Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary crystallographic analysis of a metastasis-inducing protein, human S100A4

Hongmei Zhang,[a] Zhilong Wang,[a] Yi Ding,[a] Guozhen Wang,[b] Xiaodong Wang,[a] Mark Bartlam,[a] Hong Tang,[a] Yiwei Liu,[a] Fan Jiang,[a] Roger Barraclough,[b] Philip S. Rudland[b] and Zihe Rao[a]*

[a]Laboratory of Structural Biology and MOE Laboratory of Protein Science, School of Life Science and Engineering, Tsinghua University, Beijing 100084, People's Republic of China, and [b]The Cancer and Polio Research Fund Laboratories, School of Biological Sciences, University of Liverpool, Liverpool L69 7ZB, England

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

Received 5 June 2001
Accepted 4 October 2001

A metastasis-inducing protein, human S100A4, cDNA was over-expressed in *Escherichia coli*. The recombinant protein was purified by three chromatographic procedures. Crystals were obtained using polyethylene glycol as a precipitant. The native S100A4 crystals did not diffract sufficiently well for data collection, but the SeMet S100A4 crystal, grown using similar conditions as for the native, was found to be stable during exposure to X-rays and diffracted to 4 Å resolution in-house. The crystal belongs to space group *P*6 or *P*3. The unit-cell parameters are $a = b = 47.1$, $c = 175.6$ Å, $\alpha = \beta = 90$, $\gamma = 120^\circ$. Complete data will be collected using synchrotron radiation.

1. Introduction

S100A4 (also known previously as pEL98/mts1/p9Ka/18A2/42A/calvasculin/CAPL) (Barraclough *et al.*, 1982; Goto *et al.*, 1988; Ebralidze *et al.*, 1989; Jackson-Grusby *et al.*, 1987; Masiakowski & Shooter, 1988; Watanabe *et al.*, 1992; Engelkamp *et al.*, 1992) is a member of the S100-related calcium-binding protein family (Kligman & Hilt, 1988). This family consists of 19 members that are differentially expressed in a large number of cell types (Donato, 1999). Most of the S100 proteins are of unknown function within cells (Zimmer *et al.*, 1995).

S100A4 is putatively associated with the cytoskeletal proteins (Gibbs *et al.*, 1994) and may influence cell motility, although its exact physiological role is not known. It is frequently overexpressed in metastatic breast tumours (Rudland *et al.*, 2000) as well as in normal cells with enhanced invasiveness, such as macrophages, neutrophils and T-lymphocytes. It is possible that S100A4 modulates cell motility by interacting with components of the cytoskeleton (Davies *et al.*, 1993). Filamentous actin, non-muscle myosin and non-muscle tropomyosin have been proposed as target molecules for this protein (Kriajevska *et al.*, 1994, 1998; Takenaga *et al.*, 1994, 1997; Ford & Zain, 1995; Barraclough & Rudland, 1994; Mandinova *et al.*, 1998).

The S100 proteins are a group of low molecular-weight acidic proteins that contain two Ca^{2+}-binding EF-hands and can form homodimers and heterodimers (Bhardwaj *et al.*, 1992; Kilby *et al.*, 1996). S100A4 forms a calcium- and dithiothreitol-independent homodimer (Tarabykina *et al.*, 2001). S100A4 can also interact with S100A1 and form a heterodimer *in vitro* (Wang *et al.*, 2000; Tarabykina *et al.*, 2000). Although all structures of the S100 family share common structural motifs such as dimerization and EF-hand, they have individual features: S100B undergoes dramatic conformational changes in the presence of calcium, S100A10 forms covalent tetramers, S100A7 has a zinc-binding site; structure comparision of S100A12 with the other S100 proteins proves once more that differences lie in the region of the flexible loop L11. While S100A10 and S100A11 form α-helix at the peptide-bound region of the loop L11, there is no helix in the same region of S100A12 (Matsumura *et al.*, 1998; Rety *et al.*, 1999, 2000; Brodersen *et al.*, 1999; Moroz *et al.*, 2001). To understand the structural basis of the function of S100A4, we have crystallized human S100A4 and are analyzing its three-dimensional structure.

2. Materials and methods

2.1. Bacterial expression and purification of human S100A4

A cDNA fragment encoding human S100A4 was amplified by RT PCR (Lloyd *et al.*, 1998) and cloned into the bacterial expression vector pET11a (Novagen). The recombinant plasmid was transformed into *E. coli* BL21 (DE3) (Novagen) and human S100A4 protein was expressed without any tag.

