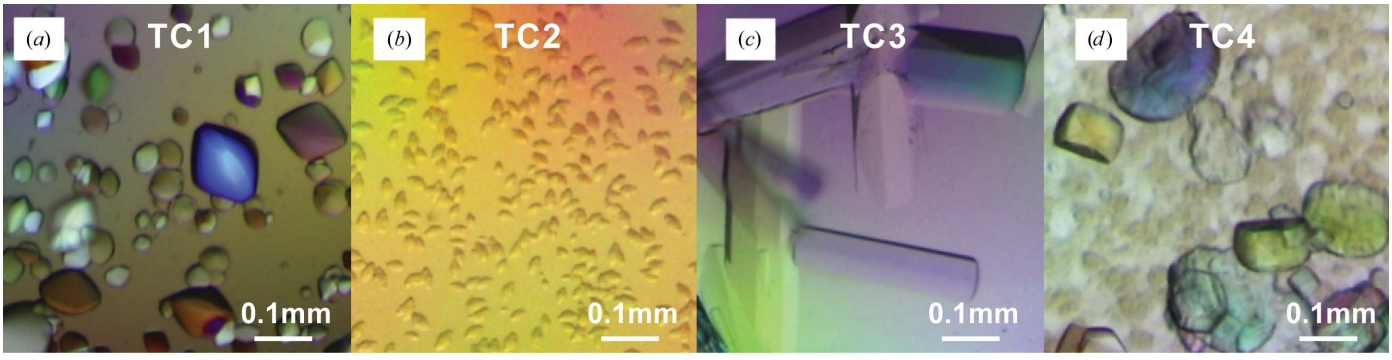


Figure 3
Sitting-drop crystals of (*a*) TC1, (*b*) TC2, (*c*) TC3 and (*d*) TC4.

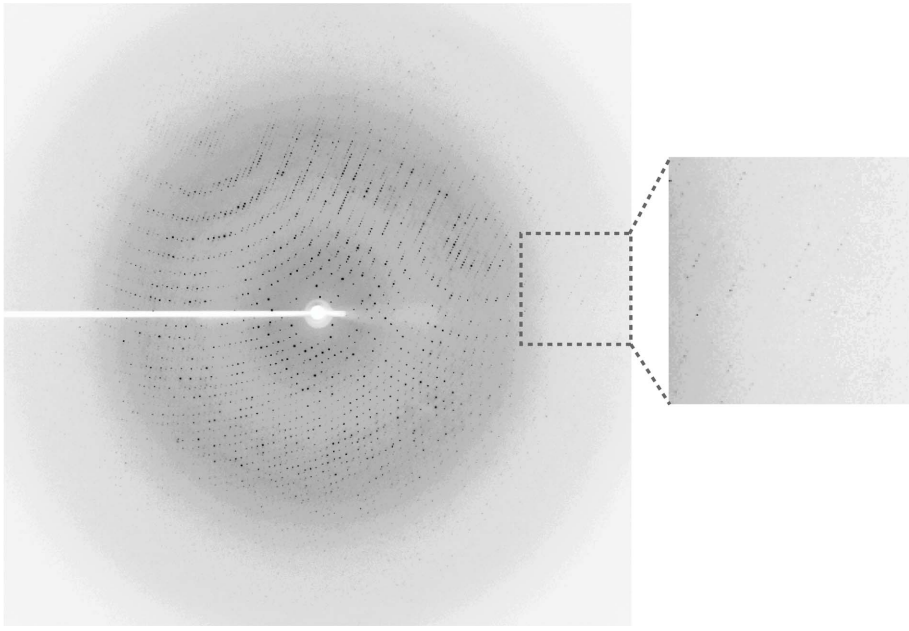


Figure 4
A representative diffraction image from the TC1 crystal. The inset shows weak spots at the edge corresponding to 2.05 Å resolution. A complete data set was collected to 2.05 Å resolution.

distinct L27 domains (such as in the TC1 protein) may have helped the tripartite complex to attain a stable conformation.

