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

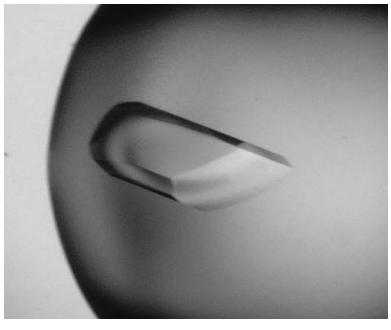
Expression, purification and crystallization of a human protein SH3BGRL at atomic resolution

Lei Yin,[a,b] De-Yu Zhu,[a,b] Na Yang,[a,b] Qiu-Hua Huang,[c] Ying Zhang[a] and Da-Cheng Wang[a]*

[a]Center for Structural and Molecular Biology, Institute of Biophysics, Chinese Academy of Science, Beijing 100101, People's Republic of China, [b]Graduate School of the Chinese Academy of Sciences, Beijing 100039, People's Republic of China, and [c]Shanghai Institute of Hematology, Rui Jin Hospital, Shanghai Second Medical University, Shanghai 200025, People's Republic of China

Correspondence e-mail: dcwang@sun5.ibp.ac.cn

Received 8 September 2004
Accepted 8 March 2005
Online 24 March 2005



The protein SH3BGRL, containing both SH3-binding and Homer EVH1-binding motifs, has been crystallized using the hanging-drop vapour-diffusion method. The crystals diffract to 0.88 Å resolution and belong to space group $P2_12_12_1$, with unit-cell parameters a = 28.8886, b = 34.9676, c = 98.0016 Å. Preliminary analysis indicates that the asymmetric unit contains one molecule and has a solvent content of about 34%.

1. Introduction

The SH3-domain-binding glutamic acid-rich protein-like protein (SH3BGRL) is a member of the new mammalian protein family SH3BGR (Egeo *et al.*, 1998; Mazzocco *et al.*, 2002). The gene is apparently widely expressed and is mapped on chromosome Xq13.3 in the region Xq13.3–Xq21.31, which is associated with an X-linked mental retardation syndrome. Using EBV-transformed lymphoblastoid cells from a mentally retarded male patient, it was confirmed that three genes were indeed deleted including SH3BGRL in the patient (Sudbrak *et al.*, 2001).

Recently, the SH3BGRL gene was cloned from human haemopoietic stem cells in our laboratories. It encoded a small protein of 114 amino acids, which was predicted to adopt the thioredoxin fold (Mazzocco *et al.*, 2002). It is characterized by the presence of a proline-rich sequence (PLPPQIF), which contains both SH3-binding (P*XX*P; Powson & Scott, 1997; Feng *et al.*, 1994; Cicchetti *et al.*, 1995) and Homer EVH1-binding (PP*XX*F; Beneken *et al.*, 2000; Niebuhr *et al.*, 1997) motifs. Proline-rich peptide sequences have been shown to play important roles in the protein–protein interactions involved in signal transduction pathways (Powson & Scott, 1997; Cohen *et al.*, 1995). Structural study of SH3BGRL and its binding to the SH3 domain is important to further our understanding of its function and its association with mental retardation. Here, we report the crystallization and preliminary crystallographic analysis of SH3BGRL at atomic resolution (0.88 Å).

2. Materials and methods

2.1. Gene cloning and expression

The coding sequence for protein human SH3BGRL (12.8 kDa) was amplified from RT-PCR products from human haemopoietic stem cells using the polymerase chain reaction (PCR) method. The PCR product was purified and restricted with *Nde*I and *Xho*I. The fragment was purified and ligated into *Nde*I- and *Xho*I-restricted sites of the pET22b(+) vector (Novagen Inc.) with a His tag. A further transformation into *Escherichia coli* strain XL1-Blue MRF′ competent cells was performed. The positive clones with an insert of the right size were identified by double digestion with *Nde*I and *Xho*I and further by DNA sequencing. The recombinant plasmid was transformed into *E. coli* strain BL21(DE3) for protein expression.

