THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 278, No. 44, Issue of October 31, pp. 43728–43735, 2003
Printed in U.S.A.

YodA from *Escherichia coli* Is a Metal-binding, Lipocalin-like Protein*

Received for publication, April 29, 2003, and in revised form, July 30, 2003
Published, JBC Papers in Press, August 8, 2003, DOI 10.1074/jbc.M304484200

Gabriel David‡, Karine Blondeau§, Marc Schiltz‡¶, Simon Penel‡‖, and Anita Lewit-Bentley‡**

From the ‡Laboratoire pour l'Utilisation du Rayonnement Electromagnétique, CNRS, CEA, MdR, BP. 34, 91898 Orsay Cedex, France and the §Institut de Génétique et Microbiologie, Bâtiment 360, Centre Scientifique Paris-Sud, 91405 Orsay Cedex, France

We have determined the crystal structure of YodA, an *Escherichia coli* protein of unknown function. YodA had been identified under conditions of cadmium stress, and we confirm that it binds metals such as cadmium and zinc. We have also found nickel bound in one of the crystal forms. YodA is composed of two domains: a main lipocalin/calycin-like domain and a helical domain. The principal metal-binding site lies on one side of the calycin domain, thus making YodA the first metal-binding lipocalin known. Our experiments suggest that YodA expression may be part of a more general stress response. From sequence analogy with the C-terminal domain of a metal-binding receptor of a member of bacterial ATP-binding cassette transporters, we propose a three-dimensional model for this receptor and suggest that YodA may have a receptor-type partner in *E. coli*.

The *Escherichia coli* protein YodA has been identified as part of the response of the bacterium to a challenge with cadmium (1). Sequence analysis of YodA has shown a remarkable similarity with the putative bacterial proteins YrpE from *Bacillus subtilis*, pXO1–130 from *Bacillus anthracis*, and the C-terminal domain of AdcA from *Streptococcus pneumoniae*, as well as some similarity with the N-terminal domain of the copper-binding amine oxidase of *E. coli* and *Klebsiella aerogenes* (2). It was therefore proposed that the former group may represent a new family of cadmium stress response proteins.

Cadmium is an extremely toxic heavy metal not known to have any function within living organisms. Its toxicity is twofold: it can replace divalent cations such as zinc in metalloenzymes (3) and DNA-binding zinc finger proteins (4), thus altering their normal activity. It can also provoke an oxidative stress by causing the formation of highly active oxygen radical species. Its redox potential is higher compared with biologically important metal ions such as zinc, copper, and nickel (5). The oxidative stress can also be caused by cadmium replacing zinc in the active sites of proteins involved in redox pathways (6), thus blocking their function.

The response of microorganisms to heavy metals, such as mercury and cadmium, has been studied in some detail (1, 7–12). The toxic elements are primarily removed by active transport, reducing their intracellular concentration to acceptable levels (13). The extrusion of cadmium is ensured in *E. coli* by the product of *orf732* (*zntA*), homologous to *cadA* in *B. subtilis*, a P-type ATPase (14). In eucaryotes, heavy metals can be neutralized by binding to specific proteins, *e.g.* metallothioneins (15). Similar compounds have recently been discovered in bacteria, although the role of their homologue in *E. coli*, designated GatA, is not yet clear (16).

In their recent study of the expression pattern of YodA, Puskarova *et al.* (12) have identified the *yodA* promotor that is activated specifically by cadmium but no other heavy metal. YodA induction was shown to be dependent on *soxS*, *fur*, and *relA*/*spoT* of the stringent response but not on *oxyR*. YodA expression was also induced by hydrogen peroxide but after a time lag compared with the cadmium response. Although primarily a cytoplasmic protein in nonstressed cells, YodA was shown to be exported to the periplasm upon exposure to cadmium.

