crystallization papers

Expression, crystallization and preliminary X-ray studies of the recombinant PTB domain of human dok-5 protein

Ning Shi,[a,b] Weihong Zhou,[a] Kaifu Tang,[b] Yu Gao,[a] Jian Jin,[b] Feng Gao,[a] Xiaozhong Peng,[b] Mark Bartlam,[a] Boqin Qiang,[b] Jiangang Yuan[b] and Zihe Rao[a,b]*

[a]MOE Laboratory of Protein Sciences and Laboratory of Structural Biology, Department of Biological Science and Biotechnology, Tsinghua University, Beijing 100084, People's Republic of China, and [b]National Laboratory of Medical Molecular Biology, Institute of Basic Medical Sciences, Chinese Academy of Medical Sciences, Peking Union Medical College, National Human Genome Center, Beijing 100005, People's Republic of China

Correspondence e-mail:
raozh@xtal.tsinghua.edu.cn

The human dok-5 PTB domain fusion protein has been over-expressed in *Escherichia coli* and crystallized in a form suitable for X-ray crystallographic study. Crystals were obtained by the vapour-diffusion method. The crystal has unit-cell parameters $a = b = 75.9$, $c = 108.0$ Å, $\alpha = \beta = 90$, $\gamma = 120°$ and belongs to space group $P3_221$. Diffraction data were collected to 2.8 Å resolution in-house. Furthermore, a selenomethionine (SeMet) derivative of dok-5 PTB domain fusion protein was overexpressed using the same expression system and was purified in a reductive environment. The derivative crystals were obtained under similar conditions. Subsequently, three different wavelength data sets were collected to 2.3 Å resolution from the derivative crystal at the Advanced Photon Source, Argonne National Laboratory.

1. Introduction

Dok-5 has been described as a novel member of the dok family of adapter proteins (Grimm *et al.*, 2001). It directly links with downstream effectors of the receptor tyrosine kinase c-Ret in neuronal differentiation (Grimm *et al.*, 2001). The dok family proteins include dok-1, dok-2, dok-3, dok-4 and dok-5 (Grimm *et al.*, 2001). Dok-1, dok-2 and dok-3 are mainly expressed in haematopoietic tissues and in negative regulation of signalling pathways activated by tyrosine kinases (Carpino *et al.*, 1997; Lemay *et al.*, 2000; Nelms *et al.*, 1998). These dok proteins inhibit mitogen-activated protein (MAP) kinase signalling, cell proliferation and cellular transformation (Cong *et al.*, 1999; Suzu *et al.*, 2000; Tamir *et al.*, 2000). Dok-4 and dok-5 are expressed in the nervous system, do not bind rasGAP and play a positive role in activation of the MAP kinase pathway (Grimm *et al.*, 2001). All members of the dok family contain an N-terminal pleckstrin homology (PH) domain, a central phospho-tyrosine-binding (PTB) domain and a C-terminal tail containing multiple tyrosine phosphorylation sites. The PH domain is known to bind to acidic phospholids and to localize protein to the plasma membrane, while the PTB domain is known to mediate protein–protein interactions *via* binding to phospho-tyrosine-containing motifs (Jacobs *et al.*, 2001). It is primed to form a multimolecular signalling complex in signal transduction (Kavanaugh *et al.*, 1995; Wang *et al.*, 1998). So far, many proteins have been identified as containing PTB domains; they fall into two major groups. The first group contains PTB domains that have primary sequence similarity to the Shc PTB domain. The second group contains proteins such as insulin-receptor substrate (IRS) proteins, which contain PTB domains with limited sequence similarity to the Shc PTB domains but similar binding characteristics (Wolf *et al.*, 1995). The PTB domain is an alternative to the src homology 2 (SH2) domain for phosphotyrosine recognition, but their recognizing character is different. PTB domains recognize pTyr in the context of the Asn-Pro-Xaa-pTyr (NP*X*pY) motif, whereas SH2 domains generally recognize the +1 to +4 positions immediately C-terminal to the pTyr (Zhou *et al.*, 1995). Recent studies show that PTB domain-like protein modules can also bind to proteins independent of tyrosine phosphorylation or even the canonical NP*X*Y motif (Dhalluin *et al.*, 2000; Forman-Kay & Pawson, 1999; Zhang *et al.*, 1997).

