FOR THE RECORD

Use of a fusion protein to obtain crystals suitable for X-ray analysis: Crystallization of a GST-fused protein containing the DNA-binding domain of DNA replication-related element-binding factor, DREF

MUNEKAZU KUGE,¹ YOSHIFUMI FUJII,¹ TOSHIYUKI SHIMIZU,¹
FUMIKO HIROSE,² AKIO MATSUKAGE,² AND TOSHIO HAKOSHIMA¹

¹Department of Molecular Biology, Nara Institute of Science and Technology (NAIST),
8916-5 Takayama, Ikoma, Nara 630-01, Japan

²Laboratory of Cell Biology, Aichi Cancer Center Research Institute, Chigusa-Ku, Nagoya 464, Japan

(RECEIVED March 3, 1997; ACCEPTED May 7, 1997)

Abstract: Crystals of glutathione-S-transferase (GST)-fused protein containing the DNA-binding domain of DNA replication-related element-binding factor, DREF, were obtained under crystallization conditions similar to those for GST. Preliminary X-ray crystallographic analysis revealed that crystals of the GST-fused protein belong to space group $P6_122$ or $P6_522$ with unit cell dimensions $a = b = 140.4$ Å, $c = 93.5$ Å and $\gamma = 120^\circ$, having one molecule in the crystallographic asymmetric unit. The crystals diffract to 2.5 Å resolution. The cell dimensions are related to those of GST crystals thus far reported. Crystallization of the DNA-binding domain that was cleaved from the fused protein by thrombin was also carried out using several methods under numerous conditions, but efforts to produce well-ordered large crystals were unsuccessful. A possible application of GST-fusion proteins for small target proteins or domains to obtain crystals suitable for X-ray structure determination is proposed.

Keywords: cell cycle; crystal engineering; crystallization; DNA-binding domain; DNA replication; DREF; fusion protein; transcription factor

Crystallization is the rate-limiting step for determination of the three-dimensional structure of protein. Much effort has been made to understand various aspects of the nucleation and crystallization of proteins, including the formation of crystal defects. Insights into the interactions of proteins in solution were obtained from dynamic light scattering (DSL) and small angle X-ray scattering studies and provide a possible guide to the search for crystallization conditions (Rosenberger et al., 1996). The screening of crys-

tallization conditions might also be performed using accumulated knowledge on successful crystallization solutions and detergents (Jancarik & Kim, 1991; Weber, 1991; Cudney et al., 1994). In the material sciences, crystal engineering has led to the rational design of polymer crystals containing three-dimensional frameworks and also to the rational design of inclusion crystals of small organic compounds. Recently, the approach of co-crystallizing proteins with antibody fragments has been applied to the production of well-ordered crystals both of a ternary complex of human immunodeficiency virus type 1 reverse transcriptase bound to double-stranded DNA (Hacobo-Molina et al., 1993) and of a membrane protein, cytochrome *c* oxidase (Ostermeier et al., 1995). In this paper, we report a novel approach to crystallization of a domain as part of a fusion protein. We easily crystallized a glutathione-S-transferase (GST) fusion protein containing the DNA-binding domain of a transcription factor regulating DNA-replication related genes.

DNA replication is one of the most important processes for cell proliferation. Expression of genes involved in DNA replication is closely correlated with the proliferating state of cells and is repressed in accordance with the progression of differentiation in various tissues during development (Matsukage et al., 1994). DNA replication-related element (DRE) consists of an eight-base pair palindromic sequence, TATCGATA. It has been found in the upstream regions of *Drosophila* genes for both proliferating cell nuclear antigen (PCNA) (Yamaguchi et al., 1991) and the DNA polymerase α (Hirose et al., 1991). Transient expression assays in cultured *Drosophila* Kc cells indicated that DRE functions as a positive *cis*-acting element (Hirose et al., 1993). It was also shown that a specific DRE-binding factor (DREF) is required for high expression of DNA polymerase α and PCNA genes. Immunocytochemical analysis demonstrated the presence of DREF in nuclei after the eighth nuclear division cycle, suggesting that nuclear accumulation of DREF is important for the coordinate zygotic

Reprint requests to: Toshio Hakoshima, Department of Molecular Biology, Nara Institute of Science and Technology (NAIST), 8916-5 Takayama, Ikoma, Nara 630-01, Japan; e-mail: hakosima@bs.aist-nara.ac.jp.

expression of DNA replication-related genes carrying DRE sequences (Hirose et al., 1996). DREF consists of 701 amino acid residues with three characteristic regions rich in basic, proline, and acidic residues, respectively. The basic region located at the N-terminus is responsible for DRE binding. GST fusion protein containing the DNA-binding domain (16–115 amino acids) exhibits the DNA-binding activity.

