Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary analysis of *Escherichia coli* YodA

Gabriel David,[a] Karine Blondeau,[b] Madalena Renouard[a] and Anita Lewit-Bentley[a]*

[a]LURE, CNRS, CEA, MR, BP 34, 91898 Orsay CEDEX, France, and [b]Institute of Genetics and Microbiology, Bâtiment 409, University Paris XI, 91405 Orsay, France

Correspondence e-mail: anita.bentley@lure.u-psud.fr

Received 21 September 2001
Accepted 8 May 2002

The *Escherichia coli* protein YodA was overexpressed, purified and crystallized in several crystal forms. The function of this protein is not known, although it has been identified under conditions of bacterial stress. Three of the four crystal forms were obtained in the presence of divalent cations (zinc, nickel and cadmium), suggesting that YodA may be a metal-binding protein.

1. Introduction

While attempting to overexpress the transacylase CyaC from *Bordetella pertussis* in *E. coli* as a host, we in fact overexpressed the *E. coli* protein YodA. We subsequently obtained good-quality crystals of this protein under several crystallization conditions. The gene *yodA* had been identified when the sequence of the *E. coli* genome was determined (Blattner *et al.*, 1997), although the function of its product, a 216-residue protein, remains unknown. Thus, the determination of the structure and subsequently the function of this protein can be considered as an exercise in the approaches developed in structural genomics projects.

To date, the *yodA* gene has been described in two studies, both concerned with the overall response of *E. coli* to different external stimuli. In the first case, Laurent-Winter *et al.* (1997) observed a decrease in YodA expression in *E. coli* mutants defective in the histone-like nucleoid-structuring protein (H-NS). In the second case, Ferianc *et al.* (1998) found an increase in YodA expression when submitting *E. coli* to increased levels of cadmium in the medium.

Bacteria have evolved mechanisms to respond to challenges from the environment over thousands of years. These include changes in temperature, oxidants, metals, toxins (*e.g.* antibiotics) and other conditions. In some cases, *E. coli* responds to stress by activating complex global regulatory systems depending on the stress stimulon (Vanbogelen *et al.*, 1987). Toxic elements are primarily removed by active extrusion, reducing their intracellular concentration to acceptable levels (Rosen, 1999). In many cases, the response involves the production of intracellular proteins capable of enzymatic degradation of the toxic molecules. In all cases, bacteria have evolved mechanisms of adaptation which allow them to resume growth after a period of stasis (Mitra *et al.*, 1975).

The response of bacteria such as *E. coli* to heavy metals, in particular cadmium, has been studied in some detail (Babai & Ron, 1998; Inbar & Ron, 1993; Ferianc *et al.*, 1998, 2000; Puskarova *et al.*, 2001). The efflux of the lethal metal is ensured by the product of *orf732* (*zntA*), homologous to *cadA* in *B. subtilis*, a P-type ATPase. While in eukaryotes heavy metals can be neutralized by binding to specific peptides (*e.g.* metallothiones), such compounds have not been found in *E. coli* (Rosen, 1999).

In their study of YodA, Ferianc and coworkers noted strong sequence similarities with YrpE from *B. subtilis* and pXO1-130 from *B. anthracis*. They suggested that these proteins form a family of bacterial cadmium stress-response proteins (Ferianc *et al.*, 1998; Puskarova *et al.*, 2001). They were also able to show a strong sequence similarity with the C-terminal part of adhesin AdcA from *Streptococcus pneumoniae* (Puskarova *et al.*, 2001). Interestingly, this protein was suggested to be a lipoprotein containing a metal-binding site and belonging to a new family of metal-binding bacterial receptors (Dintilhac & Claverys, 1997).

Finally, Ferianc *et al.* (1998) were able to show that YodA is processed in the bacterial cell by the cleavage of the N-terminal signal sequence of 24 residues. In subsequent work, the protein was localized in the periplasmic space of the bacterial cell wall (Ferianc, personal communication). In the present work, overexpression of YodA was obtained under a different type of stress conditions. The expression of this protein increased significantly during the induction in *E. coli* of a

heterologous protein (the transacylase CyaC from *Bordetella pertussis*) using a high cell-density culture strategy.

Here, we report the expression and crystallization of YodA alone and in the presence of the divalent cations cadmium, zinc and nickel.

