Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Cloning, expression, purification, crystallization and initial crystallographic analysis of transcription factor DksA from *Escherichia coli*

Marina N. Vassylyeva,[a,b] Anna A. Perederina,[c] Vladimir Svetlov,[d] Shigeyuki Yokoyama,[a,b,c,e] Irina Artsimovitch[d] and Dmitry G. Vassylyev[b,c]*

[a]Genomic Sciences Center, RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi Yokohama 230-0045, Japan, [b]Structurome Project, RIKEN Harima Institute at SPring-8, 1-1-1 Kouto, Mikazuki-cho, Sayo, Hyogo 679-5148, Japan, [c]Cellular Signaling Laboratory, RIKEN Harima Institute at SPring-8, 1-1-1 Kouto, Mikazuki-cho, Sayo, Hyogo 679-5148, Japan, [d]Department of Microbiology, The Ohio State University, 484 West 12th Avenue, Columbus, OH 43210, USA, and [e]Department of Biophysics and Biochemistry, Graduate School of Science, University of Tokyo, Bunkyo-Ku, Tokyo 113-0033, Japan

Correspondence e-mail:
dmitry@yumiyoshi.harima.riken.go.jp

Received 7 April 2004
Accepted 18 June 2004

The *Escherichia coli* gene encoding a regulator of stringent response and virulence, DksA, which contains a canonical Zn finger motif, was cloned and expressed, and the purified protein was crystallized by the hanging-drop vapor-diffusion technique in two different space groups, $P2_12_12_1$ (a = 91.32, b = 96.59, c = 117.48 Å) and $C222$ (a = 80.6, b = 115.1, c = 149.57 Å). The crystals belonging to space group $P2_12_12_1$, improved by macroseeding, diffract beyond 2.0 Å at a synchrotron. Three complete atomic resolution multiple anomalous dispersion diffraction data sets were collected from the same crystal of the $P2_12_12_1$ crystal form at the absorption edge for Zn atoms.

1. Introduction

The *dksA* gene was initially isolated in *Escherichia coli* as part of a locus that, when overexpressed, could suppress the temperature-sensitive growth and filamentation phenotypes of a *dnaK* deletion mutant (Kang & Craig, 1990). Subsequent analysis of genetic interactions of *dksA* alleles and multicopy suppression trials implicated *dksA* in a variety of cellular processes, including cell division, stringent response to amino-acid starvation, quorum sensing and expression of virulence factors in *E. coli*, *Pseudomonas aeruginosa*, *Salmonella typhimurium* and *Shigella flexneri* (Hirsch & Elliott, 2002; Jude *et al.*, 2003; Mogull *et al.*, 2001; Turner *et al.*, 1998; Webb *et al.*, 1999). The cellular target(s) of DksA and its mechanism of action remain obscure, and it has been suggested to act at the level of replication, transcription and translation. A possible link to transcription is indicated by the fact that the effect of *dksA* deletion on *rpoS* induction parallels that of the withdrawal of ppGpp (Hirsch & Elliott, 2002), an alarmone that primarily targets transcription and directly binds to RNA polymerase. On the other hand, genetic analysis suggested that DksA may primarily act during translation and tentatively mapped a DksA target to the beginning of the *rpoS* mRNA. However, DksA effects on any of the processes it has been implicated in have not been characterized *in vitro*.

DksA is a 151 amino acid long (17 kDa) polypeptide encoded by a non-essential gene found in *E. coli* and related bacteria; DksA has no notable paralogs except poorly characterized TraR proteins and uncharacterized proteins from P2 and 186 phages (Doran *et al.*, 1994). Computer-assisted analysis of the DksA sequence reveals a prominent coiled-coil in the N-terminus and a conserved C4-Zn finger motif in the C-terminus. The presence of the Zn finger motif led to a suggestion that DksA may act as a transcriptional regulator (Bass *et al.*, 1996), perhaps in conjunction with ppGpp; this motif is also consistent with the proposed mechanism of DksA action through binding to the RNA (Ishihama, 2000).