2.2. Protein expression and purification

The protein was overexpressed in *E. coli* strain BL21(DE3). The cells were grown at 310 K in 500 ml LB medium containing

100 µg ml^{-1} ampicillin until the OD_{600} value reached 0.6. Isopropyl β-thiogalactopyranoside (IPTG) was added to a final concentration of 1 m*M* and culture continued at 310 K for a further 4 h. Cells were collected by centrifugation at 4000 rev min^{-1} for 30 min. The pellet was resuspended in 20 ml lysis buffer (50 m*M* NaH_2PO_4 pH 8.0, 300 m*M* NaCl, 10 m*M* imidazole) and sonicated. The lysate was clarified by centrifugation at 16 000 rev min^{-1} for 10 min at 277 K to remove the cell debris. The supernatant was applied to an Ni^{2+}-chelating column (Novagen Inc.) and the contaminant protein was washed off with wash buffer (50 m*M* NaH_2PO_4 pH 8.0, 300 m*M* NaCl, 20 m*M* imidazole). The target protein was eluted with elution buffer (50 m*M* NaH_2PO_4 pH 8.0, 300 m*M* NaCl, 250 m*M* imidazole). The protein was concentrated with an Amicon Ultra 10 kDa molecular-weight cutoff filter unit (Millipore) and was purified on a Superdex-75 column (Pharmacia) in buffer *A* (50 m*M* Tris–HCl pH 8.0). A further purification step was performed using a Mono Q HR 5/5 column (Pharmacia) in buffer *A*. The protein was desorbed from the column in 10 m*M* NaCl.

2.3. Crystallization and data collection

The purified protein was desalted and concentrated to about 30 mg ml^{-1} in ultrapure water for crystallization. All crystallization experiments were performed with the hanging-drop vapour-diffusion method using Hampton Research Index and Crystal Screens. 1.5 µl protein solution and 1.5 µl reservoir solution were mixed and equilibrated with 400 µl reservoir solution in each well at 293 K. Screening different concentrations of precipitating agent and pH further optimized the initial crystallization condition, which was 1.4 *M* trisodium citrate and 0.1 *M* Tris pH 7.5. The best crystallization condition was obtained with a 400 µl reservoir solution containing 1.3 *M* trisodium citrate, 0.1 *M* Tris pH 8.0 at 293 K. A crystal about 0.2 × 0.3 × 0.8 mm in size was used for data collection.

All diffraction data were collected at beamline 5 of the Photon Factory (Tsukuba, Japan) using an ADSC Quantum-315 CCD detector. All crystals were briefly soaked in paraffin oil (Hampton Research) after being mounted in nylon cryoloops (Hampton Research) and were then flash-cooled in a nitrogen-gas stream at 95 K. The data set was collected at a wavelength of 0.90000 Å with a crystal-to-detector distance of 70 mm and 0.3° oscillation per frame. Each frame was exposed for only 1 s and 999 frames were collected. *MOSFLM* (v.6.2.2), *TRUNCATE* and *SCALA* from the *CCP*4 program suite v.4.2.2 (Collaborative Computational Project, Number 4, 1994) were used for processing, reduction and scaling of the diffraction data.

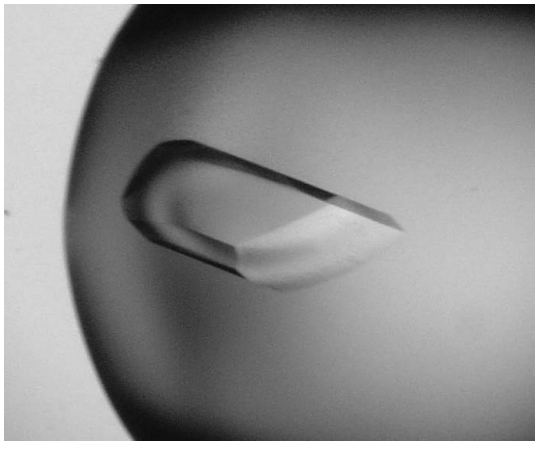


Figure 1
A crystal of SH3BGRL. The crystal is about 0.2 × 0.4 × 0.9 mm in size.