We have solved the three-dimensional structure of the mature protein in three crystal forms: in the presence of cadmium or zinc and with no added metal ions. The protein is composed of two domains: a major domain that is structurally related to the lipocalin/calycin family of proteins (17) and a smaller helical domain. The central metal-binding site lies along the side of the calycin domain, buried at the interface with the helical domain, consisting of histidine side chains, similar to metal-binding sites in some proteases and oxidoreductases. Lipocalins are a growing family of small, usually secreted proteins whose main structural feature is an up-down β-barrel (18). They have been found in all species from bacteria to man and have been associated with very diverse functions, often involved in the transport of small hydrophobic molecules. YodA can be classified as the first example of a metal-binding lipocalin implicated in response to bacterial stress.

EXPERIMENTAL PROCEDURES

Crystallography—The protein purification and crystallization has been described elsewhere (19). Multiple anomalous dispersion data were collected on crystals grown in the presence of 200 mM zinc sulfate on station BM30 (FIP) at the ESRF (Grenoble, France) to 2.7 Å resolution. A further data set to 2.4 Å resolution was collected at LURE (station DW32) at 0.953 Å wavelength and used as a fourth wavelength in the phasing. Four zinc positions per asymmetric unit were refined using SHARP (20), and after density modification with DM from the CCP4 program suite (21), the electron density map could be interpreted using O (22). The diffraction data were very anisotropic, resulting in very high temperature factors in the structure.

The first model obtained with experimental phases was used to

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (code 1OEK, 1OEE, and 1OEJ) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

¶ Present address: Ecole Polytechnique Fédérale, 1015 Lausanne, Switzerland.

‖ Present address: Laboratoire de Biométrie et Biologie Evolutive, Bât. 711,Université Lyon 1, 43 blvd. 11 Novembre 1918, 69622 Villeurbanne Cedex, France.

** To whom correspondence should be addressed. Tel.: 33-1-64468050; Fax: 33-1-64464148; E-mail: Anita.Bentley@lure.u-psud.fr.

TABLE I
Crystallographic data statistics

r.m.s. indicates root mean square. $R_{merge} = \Sigma_h\Sigma_i |I(h)_i - \langle I(h)\rangle|\Sigma_h\Sigma_i I(h)_i$, $R_{anom} = \Sigma_h |\langle I(h_+)\rangle - \langle I(h_-)\rangle|/\Sigma_h(\langle I(h_+)\rangle + \langle I(h_-)\rangle)$. Phasing power (PP) = F_H/E, where F_H is the r.m.s. mean heavy atom contribution and E is the r.m.s. residual. $R = \Sigma|P_o - F_c|/\Sigma F_o$, R_{free} is defined in Ref. 54; 5% exclusion was used. most, additional, and generous indicate the percentages of peptide bonds in the most favored, additional, and generously allowed region of the Ramachandran plot as determined by PROCHECK (55). Anom., anomalous; Isom., isomorphous; acen, acentric; cen, centric; F.O.M., figure of merit.

MAD data statistics (zinc crystal form, $P4_12_12$)[a]				
Wavelength (Å)	11.28216 (peak)	1.282667 (inflection)	1.27034 (remote)	0.953 (far remote)
Resolution (Å)	2.7	2.9	3.0	2.4 (anisotropic)
R_{merge}	0.072 (0.397)	0.089 (0.332)	0.092 (0.276)	0.086 (0.35)
R_{anom}	0.058 (0.198)	0.055 (0.127)	0.050 (0.113)	
$I/\sigma I$	7.9 (1.8)	6.5 (2.3)	6.4 (2.7)	8.5 (1.5)
Multiplicity	6.6 (5.9)	6.7 (7.0)	6.7 (6.9)	12 (9.5)
Completeness (%)	97.5 (96.9)	98 (97.6)	98.2 (97.8)	98.6 (71.1)
No. unique	7465	6156	5610	10,445
Phasing (to 2.7 Å resolution)				
Anom. PP, λ_{peak} (acen)	1.939			
Isom. PP, $\lambda_{inflection}$ (acen/cen)	0.387/0.346			
Anom. PP, $\lambda_{inflection}$ (acen)	1.927			
Isom. PP, λ_{remote} (acen/cen)	0.141/0.124			
Anom. PP, λ_{remote} (acen)	1.121			
Isom. PP, $\lambda_{far\ remote}$ (acen/cen)	0.639 / 0.548			
F.O.M. (acen/cen)	0.419/0.223			