3.3. Data collection and processing

A full native diffraction data set was collected and processed to a resolution of 2.05 Å. One of the diffraction patterns is shown in Fig. 4. The crystal belonged to space group $P6_122$ or $P6_522$, with unit-cell parameters $a = b = 145.2$, $c = 202.5$ Å. Resolution-dependent Matthews coefficient probability analysis (Kantardjieff & Rupp, 2003) suggested the presence of four molecules per asymmetric unit, with 54.25% solvent content and a V_M value of 2.69 Å^3 Da^{-1}. The data are shown in Table 1.

Molecular replacement using the coordinates of the human tripartite complex $L27_{Dlg1}$–$(L27N,L27C)_{MPP7}$–$L27_{Mals-3}$ (PDB entry 3lra; Yang *et al.*, 2010) as a template was attempted with the *Phaser* molecular-replacement program (McCoy *et al.*, 2007) as implemented in the *CCP*4 suite (Winn *et al.*, 2011). Both the entire tripartite complex and the individual L27 domains were used as search models, but no valid solution was obtained. This result may be a consequence of conformational differences of the L27 domains in the two complexes. Selenium-substituted crystals of TC1 are currently being optimized in order to solve the phase problem.

We are grateful to the staff at beamline BL17U1 of the Shanghai Synchrotron Radiation Facility for excellent technical assistance during data collection. This work was funded by the 973 Program (grants 2009CB825504 and 2007CB914301) and the TBR program (grants 08QTPTJC28200 and 08SYSYTC00200).

References

Feng, W., Long, J.-F., Fan, J.-S., Suetake, T. & Zhang, M. (2004). *Nature Struct. Mol. Biol.* **11**, 475–480.
Iden, S. & Collard, J. G. (2008). *Nature Rev. Mol. Cell Biol.* **9**, 846–859.
Jo, K., Derin, R., Li, M. & Bredt, D. S. (1999). *J. Neurosci.* **19**, 4189–4199.
Kaech, S. M., Whitfield, C. W. & Kim, S. K. (1998). *Cell*, **94**, 761–771.
Kantardjieff, K. A. & Rupp, B. (2003). *Protein Sci.* **12**, 1865–1871.
Kapsenberg, M. L., Hilkens, C. M., van der Pouw Kraan, T. C., Wierenga, E. A. & Kalinski, P. (2000). *Am. J. Respir. Crit. Care Med.* **162**, S76–S80.
Lemmers, C., Médina, E., Delgrossi, M. H., Michel, D., Arsanto, J. P. & Le Bivic, A. (2002). *J. Biol. Chem.* **277**, 25408–25415.
Li, Y., Karnak, D., Demeler, B., Margolis, B. & Lavie, A. (2004). *EMBO J.* **23**, 2723–2733.
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.
Mitra, D. K., De Rosa, S. C., Luke, A., Balamurugan, A., Khaitan, B. K., Tung, J., Mehra, N. K., Terr, A. I., O'Garra, A., Herzenberg, L. A. & Roederer, M. (1999). *Int. Immunol.* **11**, 1801–1810.
Ohno, S. (2001). *Curr. Opin. Cell Biol.* **13**, 641–648.
Olsen, O., Wade, J. B., Morin, N., Bredt, D. S. & Welling, P. A. (2005). *Am. J. Physiol. Renal Physiol.* **288**, F345–F352.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pieczynski, J. & Margolis, B. (2011). *Am. J. Physiol. Renal Physiol.* **300**, F589–F601.
Roh, M. H., Liu, C.-J., Laurinec, S. & Margolis, B. (2002). *J. Biol. Chem.* **277**, 27501–27509.
Roh, M. H., Makarova, O., Liu, C.-J., Shin, K., Lee, S., Laurinec, S., Goyal, M., Wiggins, R. & Margolis, B. (2002). *J. Cell Biol.* **157**, 161–172.
Srinivasan, K., Roosa, J., Olsen, O., Lee, S.-H., Bredt, D. S. & McConnell, S. K. (2008). *Development*, **135**, 1781–1790.
Straight, S. W., Pieczynski, J. N., Whiteman, E. L., Liu, C.-J. & Margolis, B. (2006). *J. Biol. Chem.* **281**, 37738–37747.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.
Yang, X., Xie, X., Chen, L., Zhou, H., Wang, Z., Zhao, W., Tian, R., Zhang, R., Tian, C., Long, J. & Shen, Y. (2010). *FASEB J.* **24**, 4806–4815.