Table 1
Data-collection statistics.

Values in parentheses are for the highest resolution shell (0.93–0.88 Å).

Wavelength (Å)	0.90000
Space group	$P2_12_12_1$
Unit-cell parameters (Å)	a = 28.8886, b = 34.9676, c = 98.0016
Molecules per AU	1
Matthews coefficient (Å^3 Da^{-1})	1.9
Solvent content (%)	34
Resolution limit (Å)	20.275–0.88
Total reflections	841102 (115745)
Unique reflections	78866 (11215)
Average redundancy	10.7 (10.3)
Average $I/\sigma(I)$	5.9 (2.0)
Completeness (%)	99.0 (97.6)
R_{sym} (%)	6.3 (34.0)

3. Results and discussion

The gene SH3BGRL was successfully cloned in the pET22b(+) vector (Novagen Inc.) with a His tag. By transforming the plasmid into *E. coli* strain BL21(DE3), recombinant protein was expressed at a level of about 30 mg in 0.5 l LB liquid medium.

Hampton Research Index and Crystal Screens were used to supply conditions for initial screening. The purity and concentration were found to be very important in the crystallization of this protein. The first screening experiments with SH3BRGL were unsuccessful when using a protein concentration of 10–20 mg ml^{-1}. Most of the screening conditions were still clear after two weeks. We increased the concentration to 50–60 mg ml^{-1} with similar results. The protein was then further purified using a Mono Q HR 5/5 column (Pharmacia). After passage through an Ni^{2+}-chelating column (Pharmacia) and a Superdex-75 column (Pharmacia), the purified protein showed one band on SDS–PAGE and contained three main protein peaks following ion-exchange chromatography on Mono Q. The largest protein peak was collected and used in screening trials at a concentration of 50–60 mg ml^{-1}. After two weeks, crystals were obtained in a solution containing 1.4 *M* trisodium citrate and 0.1 *M* Tris pH 7.5. The best crystal was obtained by screening different concentrations of precipitating agent and pH (Fig. 1). The crystal proved to belong to space group $P2_12_12_1$, with unit-cell parameters a = 28.8886, b = 34.9676, c = 98.0016 Å, $\alpha = \beta = \gamma = 90°$ (Table 1). These data are compatible with the presence of one protein molecule in the asymmetric unit, with a Matthews coefficient of 1.9 Å^3 Da^{-1} and a solvent content of 34%. The crystals diffracted to an ultrahigh resolution of 0.88 Å (Fig. 2).

As a potential signalling protein associated with mental retardation, the SH3-binding and EVH1-binding motifs of SH3BGRL should be the important regions for its signal transduction pathway. The SH3-binding motif (PLPP) present in SH3BGRL is similar to the SH3 peptide ligand found in the 3BP-1 protein (Cicchetti *et al.*, 1995). A proline residue located five positions upstream from the first proline of the SH3 ligand in the N-terminal direction is considered to be critical for 3BP-1 function and is also present at the same position in human SH3BGRL (Egeo *et al.*, 1998). Therefore, exploration of the structure and function of SH3BGRL is certainly important. The three-dimensional structure of SH3BGRL at atomic resolution will provide the basis for in-depth studies of its structure–function relationship. Based on the present results, the structure determination of SH3BGRL is now in progress. The phase has been determined using