For isolation of S100A4, 1 l of LB medium was mixed with an overnight culture of transformed *E. coli* strain BL21 (DE3) in the ratio 100:1 and shaken at 310 K. Isopropyl-β-D-thiogalactopyranoside (IPTG) was added to 1 m*M* final concentration and incubation was continued for 10–12 h until an OD_{600} of 0.6–0.8 was reached. The bacteria were harvested by centrifugation, resuspended in 25 m*M* Tris–

HCl pH 7.6, 50 m*M* NaCl and lysed by sonication. The soluble cell lysate obtained by centrifugation at 15000 rev min^{-1} for 30 min was loaded onto a DEAE-Sepharose fast-flow anion-exchange chromatography column (APB, Milton Keynes, England) equilibrated with 25 m*M* Tris–HCl pH 7.6, 150 m*M* NaCl. After washing away the unbound protein with two bed volumes, a linear gradient of 0.15–0.6 *M* NaCl in the same buffer was applied. S100A4 eluted at approximately 0.35 *M* NaCl. $CaCl_2$ was added to 2 m*M* and the collected fractions were centrifuged for 30 min at 15 000 rev min^{-1}. The supernatant was absorbed onto a phenyl-Sepharose column (APB) equilibrated with 25 m*M* Tris–HCl, 0.1 m*M* $CaCl_2$ pH 7.6. Bound proteins were eluted with the same buffer containing 5 m*M* EGTA [ethylene glycol bis(*β*-aminoethylether)-*N,N,N′,N′*-tetraacetic acid]. The eluted sample was concentrated using Filtron 5K to 250 µl and loaded onto a Superdex G75 size-exclusion chromatography column (APB). The target peak corresponding to a molecular weight of about 22 kDa (a dimer) was collected and its purity was suitable for crystallization.

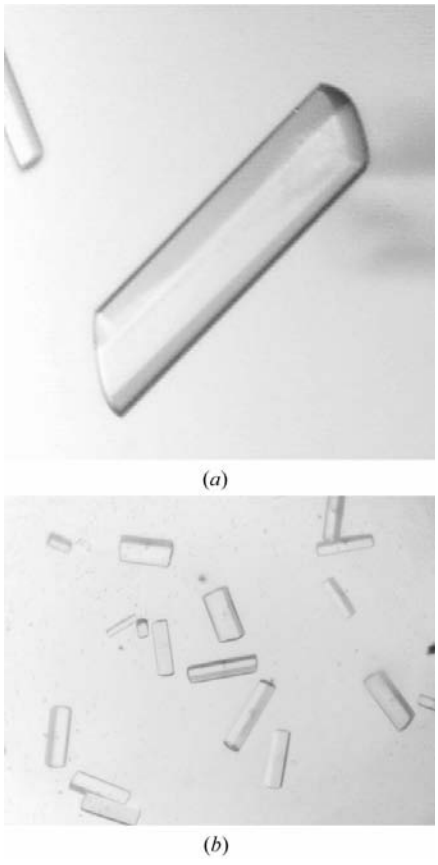

Figure 1
(*a*) Typical crystal of native human S100A4 (0.3 × 0.05 × 0.05 mm). (*b*) Typical crystals of SeMet human S100A4 (0.25 × 0.05 × 0.05 mm).

To express the SeMet S100A4 protein, the recombinant expression vector pET11a was transformed into *E. coli* B834. After incubation overnight in LB medium containing 100 µg ml^{-1} ampicillin, the bacteria were diluted with adaptive medium (20% LB medium + 80% M9 medium) and grown at 310 K to an OD_{595} of 0.6–0.8. The bacteria cells were harvested and resuspended in M9 medium, transferred to a large volume of adaptive medium and growth continued. L-SeMet at 60 mg l^{-1}, lysine, threonine and phenylalanine at 100 mg l^{-1}, leucine, isoleucine and valine at 50 mg l^{-1} and 2 m*M* IPTG were added and incubation continued for 10–12 h until the OD_{595} reached 0.6–0.8. The bacterial cells were harvested and the recombinant L-SeMet-labelled S100A4 protein was purified and crystallized as for the native protein.