Refinement	Zinc	Cadmium	"Native"
Space group	$P4_12_12$	C2	$P2_1$
Cell dimensions (Å)	$a = b = 58.4$, $c = 152.1$	$a = 107.1$, $b = 48.3$, $c = 42.65$, $\beta = 93.01°$	$a = 40.35$, $b = 64.6$, $c = 41.5$, $\beta = 117.83°$
Resolution (Å)	2.4	2.1	1.8
R	0.269	0.169	0.180
R_{Free}	0.350	0.208	0.252
Nonhydrogen atoms	1514	1521	1537
Metal ions	4	4	1
Waters	28	174	101
$\langle B\rangle$ (Å^2)	67.19	20.69	27.51
r.m.s. bonds	0.018Å	0.015	0.018
r.m.s. angles	2.278°	1.325°	1.467
r.m.s. planes	0.009Å	0.006Å	0.006Å
Ramachandran most	81.1%	88.6%	88.6%
Ramachandran additional	17.7%	10.8%	10.2%
Ramachandran generous	1.2%	0.6%	1.2%

[a] Values in parentheses refer to highest resolution range.

phase the higher resolution data from crystals grown in the presence of 20 mM cadmium chloride (cadmium data). The program ARP/wARP (23) was used to automatically rebuild the cadmium model at 2.1 Å resolution. The refined cadmium model served to phase the native data (1.8 Å, collected on station 9.5 at the SRS, Daresbury, UK) and finally to refine the zinc structure. The refinement statistics from REFMAC (24) are given in Table I.

X-ray Fluorescence Measurement—The x-ray fluorescence spectrum was recorded on the D15 station at LURE (Orsay, France). A drop of about 3 μl of the protein solution as used for crystallizations, containing 10 mg/ml YodA in 100 mM MES[1] buffer, pH 6, was deposited on a thin film of polypropylene and allowed to dry in air. The spectra were analyzed as described in Ref. 25.

E. coli Strains, Growth Media, and Culture Conditions—YodA was obtained in an Applikon fermentor system in FedBatch cultures (1500 ml) of the *E. coli* strain BL21λ DE3-Gold (Stratagene), transformed or not with pET28b, pET28b-CyaC, or pET28b-CyaC-GroES/EL expression vectors. The cultures were performed in a synthetic medium containing 10 g/liter glycerol, 27.2 g/liter KH_2PO_4, 4 g/liter $(NH_4)_2SO_4$, 8 g/liter KOH, 4 mg/liter thiamin, 0.4 g/liter $MgSO_4$ ($7H_2O$), 100 mg/liter $CaCl_2$ ($2H_2O$), 1 mg/liter $FeSO_4$ ($7H_2O$), 0.4 mg/liter $ZnSO_4$ ($7H_2O$), 0.4 mg/liter $MnSO_4$, 0.04 mg/liter $CuSO_4$ ($5H_2O$), 0.2 mg/liter Na_2MoO_4, 0.5 mg/liter H_3BO_3, and 0.1 mg/liter KI (19). For the FedBatch cultures, 50 times concentrated medium was added to obtain a specific growth rate of 0.25/h and a biomass yield of 0.5g/g of glycerol 15 h later. The cultures were carried out at 37 °C at pH 7.0 with a stirring rate of 800 rpm, 1 volume of air/1 volume of culture/min. The stirring rate was finally increased to over 1200 rpm. All of the cultures were subjected to a temperature increase to 42 °C for 20 min prior to induction, to overexpress endogenous chaperone proteins. After removing a first cell sample (noninduced), the cells were induced with 0.1 mM (in Batch cultures) or 1 mM isopropyl-β-D-thiogalactopyranoside (in FedBatch culture) during the exponential phase of bacterial growth (when the biomass attained 1 g/liter in Batch culture and 10 g/liter in FedBatch culture). The bacteria were harvested 2 h later (induced sample).