2. Materials and methods

2.1. Protein expression and purification

YodA was obtained in an Applikon fermentor system in Fed-Batch cultures of the *E. coli* strain BL21λDE3-Gold (Stratagene), transformed with pET28b-CyaC-GroESL. Cultures of 1500 ml were carried out in a synthetic medium containing 10 g l^{-1} glycerol, 27.2 g l^{-1} KH_2PO_4, 4 g l^{-1} $(NH_4)_2SO_4$, 8 g l^{-1} KOH, 0.4 g l^{-1} $MgSO_4.7H_2O$, 100 mg l^{-1} $CaCl_2.2H_2O$, 1 mg l^{-1} $FeSO_4.7H_2O$, 0.4 mg l^{-1} $ZnSO_4.7H_2O$, 0.4 mg l^{-1} $MnSO_4$, 0.04 mg l^{-1} $CuSO_4.5H_2O$, 0.2 mg l^{-1} $NaMoO_4$, 0.5 mg l^{-1} H_3BO_3 and 0.1 mg l^{-1} KI. After 15 h at 310 K, pH 7.0, 800 rev min^{-1}, 1 v.v.m. (one volume of air per volume of culture per minute), 50 times concentrated medium was added to obtain a growth rate of 0.25 h^{-1} and a biomass yield of 0.5 g per gram of glycerol. The stirring rate was then increased to 1200 rev min^{-1} and the ventilation to 1.5 v.v.m. and the culture was subjected to a temperature increase to 315 K for 20 min prior to induction in order to produce endogenous chaperone proteins. The cells were induced with 1 m*M* isopropyl β-thio-galactopyranoside (IPTG) during the exponential phase of bacterial growth (when the biomass attained was 10 g l^{-1}) and the bacteria were harvested 2 h later. Extraction using a French press was followed by purification on an Ni–NTA column (Qiagen). The final purification was carried out by HPLC gel filtration (Waters carbamate Analysis System, with Waters Pak 125 column) in 20 m*M* MES pH 6.0 and 1 m*M* DTT. Once concentrated, the protein purity was checked on 15% SDS–PAGE gels stained with silver nitrate.

2.2. Crystallization

Initial crystallization conditions were obtained using the vapour-diffusion technique with Crystal Screen I (Hampton Research, USA), with the protein at 5 mg ml^{-1} in 20 m*M* MES pH 6.0 and 1 m*M* DTT. Crystals were obtained under two conditions: 200 m*M* sodium acetate, 100 m*M* sodium cacodylate pH 6.5, 30%(*w*/*v*) polyethylene glycol (PEG) 8000 and 200 m*M* zinc sulfate, 100 m*M* sodium cacodylate pH 6.5, 18%(*w*/*v*) PEG 8000. Two other forms were obtained in the presence of cadmium or nickel. The optimized conditions for all crystal forms are summarized in Table 1. The final protein concentration was 10 mg ml^{-1}. While form I crystals formed best at higher precipitant concentrations, form II crystals (co-crystallized with zinc) required the reduction of precipitant concentration. The crystals were soaked in the respective mother liquor containing 25% glycerol for data collection at 100 K. The best data for form II crystals were, however, obtained from crystals grown in the presence of glycerol and flash-cooled directly without any exchange of mother liquor.

Table 1
Summary of crystallization conditions for the different crystal forms.

Crystal form	Salt	Buffer	Precipitant	Morphology and size
Form I, native	200 m*M* NaOAc	100 m*M* sodium cacodylate pH 6.5	30%(*w*/*v*) PEG 10 000	Needles, 50 × 20 µm
Form II, zinc	200 m*M* $ZnSO_4$	100 m*M* sodium cacodylate pH 6.5	10%(*w*/*v*) PEG 10 000, 10%(*w*/*v*) glycerol	Bipyramids, 700 × 150 µm
Form III, cadmium	20 m*M* $CdCl_2$	100 m*M* sodium cacodylate pH 6.5	10%(*w*/*v*) PEG 10 000	Plates and needles, 100 × 40 µm
Form IV, nickel	20 m*M* $Ni(OAc)_2$	100 m*M* sodium cacodylate pH 6.5	18%(*w*/*v*) PEG 10 000	Bipyramids, 200 × 50 µm

Table 2
Summary of data-collection parameters and statistics for form I, II and III crystals.

Values in italics correspond to LURE data collections. Values in parentheses are for the high-resolution range.