2.2. Crystallization

The purified protein was concentrated using a 5K ultrafiltration membrane (Filtron) to 25 mg ml^{-1} in a solution containing 50 m*M* NaCl and 2 m*M* $CaCl_2$. Initial crystallization conditions were established by the hanging-drop vapour-diffusion method. A PEG screening kit (prepared by our laboratory) was used for preliminary screening. The best crystallization conditions were 18–28% PEG 4000, 0.1 *M* Tris–HCl pH 8.0 and 14–22% PEG 8000, 0.1 *M* HEPES pH 7.5. 1.5 µl of protein solution were mixed with 1.5 µl of reservoir solution and equilibrated against 0.6 ml of the reservoir solution.

2.3. X-ray data collection and analysis

X-ray diffraction data for the native and SeMet-derivative crystals were collected at 291 K using a MAR Research image-plate detector (180 mm) and Cu *Kα* radiation from a Rigaku rotating-anode generator operating at 48 kV and 98 mA. Oscillation images were processed with *DENZO* (Otwinowski & Minor, 1997).

3. Results and discussion

After purification by DEAE-Sepharose fast flow, phenyl-Sepharose and Superdex G75 size-exclusion chromatography, one 11 kDa band as determined by 16.5% SDS–PAGE dialysis was obtained corresponding to a molecular weight of human S100A4 monomer.

In order to crystallize S100A4, we tried screening using crystallization kits from Hampton Research using different protein concentrations and obtained some crystals from conditions containing ammonium salts. However, these crystals diffracted X-rays as salt crystals. We discontinued this screening and continued to screen using our home PEG kit. A full range of PEG (400–8000) and pH (6.0–9.0) was used for different crystallization trials. Some crystals were produced in buffers containing 18–28% PEG 4000, 0.1 *M* Tris pH 8.0 and 14–22% PEG 8000, 0.1 *M* HEPES pH 7.5. Based on the preliminary results, further refinement was carried out. 2% isopropanol was added to the reservoir solution as an additive, with the best results obtained with 18% PEG 8000, 0.1 *M* HEPES pH 7.5. Crystals grew in 3 d at 291 K and reached their final size in about a week (Fig. 1). The native crystals diffracted very poorly and a feasibility study using the synchrotron-radiation source at SRS (Daresbury, England) was carried out, where the low resolution of 7–9 Å was confirmed. The space group could not be determined because of the low resolution. The SeMet crystal diffracted well to a resolution of 4 Å in-house. The crystal belongs to space group *P*6 or *P*3 with unit-cell parameters $a = b = 47.1$, $c = 175.6$ Å (insufficient data were collected and the space group cannot be determined correctly). The protein consists of 101 amino acids. According to the result of Superdex G75 size-exclusion chromatography, a dimer should be crystallized. Assuming the space group to be *P*3, the asymmetric unit contains four or six 10 414 Da monomers and the specific volume V_M is 2.81 or 1.87 Å^3 Da^{-1}; assuming the space group to be *P*6, the asymmetric unit contains two 10 414 Da monomers and the specific volume V_M is 2.81 Å^3 Da^{-1}. MAD data will be collected using synchrotron radiation and the final structure should be determined shortly.

This research was supported by the following grants: NSFC Nos. 39870174 and 39970155, Project '863' No. 103130306, and Project '973' Nos. G1999075602, G1999011902 and 1998051105

References

Barraclough, R., Dawson, K. J. & Rudland, P. S. (1982). *Eur. J. Biochem.* **129**, 335–341.
Barraclough, R. & Rudland, P. S. (1994). *Eur. J. Cancer*, **30**A, 1570–1576.
Bhardwaj, R. S., Zotz, C., Zwadloklarwasser, G., Roth, J., Goebeler, M., Mahnke, K., Falk, M., Meinardus-Hager, G. & Sorg, C. (1992). *Eur. J. Immunol.* **22**, 1891–1897.
Brodersen, D. E., Nyborg, J. & Kjeldgaard, M. (1999). *Biochemistry*, **38**, 1695–704.
Davies, B. R., Davies, M. P., Gibbs, F. E., Barraclough, R. & Rudland, P. S. (1993). *Ocogene*, **8**, 999–1008.