Results: Microcrystals of the DNA-binding domain with the maximum size of smaller than 50 μm were obtained by microdialysis from the solution containing 25 mg/mL of the protein, 5 mM Bis-Tris-HCl (pH 6.0), 10 mM dithiothreitol (DTT), and 100 mM KCl, against a solution containing the same buffer with 0–70 mM KCl. Several efforts to produce well-ordered and bigger crystals, however, were unsuccessful. By contrast, crystallization of GST-DREF(105) fusion protein was successful. The best crystals grew to a maximum size of 150 μm $\times$ 150 μm $\times$ 2000 μm from a droplet containing 7.5 mg/mL protein, 25 mM potassium phosphate buffer (pH 5.2), 50 mM KCl, 2.5% (w/v) polyethylene glycol (PEG) 3350, and 5% ethylene glycol, against a reservoir solution containing 5% (w/v) PEG 3350, 10% ethylene glycol, and 50 mM potassium phosphate buffer. Gel electrophoresis of a solution in which crystals were redissolved showed that the crystals contain the fusion protein. The crystallization condition and the rodlike shape of crystals resemble those of the parasitic helminth GST molecule fused with a six amino acid of gp41, which was crystallized in a solution containing potassium phosphate buffer (pH 6.5) with PEG 3350 (Lim et al., 1994). Needle crystals were also obtained from droplets under similar conditions in a pH range of 4.6 to 6.0. Crystals for cryo-crystallography were transferred into cryo-protectant droplets containing 10% PEG 3350 and 25% ethylene glycol.

The crystals of the fusion protein diffracted to 2.5 Å resolution at -170°C , whereas the resolution limit of the crystals at 10°C was 3.0 Å. The crystals belong to space group $P6_122$ or $P6_522$ with unit cell dimensions $a = b = 140.42$ Å, $c = 93.51$ Å and $\gamma = 120^\circ$. The unit cell dimensions at 10°C were $a = b = 141.6$ Å, $c = 96.2$ Å and $\gamma = 120^\circ$. The space group was different from that of parasitic helminth GST, which belongs $P4_12_12$ with unit cell dimensions, $a = b = 94.7$ Å, and $c = 58.1$ Å (Lim et al., 1994). However, the lattice was somewhat related to that of the crystal of squid gland GST bound to S-(3-iodobenzene)-glutathione: $a = b = 73.15$ Å, and $c = 94.88$ Å in space group of $P3_121$ (Jin et al., 1995). The difference between the c -axes of these crystals is less than 1.5%, and the a - and b -axes of the present crystal are almost double those of squid GST crystal. The crystal packing of the squid GST shows large channel around the 3_1 axis at the origin. The C-terminus of each GST molecule faces to this channel, whose space seems to be enough to accommodate an additional 100 residues. The crystal has a rather large volume of solvent, 58% of the cell volume (Matthews, 1968). The addition of 100 residues to each GST gives a V_{solv} of 0.41, which is still an acceptable value for protein crystals. The crystallization conditions of these crystals are different. Squid GST was crystallized from a solution containing ammonium sulfate and Tris-HCl at pH 7.0. Diffraction data were collected at 2.5 Å resolution with an R -merge of 7.76% and a completeness of 86.9%; that of the outer shell from 3 Å to 2.5 Å is 77.1%. This crystal was estimated to contain one molecule in the asymmetric unit with a V_m value of 3.54 Å³/Da and a V_{solv} value of 65.3% (Matthews, 1968).