Western Immunoblot Assay—Polyclonal antiserum against the purified YodA protein was prepared by injecting a rabbit subcutaneously three times at 3-week intervals. After collecting the blood, the antibody was purified from the serum on a Hi-Trap NHS-activated column (Amersham Biosciences) coupled with YodA according to the manufacturer's protocol. The specificity of the antibody was checked on a purified sample of YodA by a Western blot (26). All of the Western blots were stained using the nitrotoluedine blue and 5-bromo-4-chloro-3-indolyl phosphate system (Bio-Rad).

RESULTS

YodA Structure—The three-dimensional structure of YodA consists of two domains: an antiparallel, up-down β-barrel flanked by one α-helix (the "calyx" domain), a compact domain about 20 Å × 23 Å × 26 Å in size, and a helical domain that opens out at the side of the calyx β-barrel (the "helix" domain). The latter domain is very open, with an overall size of 30 Å × 20 Å × 23 Å, with a deep cleft running down through it (Fig. 1). A visual inspection of the β-barrel domain suggests a relationship with the lipocalin-calycin family of proteins (17). This impression is confirmed by the results of a DALI search of existing protein structural data (27), which returns several lipocalin domains, including the epididymal retinoic acid-binding protein (EPA; Protein Data Bank identification code 1EPA) (28), in the list of structurally related protein domains. We used the program TOPP from the CCP4 package (29), which allows the superposition of molecules using their secondary structure

[1] The abbreviations used are: MES, 4-morpholineethanesulfonic acid; ABC, ATP-binding cassette; EPA, epididymal retinoic binding protein.