Donato, R. (1999). *Biochim. Biophys. Acta*, **1450**, 191–231.
Ebralidze, A., Tulchinsky, E., Grigorian, M., Afanayeva, A., Senin, V., Revazova, E. & Lukanidin, E. (1989). *Genes Dev.* **3**, 1086–1093.
Engelkamp, D., Schafer, B. W., Erne, P. & Heizmann, C. W. (1992). *Biochemistry*, **31**, 10258–10264.
Ford, H. L. & Zain, S. B. (1995). *Oncogene*, **10**, 1597–1605.
Gibbs, F. E. M., Wilkinson, M. C., Rudland, P. S. & Barraclough, A. (1994). *J. Biol. Chem.* **269**, 18992–18999.
Goto, K., Endo, H. & Fujiyoshi, T. (1988). *J. Biochem. (Tokyo)*, **103**, 48–53.
Jackson-Grusby, L. L., Swiergiel, J. & Linzer, D. I. H. (1987). *Nucleic Acids Res.* **15**, 6677–6690.
Kilby, P. M., Van Eldik, L. J. & Roberts, G. C. (1996). *Structure*, **4**, 1041–1052.
Kligman, D. & Hilt, D. C. (1988). *Trends Biochem. Sci.* **13**, 437–443.
Kriajevska, M. V., Cardenas, M. N., Grigorian, M. S., Ambartsumian, N. S., Geogiev, G. P. & Lukanidin, E. M. (1994). *J. Biol. Chem.* **269**, 19679–19682.
Kriajevska, M., Tarabykina, S., Bronstein, I., Maitland, N., Lomonosov, M., Hansen, K., Georgiev, G. & Lukanadin, E. (1998). *J. Biol. Chem.* **273**, 9852–9856.
Lloyd, B. H., Platt-Higgins, A., Rudland, P. S. & Barraclough, R. (1998). *Oncogene*, **17**, 465–573.
Mandinova, A., Atar, D., Schaffer, B. W., Spiess, M., Aebi, U. & Heizmann, C. W. (1998). *J. Cell Sci.* **111**, 2043–2054.
Masiakowski, P. & Shooter, E. M. (1988). *Proc. Natl Acad. Sci. USA*, **85**, 1277–1281.
Matsumura, H., Shiba, T., Harada, S. & & Kai, Y. (1998). *Structure*, **6**, 233–241.
Moroz, O. V., Antson, A. A., Murshudov, G. N., Maitland, N. J., Dodson, G. G., Wilson, K. S., Skibshoj, I., Lukanidin, E. M. & Bronstein, I. B. (2001). *Acta Cryst.* D**57**, 20–29.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Rety, S., Osterloh, D., Arie, J. P., Tabaries, S., Seeman, J., Russo-Marie, F., Gerke, V. & Lewit-Bentley, A. (2000). *Structure Fold. Des.* **8**, 175–184.
Rety, S., Sopkova, J., Renouard, M., Osterloh, D., Gerke, V., Tabaries, S., Russo-Marie, F. & Lewit-Bentley, A. (1999). *Nature Struct. Biol.* **6**, 89–95.
Rudland, P. S., Platt-Higgins, A., Renshaw, C., West, C. R., Winstanley, J. H., Robertson, L. & Barrraclough, R. (2000). *Cancer Res.* **60**, 1595–1603.
Takenaga, K., Nakamura, Y. & Sakiyama, S. (1997). *Oncogene*, **14**, 331–337.
Takenaga, K., Nakamura, Y., Sakiyama, S., Hasegawa, Y., Sato, K. & Endo, H. (1994). *J. Cell Biol.* **124**, 757–768.
Tarabykina, S., Kriajevska, M., Scott, D. J., Hill, T. J., Lafitte, D., Derrick, P. J., Dodson, G. G., Lukanidin, E. & Bronstein, I. (2000). *FEBS Lett.* **475**, 187–191.
Tarabykina, S., Scott, D. J., Herzyk, P., Hill, T. J., Tame, J. R. H., Kriajevska, M., Lafitte, D., Derrick, P. J., Dodson, G. G. & Maitland, N. J. (2001). *J. Biol. Chem.* **276**, 24212–24222.
Wang, G., Rudland, P. S., White, M. R. & Barraclough, R. (2000). *J. Biol. Chem.* **275**, 11141–11146.
Watanabe, Y., Kobayashi, R., Ishikawa, T. & Hidaka, H. (1992). *Arch. Biochem. Biophys.* **292**, 563–569.
Zimmer, D. B., Cornwall, E. H., Landar, A. & Song, W. (1995). *Brain Res. Bull.* **37**, 417–429.