The dok-5 PTB domain belongs to the second group. It is 25% identical in sequence to IRS PTB domain and can also recognize sequences that contain the NP*X*pY motif. Whether the binding specificity of the dok-5 PTB domain is same as the other proteins is not known. Although the function of dok-5 is not clear, it has been proposed that the dok-5 PTB domain has an important role in signal transduction during nerve development (Grimm *et al.*, 2001). Here, we report our results in the cloning, expression, purification, crystallization and preliminary X-ray crystal-

lographic analysis of natural and seleno-methionyl recombinant dok-5 PTB domain protein.

2. Materials and methods

2.1. Plasmid construction

The *dok*-5 gene was found in a large-scale sequencing of the human brain cDNA library. A PCR product containing the coding sequence of *dok*-5 PTB domain was generated from the recombinant pUC plasmid carrying the *dok*-5 gene. Two PCR primers, 5′-CG GAA TTC CTG GCC ACT GGG GTT GAG AGA GAA C-3′ and 5′-CCG CTC GAG CTC GGC TAT GGC CAA GGC AGC AG-3′, were designed. The PCR product was restricted with *Eco*RI and *Xho*I, purified and ligated into *Eco*RI- and *Xho*I-restricted sites of the pGEX-6P-1 plasmid vector (Amersham Pharmacia Biotech.) with T4 DNA ligase. A further transformation into *E. coli* DH5α competent

cells was performed and the positive clones with an insert of the right size were identi-fied by double digestion with *Eco*RI and *Xho*I. The sequence of the insert was veri-fied by sequencing. The clones harbouring the expected recombinant plasmid were used for protein expression.

2.2. Overexpression and purification

The 106-amino-acid dok-5 PTB domain, whose N-terminus was fused with the glutathione *S*-transferase in frame, were overexpressed in *Escherichia coli* DH5α (Novagen) after induction with 0.2 mM isopropyl-β-D-thiogalactoside (IPTG). The cells were cultured for 18 h at 303 K and then harvested by centrifugation at 7700g for 10 min at 277 K. The precipitated cells were resuspended in 0.14 mM cold PBS buffer (140 mM NaCl, 2.7 mM KCl, 10 mM Na$_2$HPO$_4$, 1.8 mM KH$_2$PO$_4$, 5 mM EDTA, 1 mM PMSF) pH 7.3 and homogenized by sonication. The lysate was centrifuged at 20 000g for 20 min at 277 K to remove the cell debris. The supernatant was applied to a column (containing 2.5 ml glutathione–agarose 4B, Pharmacia) pre-equilibrated with 0.14 mM PBS. The contaminant proteins were washed out with 20 ml cold PBS. After three washes with PBS, 2.5 ml of PBS and 40 units of PreScission Protease were added to the column. The mixture was incubated with gentle agitation for 16 h at 277 K. The target protein was washed down with buffer *A* (20 mM Tris–HCl pH 8.2). This sample was subsequently applied to a 1 ml Resource Q column (Pharmacia) and the column was then developed with an NaCl linear gradient from 0 to 400 mM in

buffer *A*. The target protein peak that was eluted at 60 mM NaCl was collected and concentrated. Further purification was achieved by gel-filtration chromatography on Superdex G75 (Pharmacia) in buffer *A* containing 50 mM NaCl. The purified protein from Superdex G75 was concen-trated to 5–10 mg ml^{-1} using a 5K ultra-filtration membrane and was immediately used for crystallization. Homogeneity was demonstrated by 15% SDS–PAGE.