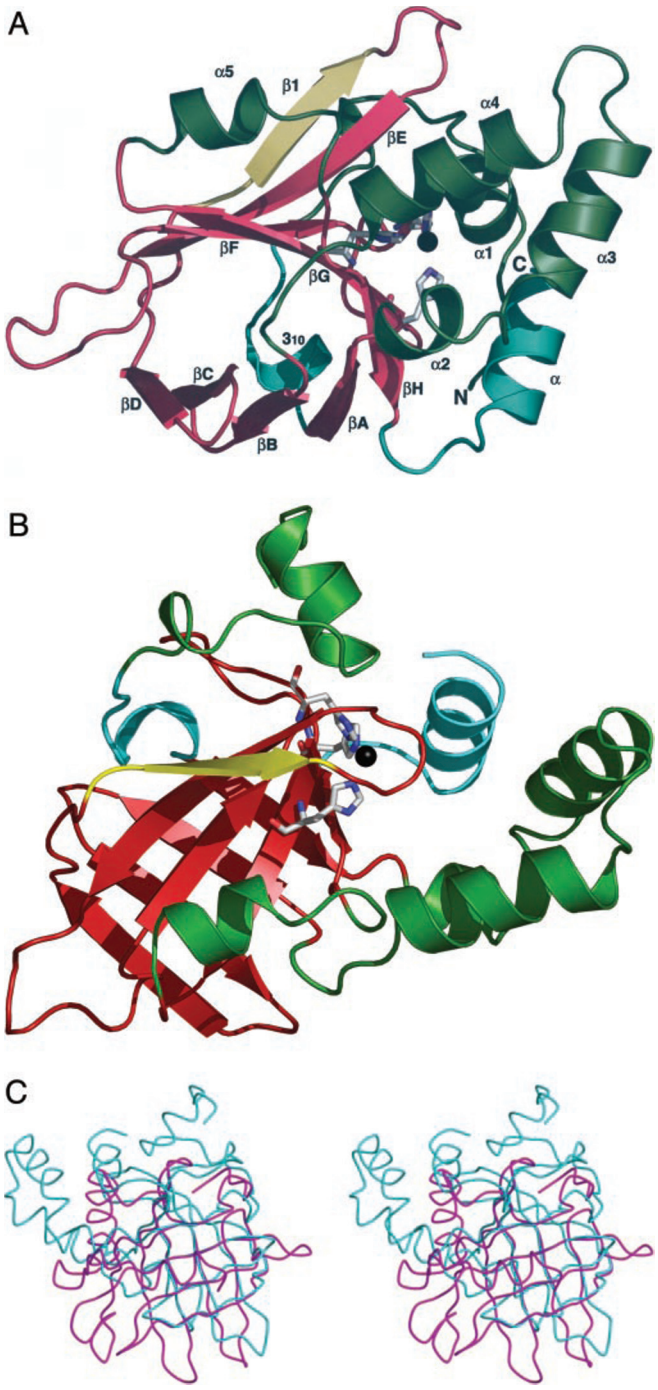

FIG. 1. *Top* and *middle panels*, two perpendicular, schematic views of the YodA molecule. The calyx domain is in *brick red* (β-barrel) and *cyan* (helices), whereas the insertions are colored in *green* (helices) and *yellow* (β-strand). The secondary structure elements are labeled in the *top panel* using the same convention as in Fig. 2 for clarity. The three histidines and the metal ion in the central metal-binding site are *highlighted*. *Bottom panel*, stereo view of a superposition of EPA (*magenta*) on YodA (*light blue*) backbones.

elements regardless of sequence similarity, to carry out a careful alignment of EPA on YodA. The procedure matched all eight β-strands of the calyx domain, corresponding to 60 residues, with an root mean square distance on Cα atoms of 1.53 Å. This results in a sequence alignment with only seven identical residues, perhaps one of the lowest sequence similarities within the lipocalin family; yet all of the elements of the calycin structure are matched (Fig. 2).

The relationship with lipocalins is most conserved for the N-terminal half of the structure and especially within the first signature motif (structurally conserved region 1) (17), where the strict requirement of a conserved Gly-Xaa-Trp sequence at the start of the first β-strand is fully obeyed. Even the Trp^{39} side chain superimposes onto the corresponding EPA Trp^{15} extremely well. All of the β-strands on the N-terminal side of the β-barrel superimpose rather well, although the YodA strands tend to be shorter. The superposition is somewhat less good for the second half of the β-barrel, where in the case of strands E and H, it is valid for only four residues. Conversely, the curvature of the β-strands matches more closely in the C-terminal half of the calyx domain than in the N-terminal half. Note that the conserved Arg of the lipocalin structurally conserved region 3 motif is replaced by a Pro in YodA, leaving the indole ring of Trp^{39} exposed to solvent. Unlike Trp^{15} in EPA, however, the Trp^{39} side chain in YodA packs against several aromatic residues filling the interior of the barrel (Fig. 3).

An ordered water molecule (B = 21 $Å^2$ in the native form, B = 7.9 $Å^2$ in the cadmium form) lies at the top of the calyx domain, between strands A and H and the base of helix 4, securing the α-helical loop against the calyx. It forms a hydrogen bond with two backbone oxygens, one backbone nitrogen and Oγ of Ser^{41}. This water molecule is the most buried within the interior of YodA. The two crystal forms, cadmium and "native," superimpose nearly exactly in this region of the structure.

YodA is larger than some canonical lipocalins such as EPA (193 residues for YodA compared with 162 for EPA), which is due to several sequence insertions. These occur, furthermore, between the conserved structural elements of the calyx domain: one extra β-strand extending strand D and several helices that constitute the helical domain. Helix 1 (α-helix) and an extra 3_{10} helix form an extension at the N terminus that lies opposite α-helices 2, 3, and 4, all three forming an insertion in the sequence between β-strands A and B (Figs. 1 and 2). There is another insertion between strands G and H that forms a short α-helix at the side of the barrel (helix 5).

Although most lipocalins described to date bind a hydrophobic molecule in the barrel interior, this does not seem to be the case for YodA. It is therefore interesting to compare YodA with a lipocalin with a different function, such as triabin (Protein Data Bank code 1AVG) (30), an inhibitor of thrombin that acts *via* its surface to block the active site of the protease. The calyx barrel of triabin is very hydrophobic, with the exception of one stabilizing salt bridge, and does not contain a cavity that might accommodate other molecules. In a similar manner, the interior of the YodA calyx barrel is entirely hydrophobic with several densely packed aromatic residues. Unlike triabin, however, there are no internal polar residues within the barrel. The entrance to the wider side of the calyx is closed by a loop between helix 3 and β-strand B and a loop between β-strands E and F, restricting the opening. This loop is held in place by a disulfide bridge between strands D and E. This is very similar to triabin, and the respective cysteines lie indeed quite close in the superposition (Fig. 3). The opening is further obstructed in YodA by the side chain of Arg^{96} that lies across it. The narrow end of the calyx barrel is more narrow in YodA than in EPA or even triabin, with the loop between strands F and G closing over it like a lid. Interestingly, the helix 5 insertion of YodA overlaps with the N terminus of triabin in the optimized superposition.