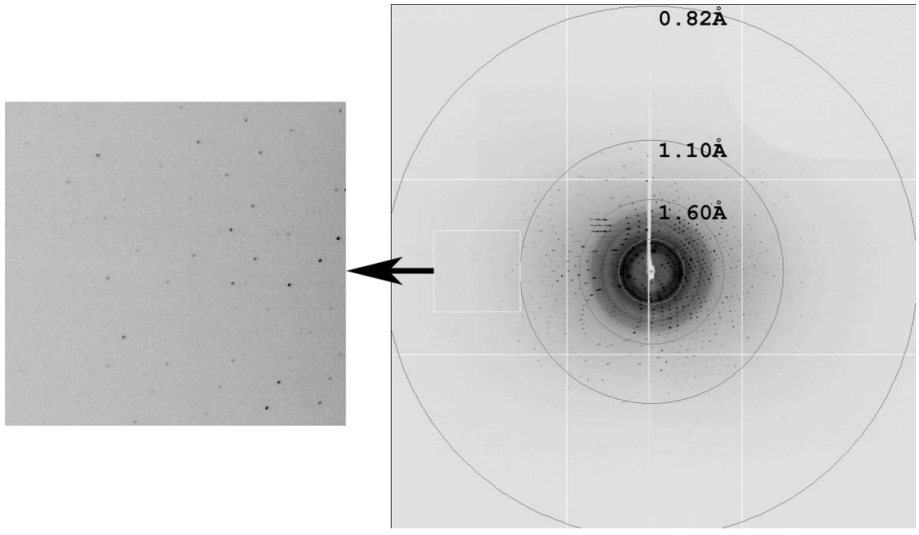


Figure 2
A typical diffraction pattern of an SH3BGRL crystal. The image on the left extends to a resolution limit of 0.88 Å (data were collected with a crystal-to-detector distance of 70 mm, 0.3° oscillation and 1 s exposures).

an iodine derivative using the SIRAS method and the model is now being refined.

This work was supported by grants from Chinese Academy of Science (KSCX2-SW-322, KSCX1-SW-17) and the MOST of China (863-2002BA711A13, 2004CB520801). The X-ray data collection was supported by the Photon Factory, KEK, Japan (00G209). We thank Professor F. Jiang and his laboratory for providing the in-house X-ray facility for preliminary X-ray analysis.

References

Beneken, J., Tu, J. C., Xiao, B., Nuriya, M., Yuan, J. P., Worley, P. F. & Leahy, D. J. (2000). *Neuron*, **26**, 143–154.
Cicchetti, P., Mayer, B. J., Thiel, G. & Baltimore, D. (1995). *Science*, **257**, 803–806.
Cohen, G. B., Ren, R. & Baltimore, D. (1995). *Cell*, **80**, 237–248.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Egeo, A., Mazzocco, M., Arrigo, P., Vidal-Taboada, J. M., Oliva, R., Pirola, B., Giglio, S., Rasore-Quartino, A. & Scartezzini, P. (1998). *Biochem. Biophys. Res. Commun.* **247**, 302–306.
Feng, S., Chen, J. K., Yu, H., Simon, J. A. & Schreiber, S. L. (1994). *Science*, **266**, 1241–1247.
Mazzocco, M., Maffei, M., Egeo, A., Vergano, A., Arrigo, P., Lisi, R. D., Ghiotto, F. & Scartezzini, P. (2002). *Gene*, **291**, 233–239.
Niebuhr, K., Ebel, F., Frank, R., Reinhard, M., Domain, E., Carl, U. D., Walter, U., Gertler, F. B., Wehland, J. & Chakraborty, T. (1997). *EMBO J.* **16**, 5433–5444.
Powson, T. & Scott, J. D. (1997). *Science*, **278**, 2075–2080.
Sudbrak, R., Wieczorek, G., Nuber, U. A., Mann, W., Kirchner, R., Erdogan, F., Brown, C. J., Wöhrle, D., Sterk, P., Kalscheuer, V. M., Berger, W., Lehrach, H. & Ropers, H. H. (2001). *Hum. Mol. Genet.* **10**, 77–83.