To elucidate the mechanism of action of DksA we sought to determine its three-dimensional structure. To this end, we have cloned and expressed the *dksA* gene from *E. coli*. The purified DksA protein has been crystallized.

2. Experimental procedures and results

2.1. Cloning, expression and purification

DksA expression vector pVS11 was constructed by cloning of the *E. coli dksA* gene into a pTYB12 vector (New England Biolabs). A genomic copy of the *dksA* gene from DH5α strain was amplified using *Pfu* DNA polymerase-driven high-fidelity PCR with primers IA323 (5′-GACCGTGGCTGCAGT-TAGCCAGCCATCTGTTTTTC) and IA324 (5′-GGTGGTGGAATGCTATGCAAGAAG-GGCAAAACCGT), which include the recognition sites for the *Bsm*I and *Pst*I restriction endonucleases (New England Biolabs), respectively. *Bsm*I- and *Pst*I-digested PCR product was ligated to similarly treated dephosphorylated pTYB12 vector DNA; the sequence of the resulting construct was verified by automated sequencing (Genewiz Inc., NJ). In the resulting plasmid, DksA is expressed under control of the T7 gene 10 promoter and is fused at its N-terminus to a chitin-binding protein/intein moiety. The DksA protein that is released upon intein-induced cleavage carries an additional Ala residue at its N-terminus.

© 2004 International Union of Crystallography
Printed in Denmark – all rights reserved

 DOI: 10.1107/S0907444904015033

For overexpression of the native protein, pVS11 was transformed into *E. coli* strain BL21(λDE3). DksA was expressed using the Overnight Express™ autoinduction system (Novagen); the majority of the fusion protein was found in and purified from the soluble fraction. Cells were harvested and disrupted by sonication in ImpactCN100 buffer (50 m*M* Tris–HCl pH 8.8, 100 m*M* NaCl, 1 m*M* EDTA, 0.1 m*M* PMSF) with EDTA-free Complete Protease Inhibitor cocktail (Roche) and 0.1% Tween-20. The cell lysate was loaded onto a disposable gravity column with chemically cross-linked chitin polymer (New England Biolabs) equilibrated with the same buffer. The column was washed with ten volumes of the ImpactCN100 buffer. The column was finally equilibrated with three volumes of ImpactCN100 with 50 m*M* 2-mercapto-ethanol (ME) and incubated overnight at room temperature to elicit an intein-mediated cleavage reaction. The cleaved-off protein was eluted with two to three column volumes of HepA buffer (50 m*M* Tris–HCl pH 6.9, 5% glycerol, 1 m*M* ME, 1 m*M* PMSF).

Fractions containing the protein of interest were combined and loaded onto a Q 16/10 column using an AKTA Prime system at 2 ml min^{-1}. The column was washed with five volumes of HepA buffer followed by a 200 ml NaCl gradient from 0 to 600 m*M*; DksA eluted at ~250 m*M* NaCl. The purified protein was concentrated to approximately 6 mg ml^{-1} using an Vivaspin 10 filter (Millipore). The yield was 13 mg of purified protein per litre of culture.