Metal-binding Sites—We have solved the structure of YodA in three crystal forms, each corresponding to different crystallization conditions. Two forms were obtained in the presence of added divalent cations, Zn^{2+} and Cd^{2+}. The latter choice was

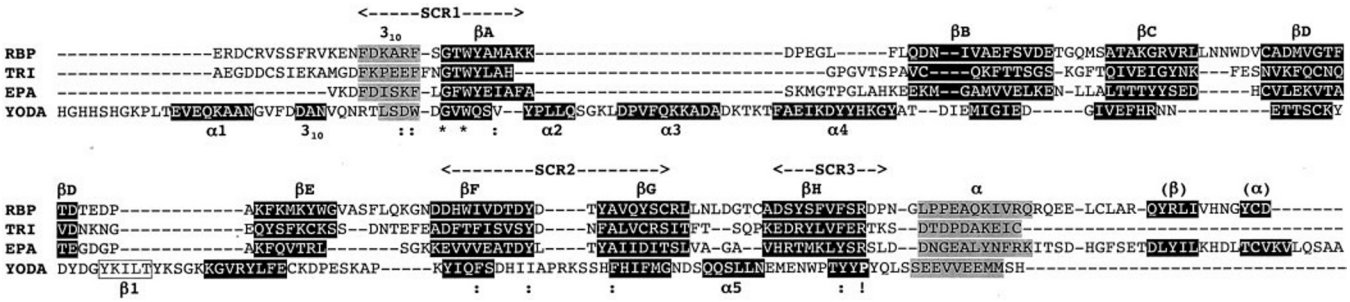

FIG. 2. **Sequence alignment of YodA with the following lipocalins: retinol-binding protein (*RBP*) (56), triabin (*TRI*) (30), and epididymal retinoic acid-binding protein (*EPA*) (28).** Sequence alignment of YodA with the following lipocalins: retinol-binding protein (*RBP*) (56), triabin (*TRI*) (30), and epididymal retinoic acid-binding protein (*EPA*) (28). The lipocalin secondary structural elements are depicted as follows: β-strands as *white characters* on a *black background*, conserved α-helices as *black* on *gray*. The insertions in YodA are depicted in *dark gray* on *light gray* (β-strand) and *white* on *gray* (helices). The structurally conserved regions (*SCR1–3*) and the numbering of secondary structures as defined for lipocalins (17) are indicated *above* the sequences, conserved residues, and YodA-specific secondary structures are indicated *below* the sequences. The replacement of the conserved Arg by a Pro in YodA is highlighted by an *exclamation point*.

influenced by our knowledge of the cadmium-induced expression of the protein.

The overall structure of the three forms is practically identical (root mean square differences on Cα atoms are 0.286 Å between the native and cadmium structures and 0.594 Å between the native and zinc structures). In all of them, the N- and C-terminal helical extensions are less well ordered than the calyx domain. The N terminus is most ordered in the native form, where the electron density was interpretable from residue 7 (Gly), and least well defined in the zinc form. The latter crystal form gave highly anisotropic data with the lowest resolution, yet in this crystal form the C terminus is relatively well defined up to the last residue (His^{193}), which is poorly ordered in the other forms.

The zinc and cadmium crystal forms contain several metal ions bound to the protein. All of the sites were confirmed using the anomalous signal from data collected at appropriate wavelengths (the multiple anomalous dispersion data for the zinc form, a low resolution data-set collected at 1.7712 Å for the cadmium form) (19). One central site, common to all forms, lies along the side of the calyx domain, enclosed by the helix domain. In the case of cadmium, the metal is coordinated by three histidine side chains (His^{144}, His^{153}, and His^{155}, from strands F and G of the calyx domain) and three water molecules in a typical octahedral geometry (Fig. 4*A*). The same site accommodates two zinc ions; one is coordinated by His^{144} and His^{155}, whereas the other interacts with His^{153}, His^{193}, and the caboxyl moiety of Glu^{189} (Fig. 4*B*). Several water molecules lie in close contact with the zinc ions, but we cannot give a precise description of the coordination geometry, because the metal ions are disordered. His^{153} moves slightly to accommodate the second zinc ion, whereas the C-terminal His^{193} becomes more ordered by binding to the metal ion.