The selenomethionine derivative of the dok-5 PTB domain was expressed using *E. coli* DH5α cultured in minimal media M9 containing 30 mg l^{-1} L-SeMet. A further six amino acids (lysine, threonine, phenyl-alanine, leucine, isoleucine and valine) were added to the culture for inhibited Met biosynthesis of the DH5α expression strain. Purification of the selenomethionine dok-5 PTB domain derivative was performed as for the native dok-5 PTB domain except for the addition of a reducing environment provided by 5 mM DTT and 0.5 mM EDTA. The incorporation of selenium was confirmed by mass-spectrometric analysis.

2.3. Crystallization

Initial crystallization trials were carried out by the hanging-drop vapour-diffusion method using Crystal Screen reagent kits I and II (Hampton Research; Jancarik & Kim, 1991). Each drop was formed by mixing equal volumes (1–3 µl) of protein and reservoir solutions and was allowed to equilibrate *via* vapour diffusion over 200 µl of reservoir solution at 291 K. The protein concentration was 6 mg ml^{-1} prior to mixing with the reservoir solution. Conditions
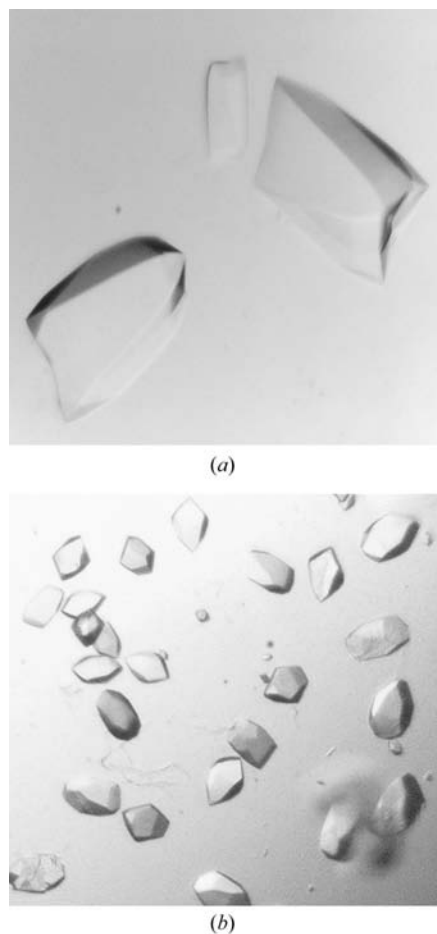

(a)

(b)

Figure 1
Crystals of dok-5 PTB domain. (*a*) Native crystal (dimensions of a single crystal are approximately 0.7 × 0.4 × 0.2 mm). (*b*) Selenomethionine-deriva-tive crystals (maximum dimensions of about 0.4 × 0.2 × 0.1 mm).
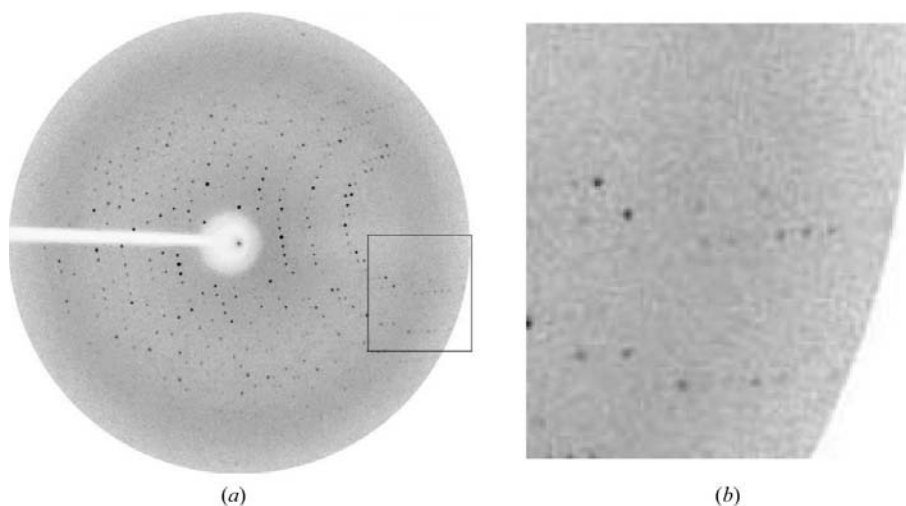