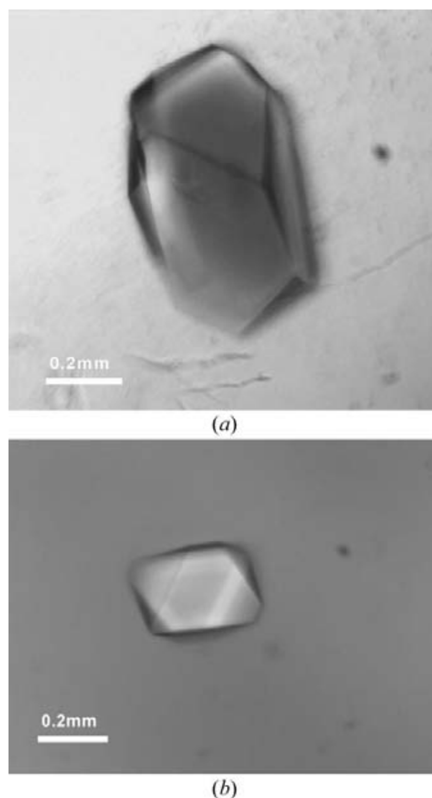

Figure 1
Crystals of the DksA protein. (*a*) $P2_12_12_1$ crystal form. (*b*) *C*222 crystal form.

2.2. Crystallization and data collection

A Hampton Crystal Screen kit (Jancarik & Kim, 1991) was used to determine the initial crystallization conditions for the DksA protein. Crystallization was carried out by the sitting-drop vapor-diffusion technique at 293 K, and the Crystal Screen precipitant solutions were diluted twofold by water in the initial screening. Aliquots (2 µl) of these diluted precipitant solutions were added to the 2 µl drops containing the protein solution, at a protein concentration of 6.8 mg ml^{-1}. After one week of equilibration, small plate-like crystals had occasionally grown in Crystal Screen I, solution 40. The crystals were subjected to micro- and then to macroseeding using drops prepared under the same conditions. After a few days of equilibration, crystals grew up to typical dimensions of 0.3 × 0.3 × 0.5 mm (Fig. 1*a*), and diffracted beyond 2.5 Å resolution at the in-house X-ray generator. They appeared to belong to space group $P2_12_12_1$, with unit-cell dimensions *a* = 91.32, *b* = 96.59, *c* = 117.48 Å. The final well crystallization solution was composed of 9% PEG 4000, 9% isopropanol, 45 m*M* Na citrate pH 5.6 and 0.15 *M* NaCl. At the same time, plate-like crystals (dimensions 0.3 × 0.3 × 0.05 mm) were grown from the same conditions without using seeding techniques (Fig. 1*b*). These crystals belong to space group *C*222, with unit-cell dimensions *a* = 80.6, *b* = 115.1, *c* = 149.57 Å. The final reservoir solution used for growing the *C*222 form crystals contained 10% PEG 4000, 10% isopropanol, 50 m*M* Na citrate pH 5.6 and 0.15 *M* NaCl. However, the diffraction quality of these crystals was substantially worse than that of the $P2_12_12_1$ crystal form, and the *C*222 form crystals were characterized by a high mosaicity of about 1.3°.

The DksA protein has four cysteine residues, indicating the presence of a canonical Zn finger motif. Zn atoms are known to provide a strong X-ray anomalous signal. Therefore, to phase the diffraction data using the multiple anomalous dispersion technique, the large DksA crystals were transferred into the standard protein crystallization drop to which 2 m*M* $ZnCl_2$ was added. Crystals continued growing in these drops for about two days. Among the crystals prepared for data collection at the synchrotron, one was substantially larger and diffracted apparently better than the others. Given the high intensity of the beamline used (NW12 beamline at Photon Factory, Tsukuba, Japan) we could not rule out the possibility of radiation damage during the data collection of the three complete data sets required for multiple anomalous dispersion phasing. Therefore, we chose to collect the two first data sets (peak and edge) with short exposures at medium resolution, while increasing the exposure time threefold for the last data set at the remote wavelength, if the two first data sets were not affected by radiation damage. This strategy proved correct, as the last (remote) high-resolution data set showed a non-significant but systematic increase in *B* factors during scaling, indicating the beginning of radiation damage that could be far more prominent if all data sets were collected with long exposures. Thus, a complete diffraction data set for the $P2_12_12_1$ crystal form was collected at 2.0 Å resolution at the remote wavelength (Table 1) for the Zn atoms, and two other complete data sets, at 2.65 Å resolution (data not shown), were collected from the same crystal at the peak and edge wavelengths using an ADSC Quantum-210 CCD detector. The data were collected at 100 K, using the mother liquor solution with 25% butandiol as cryoprotectant. The data were processed using the *HKL*2000 program suite

Table 1
Data-collection statistics for the DksA protein.