	Form I (native)	Form II (zinc)	Form III (cadmium)
Site and source	Station 9.5 (DL)	BM30 (ESRF)	BM30 (ESRF)
	DW32 (LURE)	*DW32 (LURE)*	*DW32 (LURE)*
Wavelength (Å)	1.27	λ_1 = 1.28216, λ_2 = 1.28266, λ_3 = 1.27034	1.77122
	0.948	*0.953*	*0.966*
Crystal-to-detector distance (mm)	100	150	132
	200	*320*	*180*
Space group	$P2_1$	$P4_12_12$	$C2$
Unit-cell parameters (Å, °)	a = 40.35, b = 64.6, c = 41.5, β = 117.83	a = b = 58.4, c = 152.1	a = 107.1, b = 48.3, c = 42.65, β = 93.01
No. of molecules in asymmetric unit	1	1	1
Matthews coefficient V_M ($Å^3$ Da^{-1})	2.14	2.90	2.47
Resolution limit (Å)	1.81	2.7	3.05
	2.2	*2.27*	*2.1*
No. of unique reflections	12793	7599	3781
	7938	*6777*	*11894*
Completeness (%)	74.5 (74.5)	97.7 (97.7)	46 (45)
	100 (97.3)	*100 (71)*	*98.3 (90.6)*
R_{sym}	0.06 (0.43)	0.088 (0.246)	0.077 (0.07)
	0.132 (0.49)	*0.074 (0.19)*	*0.081 (0.272)*
$\langle I \rangle$	10.5 (1.6)	5.1 (2.9)	2.8 (14.7)
	8.5 (4.0)	*5.4 (1.5)*	*8.9 (2.7)*

† $R_{sym} = \sum_h \sum_i |I(h)_i - \langle I(h) \rangle| / \sum_h \sum_i [I(h)]$.

2.3. Data collection

All data from the three crystal forms were collected at cryogenic temperatures on the DW32 station at LURE using a MAR345 imaging plate. Furthermore, data from a form I crystal were collected on station 9.5 at the SRS, Daresbury, England with a MAR CCD detector. MAD data from form II crystals and SAD data from form III crystals were collected on beamline BM30 (FIP) at the ESRF, Grenoble, France with a 165 mm MAR CCD detector. The three wavelengths for the MAD experiment were determined using a fluorescence scan of the crystal. The ESRF data were treated using the *HKL* package (Otwinowski & Minor, 1997); the data from SRS and LURE were treated using *MOSFLM* (Leslie, 2001). Data statistics are presented in Table 2.

3. Results and discussion

The protein was identified as YodA using MALDI–TOF mass spectrometry and peptide digestion. The determination of its molecular weight as 22.3 kDa, *i.e.* 24 amino acids shorter than the sequence deduced from the gene, corroborates the finding of Ferianc *et al.* (1998) that YodA undergoes N-terminal processing.

Furthermore, the removal of the 24 N-terminal residues transforms YodA into a natural His-tag protein. The mature N-terminal sequence is HGHHSH, a nearly

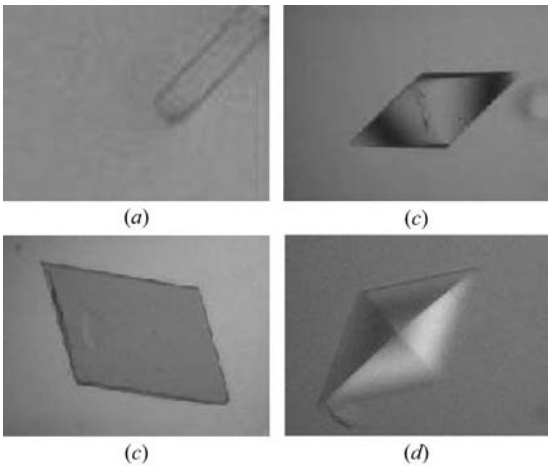

Figure 1
Photographs of the four crystal forms. (*a*) Form I, native (dimensions 50 × 20 × 20 µm), (*b*) form II, co-crystallized with zinc (dimensions 700 × 300 × 200 µm), (*c*) form III, co-crystallized with cadmium (dimensions 500 × 200 × 30µm), (*d*) form IV, co-crystallized with nickel (dimensions 300 × 80 × 30 µm).

perfect replica of the 6×His sequence commonly used as an affinity tag. This explains why YodA replaced the recombinant transacylase CyaC from *Bordetella pertussis* in the preparations: the purification was performed on an Ni–NTA affinity column because an N-terminal 6×His-tag had been added to the CyaC sequence. CyaC is, however, less well expressed under the conditions used and the more abundant (and more soluble) YodA was preferentially selected by the affinity column.