(a)

(b)

Figure 2
(*a*) A typical X-ray diffraction pattern from a crystal of dok-5 PTB domain. The diffraction image was collected on a MAR Research image-plate detector with a crystal-to-film distance of 140 mm. The oscillation range is 1°. The detector edge corresponds to 2.7 Å. (*b*) An enlarged image of the area indicated in (*a*).

crystallization papers

Space group			$P3_221$
Unit-cell parameters			
$a = b$ (Å)			75.9
c (Å)			108.0
γ (°)			120

	Peak	Inflection	Remote
Wavelength	0.97945	0.97959	0.95372
Resolution	50–2.3	50–2.3	50–2.3
Completeness (%)	99.3 (98.9)	99.3 (99.0)	99.3 (99.4)
Mean redundancy	7.8 (4.9)	7.8 (4.9)	8.0 (5.9)
$R_{merge}(I)$	0.075 (0.399)	0.067 (0.420)	0.069 (0.573)
Mean $I/\sigma(I)$	15.4 (4.6)	15.2 (4.4)	14.5 (3.9)
Total observations	127639	127777	132314
Unique reflections	16437 (1602)	16365 (1571)	16349 (1573)

† $R_{merge}(I)$ is defined as $\sum |I - \langle I \rangle| / \sum I$.

yielding small crystals were further optimized by refinement of precipitants, buffer pH, protein concentration, additives and temperature.

The purified selenomethionine derivative was concentrated to 4 mg ml^{-1}. Crystallization trials were set up based on the optimum conditions used for native protein.

2.4. Data collection and processing

Data were collected in-house at 115 K using a MAR Research image plate and a 4.8 kW Rigaku rotating-anode generator producing Cu $K\alpha$ radiation of wavelength 1.5418 Å. Multiple-wavelength anomalous dispersion (MAD) data were collected on beamline 19ID under cryoconditions at the Advanced Photon Source (APS), Argonne National Laboratory. In each case, crystals were frozen in the crystallization buffer with 20% glycerol as cryoprotectant before data were collected.

Data processing was performed using the program *DENZO* and data sets were scaled and merged using *SCALEPACK* (Otwinowski & Minor, 1997).

3. Results and discussion

The final purified protein was confirmed to be homogenous by SDS–PAGE analysis and to be suitable for use in crystallization trials. Small hexagonal rod-shaped crystals appeared after 1 d from several different conditions of Crystal Screen kits I and II (Hampton Research) that contained lithium sulfate or ammonium sulfate (kit I conditions Nos. 16, 32, 39 and kit II conditions Nos. 15, 25, 32). The conditions were further optimized by variation of precipitants, buffer pH and protein concentration. Large tooth-shaped crystals were obtained from a reservoir consisting of 1.4 *M* (NH$_4$)$_2$SO$_4$, 0.1 *M* HEPES pH 7.5. However, these crystals only diffracted to 4.0 Å in-house.

In order to increase the resolution, we slowed the growth of crystals by regulating the protein concentration, the volume of the reservoir and the temperature. The best crystals (Fig. 1*a*) grew to dimensions of 0.7 × 0.4 × 0.2 mm at 277 K in one week and diffracted to 2.8 Å in-house. Selenomethionine-derivative crystals (Fig. 1*b*) were obtained using similar conditions to those for native protein, except that a lower pH (0.1 *M* MES pH 6.0) and protein concentration (4 mg ml^{-1}) were used.