The data for the highest-resolution shell are shown in brackets.

Space group	$P2_12_12_1$
a (Å)	91.32
b (Å)	96.59
c (Å)	117.48
Matthews coefficient (Å^3 Da^{-1})	3.05
Solvent content (%)	58
Synchrotron beamline	NW12, Photon Factory, Tsukuba, Japan
Wavelength (Å)	1.0
Temperature (K)	100
Molecules in the unit cell	20
Molecules in asymmetric unit	5
Solvent content (%)	40–60
Resolution (Å)	40–2.0 (2.07–2.0)
Observations	353 953
Unique reflections	69 465
Multiplicity	5.1 (3.2)
R_{merge}†	0.046 (0.400)
Completeness (%)	98.4 (96.2)
$I/\sigma(I)$	28.0 (3.4)
Reflections with $I > 3\sigma(I)$ (%)	83.3 (54.2)

† $R_{merge} = \sum |I_j - \langle I_j \rangle| / \sum \langle I_j \rangle$, where I_j is the intensity of reflection *j* and $\langle I_j \rangle$ is the average intensity of reflection *j*.

(Otwinowski & Minor, 1997; Table 1). On the basis of a calculation of the Matthews (1968) coefficient, there may be between four and six DksA molecules in the asymmetric unit of the crystal, with the solvent content ranging from 40 to 65%, respectively. Given that one Zn^{2+} ion should be bound to one DksA monomer, the SnB (Weeks & Miller, 1999) calculations were performed probing different numbers (three to six) of Zn^{2+} ions in the search trials. The best result was obtained using five Zn^{2+} ions as a search model, indicating five DksA molecules in the asymmetric unit of the crystal. The structure determination is now in progress.

We thank Drs N. Igarashi, N. Matsugaki and M. Suzuki from the laboratory of Professor S. Wakatsuki (Photon Factory, Tsukuba, Japan) for assistance in data collection at the NW12 synchrotron beamline. This work was supported in part by grants from the National Institutes of Health (to IA), and from the RIKEN Structural Genomics/Proteomics Initiative (RSGI, the National Project on Protein Structural and Functional Analyses, Ministry of Education, Culture, Sports, Science and Technology of Japan) to SY and DGV.

References

Bass, S., Gu, Q. & Christen, A. (1996). *J. Bacteriol.* **178**, 1154–1161.
Doran, T. J., Loh, S. M., Firth, N. & Skurray, R. A. (1994). *J. Bacteriol.* **176**, 4182–6.
Hirsch, M. & Elliott, T. (2002). *J. Bacteriol.* **184**, 5077–5087.
Ishihama, A. (2000). *Annu. Rev. Microbiol.* **54**, 499–518.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Jude, F., Kohler, T., Branny, P., Perron, K., Mayer, M. P., Comte, R. & van Delden, C. (2003). *J. Bacteriol.* **185**, 3558–3566.
Kang, P. J. & Craig, E. A. (1990). *J. Bacteriol.* **172**, 2055–2064.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Mogull, S. A., Runyen-Janecky, L. J., Hong, M. & Payne, S. M. (2001). *Infect. Immun.* **69**, 5742–5751.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Turner, A. K., Lovell, M. A., Hulme, S. D., Zhang-Barber, L. & Barrow, P. A. (1998). *Infect. Immun.* **66**, 2099–2106.
Webb, C., Moreno, M., Wilmes-Riesenberg, M., Curtiss, R. III, & Foster, J. W. (1999). *Mol. Microbiol.* **34**, 112–123.
Weeks, C. M. & Miller, R. (1999). *J. Appl. Cryst.* **32**, 120–124.