Discussion: Several GSTs from different sources and classes were recently crystallized and most of them formed well-ordered crystals that diffracted at a high resolution. These GSTs include parasitic helminth GST at 2.5 Å resolution (Lim et al., 1994), squid gland class-sigma GST at 2.2 Å resolution (Jin et al., 1995), rat liver class-mu GST at 2.2 Å resolution (Jin et al., 1992), human muscle class-mu GST at 1.85 Å resolution (Raghunathan et al., 1994), mouse liver class-pi GST at 1.8 Å resolution (Garcia-Saez et al., 1994), and human class-alpha GST at 2.6 Å resolution (Sinning et al., 1993). In general, sequence identities are high among GSTs within the same class, typically in the range of 50–80%, while inter-class identities are considerably less, about 25–35%. Parasitic helminth GST shares only 24% identities with squid GST. At the present time, it is unclear that use of GST fusion proteins always increases the chances for growing well-ordered crystals through intermolecular interactions of GST portions. However, comparison of the crystal structure of the present GST fusion protein with known crystal-packing structures of GSTs could elucidate which surfaces of the GST proteins are responsible for establishing the ordered three-dimensional crystal lattice. Such knowledge may enable further crystal engineering for obtaining well-ordered crystals of GST fusion proteins containing different target domains.

Since the molecular replacement method using the known structure of GST molecule can be applied for these crystals of fusion proteins, structure determination can be made much easier than for the crystals of target proteins alone. A recent structure determination by molecular replacement method revealed the structure of an additional domain of more than 110 residues that were not included in the initial model comprising 60% of the molecular mass of the target protein (Rudenko et al., 1996).

The expression system of GST fusion proteins in *Escherichia coli* cells was originally developed for single-step purification of proteins or domains using GST-affinity column chromatography (Smith & Johnson, 1988; Smith, 1993). More recently, GST fusion proteins immobilized on a glutathione affinity matrix were used in a GST pulldown assay to determine whether a protein directly interacts with another protein. Furthermore, GST fusion proteins have often been used in protecting expressed unstable proteins from proteolysis. For these reasons, the expression system of GST fusion proteins in *E. coli* cells has become popular among molecular and cell biologists. The system is more convenient than large-scale preparation of antibody fragments for co-crystallization (Hacobo-Molina et al., 1993; Ostermeier et al., 1995).

A structure analysis of the crystal using both the molecular replacement and multiple isomorphous replacement methods is in progress.

Materials and methods: Protein purification: Construction of *E. coli* strain and plasmid used in this study was described previously (Hirose et al., 1996). The plasmid pET GST-DREF(105) carries a gene encoding the DNA-binding domain of DREF under the T7 promoter. The fusion protein includes parasitic helminth class-alpha GST from *Schistosoma japonicum* (Smith & Johnson, 1988; Smith, 1993). The linker between the GST portion and the DNA-binding domain contains a cutting site for thrombin, consisting of seven amino acid residues, Asp-Leu-Val-Pro-Arg-Gly-Ser, having a cleaved peptide bond between Arg and Gly. *E. coli* BL21 (ADE3) cells harboring plasmid pET GST-DREF(105) were grown in 8 L of YT broth medium containing 50 $\mu\text{g/mL}$ ampicillin until

the OD₆₆₀ was about 0.45. Then, cells were induced with 0.5 mM isopropyl-1-thio- β -D-galactoside and cultured at 20°C for 12 h. These procedures enabled the expressed protein to remain in the soluble fraction. Cells (20 g) were suspended in 130 ml of 100 mM Tris-HCl buffer (pH 7.2) containing 400 mM KCl, 10 mM DTT, 0.1 mM phenylmethane sulfonyl fluoride PMSF, 5% glycerol, and 1% Triton X-100, and were disrupted by sonication at 4°C. The supernatant was applied onto a DEAE sepharose FF column (Pharmacia-LKB). The flow-through fractions were recovered and were applied onto a GST sepharose 4B column (Pharmacia-LKB). The column was washed a few times with 50 mM phosphate buffer containing 1 M KCl. Fractions containing GST-DREF(105) were eluted with 50 mM glutathione and dialyzed overnight. The fractions were then applied onto a Mono S column (Pharmacia-LKB). Protein was eluted with a 60 ml linear NaCl gradient from 0 to 1 M. At every step, the protein was monitored by 15% sodium dodecyl sulfate polyacrylamide gel electrophoresis and gels were stained with Coomassie Brilliant Blue. The DNA-binding domain was also purified by a similar procedure, in which the GST sepharose 4B column that adsorbed the GST-DREF(105) fusion protein was incubated with thrombin. Amino-terminal analysis of the DNA-binding domain was carried out with an automatic analyzer 475A (Applied Biosystems). The analysis showed that the DNA-binding domain was cleaved at the proper site of the linker peptide, i.e., between Arg and Gly as described above.