The protein crystallized readily under several conditions of the Hampton screens and, in particular, in the presence of zinc (Fig. 1). We therefore assume that the N-terminal histidine-rich sequence is very accessible to solvent. On the other hand, the presence of chlorides, in particular NaCl, caused precipitation of the protein. The fact that the crystal form is dependent on the presence and type of divalent cations indicates that their binding sites may be different. Patterson maps show three strong and two weaker zinc sites and one cadmium site per molecule of YodA.

We have used the programs *JPRED* (Cuff & Clamp, 1998) and *PRODOM* (Rost, 1996) for secondary-structure assignments. These indicate the presence of four α-helices and six β-strands in the sequence. All four α-helices and three of the β-strands occur in the sequences conserved in the three related proteins (*B. subtilis* YrpE and *B. anthracis* pXO1-130).

We have obtained YodA under strictly controlled conditions in fed-batch mode using fermentor cultures. When using the same medium in batch cultures in Erlenmeyer flasks, expression of YodA was not observed. We are therefore studying the effect of different parameters on *E. coli* cultures in order to understand the regulation of this protein.

Our serendipitous observation may be of dual interest in the context of structural genomics. Firstly, YodA is a protein of unknown function and in order to elucidate it we shall be employing the tools now being developed for structural genomics projects. Secondly, the combination of 6×His affinity tag, *E. coli* as a host organism and fermentor culture is being widely used in structural genomic projects. The expression of a rogue protein that mimics the desired gene product has to be understood in order to be avoided. Finally, since YodA expression has been linked to heavy-metal (or oxidative) stress, we hope to be able to shed further light on the physiology of this commonly used bacterium.

We thank Peter Sebo and Petr Halada, Microbiology Institute, Prague, Czech Republic and Paulette DeCotignies, IBBC, Université Paris-Sud, Orsay for carrying out mass-spectrometry measurements, and Robert Aufrère and Gilles Henckes, IGM, Université Paris-Sud, Orsay for their help with molecular biology. We are grateful to Peter Ferianc, Institute of Molecular Biology and Microbiology, Bratislava, Slovakia for useful discussions and for providing unpublished results.

References

Babai, R. & Ron, E. Z. (1998). *FEMS Microbiol. Lett.* **167**, 107–111.
Blattner, F. R., Plunkett, G., Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A., Goeden, M. A., Rose, D. J., Mau, B. & Shao, Y. (1997). *Science*, **277**, 1453–1474.
Cuff, J. A. & Clamp, M. E. (1998). *Bioinformatics*, **14**, 892–893.
Dintilhac, A. & Claverys, J.-P. (1997). *Res. Microbiol.* **148**, 119–131.
Ferianc, P., Farewell, A. & Nystrom, T. (1998). *Micobiology*, **144**, 1045–1050.
Ferianc, P., Puskarova, A., Godocikova, J., Polek, B. & Toth, D. (2000). *Biologica (Bratislava)*, **55**, 653–659.
Inbar, O. & Ron, E. Z. (1993). *FEMS Microbiol. Lett.* **113**, 197–200.
Laurent-Winter, C., Savaruth, N., Danchin, A. & Bertin, P. (1997). *Eur. J. Biochem.* **244**, 767–773.
Leslie, A. G. W. (2001). *MOSFLM Program for Autoindexing and Integrating X-ray Diffraction Data*, Version 6.11. MRC Laboratory of Molecular Biology, Hills Road, Cambridge, England.
Mitra, R. S., Gray, R. H., Chin, B. & Bernstein, I. A. (1975). *J. Bacteriol.* **121**, 1180–1188.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Puskarova, A., Janecek, S., Ferianc, P. & Polek, B. (2001). *Biologica (Bratislava)*, **56**, 337–339.
Rosen, B. P. (1999). *Essays Biochem.* **34**, 1–15.
Rost, B. (1996). *Methods Enzymol.* **266**, 525–539.
Vanbogelen, R. A., Kelley, P. M. & Neidhardt, F. (1987). *J. Bacteriol.* **169**, 26–32.