Diffraction data to 2.8 Å resolution were collected in-house from a native crystal (Fig. 2). The crystal belongs to space group $P3_221$, with unit-cell parameters $a = b = 76.2$, $c = 107.7$ Å, $\alpha = \beta = 90$, $\gamma = 120°$. Assuming three monomers (13.5 kDa) per asymmetric unit, the Matthews coefficient (volume-to-mass ratio; V_M) is 2.2 Å^3 Da^{-1}, giving a solvent content of 44%.

Our initial efforts to determine the structure by the molecular-replacement method using other PTB-domain structures were unsuccessful. Since there are two Met residues in the dok-5 PTB-domain sequence, the MAD method was considered. Statistics of data collection from selenomethionine-derivative crystals are shown in Table 1.

MAD data were collected from a single selenomethionine-derivative crystal at peak (λ_1; 0.97945 Å), inflection (λ_2; 0.97959 Å) and remote (λ_3; 0.95372 Å) wavelengths to 2.3 Å resolution. Structure determination of the dok-5 PTB domain is currently under way.

References

Carpino, N., Wisniewski, D., Strife, A., Marshak, D., Kobayashi, R., Stillman, B. & Clarkson, B. (1997). *Cell*, **88**, 197–204.
Cong, F., Yuan, B. & Goff, S. P. (1999). *Mol. Cell Biol.* **19**, 8314–8325.
Dhalluin, C., Yan, K., Plotnikova, O., Lee, K. W., Zeng, L., Kuti, M., Mujtaba, S., Goldfarb, M. P. & Zhou, M. M. (2000). *Mol. Cell*, **6**, 921–929.
Forman-Kay, J. D. & Pawson, T. (1999). *Curr. Opin. Struct. Biol.* **9**, 690–695.
Grimm, J., Sachs, M., Britsch, S., Di Cesare, S., Schwarz-Romond, T., Alitalo, K. & Birchmeier, W. (2001). *J. Cell Biol.* **154**, 345–354.
Jacobs, A. R., LeRoith, D. & Taylor, S. I. (2001). *J. Biol. Chem.* **276**, 40795–40802.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Kavanaugh, W. M., Turck, C. W. & Williams, L. T. (1995). *Science*, **268**, 1177–1179.
Lemay, S., Davidson, D., Latour, S. & Veillette, A. (2000). *Mol. Cell Biol.* **20**, 2743–2754.
Nelms, K., Snow, A. L., Hu-Li, J. & Paul, W. E. (1998). *Immunity*, **9**, 13–24.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Suzu, S., Tanaka-Douzono, M., Nomaguchi, K., Yamada, M., Hayasawa, H., Kimura, F. & Motoyoshi, K. (2000). *EMBO J.* **19**, 5114–5122.
Tamir, I., Stolpa, J. C., Helgason, C. D., Nakamura, K., Bruhns, P., Daeron, M. & Cambier, J. C. (2000). *Immunity*, **12**, 347–358.
Wang, H. Y., Zamorano, J. & Keegan, A. D. (1998). *J. Biol. Chem.* **273**, 9898–9905.
Wolf, G., Trub, T., Ottinger, E., Groninga, L., Lynch, A., White, M. F., Miyazaki, M., Lee, J. & Shoelson, S. E. (1995). *J. Biol. Chem.* **270**, 27407–27410.
Zhang, Z., Lee, C. H., Mandiyan, V., Borg, J. P., Margolis, B., Schlessinger, J. & Kuriyan, J. (1997). *EMBO J.* **16**, 6141–6150.
Zhou, M. M., Harlan, J. E., Wade, W. S., Crosby, S., Ravichandran, K. S., Burakoff, S. J. & Fesik, S. W. (1995). *J. Biol. Chem.* **270**, 31119–31123.