Protein characterization: Gel filtration of the DNA-binding domain of DREF and its GST fusion protein using a Superose 12 HR column (Pharmacia-LKB) indicated that these proteins exist as a dimer in solutions (data not shown). Dynamic light scattering (DSL) was measured with DynaPro 801 (Protein Solutions Co., USA). The DSL of both proteins suggested that solutions of these proteins exhibit monodispersity with less than 15% deviations of the hydrodynamic radii on the monomodal analyses. Hydrodynamic radius of the DNA-binding domain was estimated between 20 Å and 23 Å, which corresponds to a globular dimer with an estimated molecular mass of 20 kDa. Circular dichroism (CD) of the proteins was measured with a JASCO 720W spectrometer. CD spectra of the DNA-binding domain indicated that the protein structure is stable at a pH value from 5.0 to 9.0.

Protein crystallization: Screening of crystallization conditions of the DNA-binding domain of DREF was performed by several methods—including microbatch by the IMPAX system, microdialysis, and vapor-diffusion techniques—using numerous kinds of crystallization solutions containing different precipitants, metal ions, and detergents, and different buffers with different pH. Crystallization of GST-DREF(105) fusion protein was carried out by a microbatch method using the IMPAX system with standard crystallization solutions (Hampton Research). Then, crystallization conditions were refined by a hanging-drop vapor-diffusion method. Crystals obtained from solutions containing the fusion protein were thoroughly washed with the reservoir and redissolved in an aliquot of a buffer. Then, the solution was analyzed by gel electrophoresis.

Preliminary X-ray analysis: X-ray diffraction data of the fusion protein crystal were collected with a Rigaku imaging-plate area-detector (R-Axis IV) using Cu-K α radiation ($\lambda = 1.54178$ Å) generated by a rotating anode generator (Rigaku FR-C) operating at 50 kV and 60 mA. The focus size of the X-ray beam was

100 μ m, and the beam was focused using a Supper double-focusing mirror. The distance from crystal to imaging plate was 150 mm. The data on 180° rotation were collected at a rate of 1.4° per 40 min for each image. The loop-mounted crystal (diameter, 1.5 mm) was flash-cooled using a Rigaku cooling device with liquid nitrogen gas. Intensities were evaluated with the program PROCESS (Higashi T, Rigaku).

Acknowledgments: We thank Mr. T. Ikegami for his technical assistance. This work was supported by Grants-in-Aid for Scientific Research on Priority Areas from the Ministry of Education, Science, Sports and Culture of Japan to A.M. and T.H.

References

- Cudney R, Patel S, Weisgraber K, Newhouse Y, McPherson A. 1994. Screening and optimization of strategies for macromolecular crystal growth. *Acta Cryst D50*:414–423.
- Garcia-Saez I, Parraga A, Phillips MF, Mantle TJ, Coll M. 1994. Molecular structure at 1.8 Å of mouse liver class pi glutathione S-transferase complexed with S-(pi-nitrobenzyl)glutathione and other inhibitors. *J Mol Biol* 237:298–314.
- Hacobo-Molina A, Clark AD, Williams RI, Nanni RG, Clark P, Ferris I, Hughes SH, Arnold F. 1993. Crystal of a ternary complex of human immunodeficiency virus type 1 reverse transcriptase with a monoclonal antibody Fab fragment and double-stranded DNA diffract X-rays at 3.5 Å resolution. *Proc Natl Acad Sci USA* 88:10895–10899.
- Hirose F, Yamaguchi M, Nishida Y, Matsutani M, Miyazawa H, Hanaoka F, Matsukage A. 1991. Structure and expression during development of *Drosophila melanogaster* gene for DNA polymerase α . *Nucleic Acids Res* 19:4991–4998.
- Hirose F, Yamaguchi M, Handa H, Inomata Y, Matsukage A. 1993. Novel 8-base pair sequence (*Drosophila* DNA replication-related element) and specific binding factor involved in the expression of *Drosophila* genes for DNA polymerase and proliferating cell nuclear antigen. *J Biol Chem* 268:2092–2099.
- Hirose F, Yamaguchi MK, Kuroda K, Omori A, Hachiya T, Ikeda M, Nishimoto Y, Matsukage A. 1996. Isolation and characterization of cDNA for DREF, a promoter-activating factor for *Drosophila* DNA replication-related genes. *J Biol Chem* 271:3930–3937.
- Jancarik J, Kim SH. 1991. Sparse matrix sampling: A screening method for crystallization of protein. *J Appl Cryst* 24:409–411.
- Jin X, Zhang P, Armstrong RN, Gilliland GL. 1992. The three-dimensional structure of a glutathione S-transferase from the mu gene class. Structural analysis of the binary complex of isozyme 3-3 and glutathione at 2.2-Å resolution. *Biochemistry* 31:10169–10184.
- Jin X, von Rosenvinge EC, Johnson WW, Tomarev SI, Piatigorsky J, Armstrong RN, Gilliland GL. 1995. Three-dimensional structure, catalytic properties, and evolution of a sigma class glutathione transferase from squid, a progenitor of the lens S-crystallins of cephalopods. *Biochemistry* 34:5317–5328.
- Lim K, Ho JX, Keeling K, Gilliland GL, Ji X, Ruker F, Carter DC. 1994. Three-dimensional structure of *Schistosoma japonicum* glutathione S-transferase fused with six amino acid conserved neutralizing epitope of gp41 from HIV. *Protein Sci* 3:2233–2244.
- Matthews, BW. 1968. Solvent content of protein crystals. *J Mol Biol* 33:491–497.
- Matsukage A, Hirose F, Yamaguchi M. 1994. Transcriptional regulation of DNA replication-related genes in cell growth, differentiation and oncogenesis. *Jpn J Cancer Res* 85:1–8.
- Ostermeier C, Essen IO, Michel H. 1995. Crystals of an antibody Fv fragment against an integral membrane protein diffracting to 1.28 Å resolution. *Proteins* 21:74–77.
- Raghuathan S, Chandross RJ, Kretsinger RH, Allison TJ, Penington CJ, Rule GS. 1994. Crystal structure of human class mu glutathione S-transferase GSTM2-2: Effects of lattice packing on conformational heterogeneity. *J Mol Biol* 338:815–832.
- Rosenberger F, Vekilov PG, Muschol M, Thomas BR. 1996. Nucleation and crystallization of globular proteins: What we know and what is missing. In: Ataka A, Miyashita T, eds. *Crystallization of biological macromolecules*, Tokyo, Japan: North-Holland, Elsevier. pp 1–27.
- Rudenko G, Bonten E, D'Azzo A, Hol WGJ. 1996. Structure determination of the human protective protein: Twofold averaging reveals the three-dimensional

- structure of a domain which was entirely absent in the initial model. *Acta Crystallogr D* 52:923–936.
- Sinning I, Kleywegt GJ, Cowan SW, Reinemer P, Dirr H, Huber R, Gilliland GL, Armstrong RN, Ji X, Board PG, Olin B, Mannervik B, Jones TA. 1993. Structure determination and refinement of human alpha class glutathione transferase A1-1, and a comparison with mu and pi class enzymes. *J Mol Biol* 232:192–212.
- Smith DB. 1993. Purification of glutathione S-transferase fusion proteins. *Meth Mol Cell Biol* 4:220–229.
- Smith DB, Johnson KS. 1988. Single-step purification of polypeptides expressed in *Escherichia coli* as fusions glutathione S-transferase. *Gene* 67:31–40.
- Yamaguchi M, Nishida Y, Moriuchi T, Hirose F, Hui C-C, Suzuki Y, Matsukage A. 1991. *Drosophila* PCNA/cyclin gene: Structure and expression during development and specific binding of homeodomain proteins to its 5'-flanking region. *Mol Cell Biol* 10:872–879.
- Weber PC. 1991. Physical principles of protein crystallization. *Adv Prot Chem* 41:1–36.