The remaining metal ions are bound on the protein surface at intermolecular sites formed by residues from two neighboring molecules. Because the packing in the two crystal forms is totally different, these sites are formed differently. Nevertheless, some of these sites have common features. One zinc ion is coordinated by His^{95} of one molecule and the carboxyl group of Glu^{93} of a symmetry-related molecule. In the cadmium form, we find one cadmium ion coordinated by His^{95} of one molecule and the carboxyl of Glu^{185} of a symmetry-related molecule, and another one is coordinated by carboxyls of Glu^{93} from one molecule and Glu^{188} of a symmetry-related molecule. The remaining zinc ion mediates contacts between the helix domains of symmetry-related molecules (carboxyl group of Asp^{24} of one and His^{75} and an oxygen of Gln^{47} amide group of the other molecule). The last cadmium ion also lies between symmetry-related molecules mediating a contact between the helices

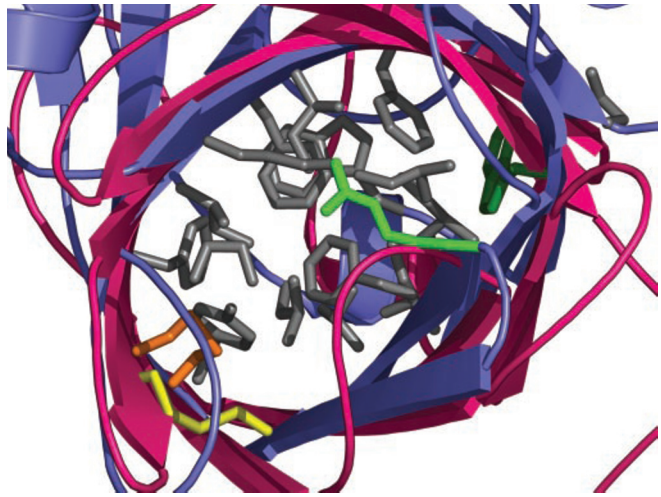

FIG. 3. **Superposition of the calyx domains of triabin and YodA viewed down the wider end of the barrel.** The backbone of triabin is in *magenta*, YodA is in *slate blue*, and the hydrophobic side chains within YodA interior are in *gray*. The only disulfide in YodA is in *yellow*, the corresponding one in triabin (Cys^{69} and Cys^{84}) is in *orange*. Trp^{39} and Arg^{69}, lying at opposite entrances to the calyx barrel, are highlighted in *dark* and *light green*, respectively.

flanking their calyx domains (carboxyl groups of Glu^{169} of one and Gly^{171} of the other molecule). These intermolecular sites all correspond to negatively charged regions on the protein molecule (Fig. 5).

To our surprise, the native crystal form revealed the presence of one metal ion bound in the central, common, metal-binding site. The ion position corresponds to that of cadmium and it is coordinated by the same three histidines (His^{144}, His^{153}, and His^{155}). In this case, however, only one water molecule completes the coordination, forming a somewhat distorted tetrahedron (Fig. 4*C*). The presence of a metal cation indicates that this site is a high affinity metal-binding site, because no significant amount of divalent cations was added during the purification and crystallization of the protein. The binding geometry would suggest zinc as a prime candidate, although iron and copper are a possibility (all of these elements were present in the bacterial growth medium) (19). Another ion that was present during protein purification is nickel, because we used a nickel-nitrilotriacetic acid affinity column.

The crystal data were collected at a wavelength of 1.3 Å, and we found a significant anomalous peak (5.5 σ) centered on the metal site. The wavelength used corresponds to the low energy side of the zinc absorption threshold but to the high energy side for nickel, copper, and iron. To confirm the identity of the ion,

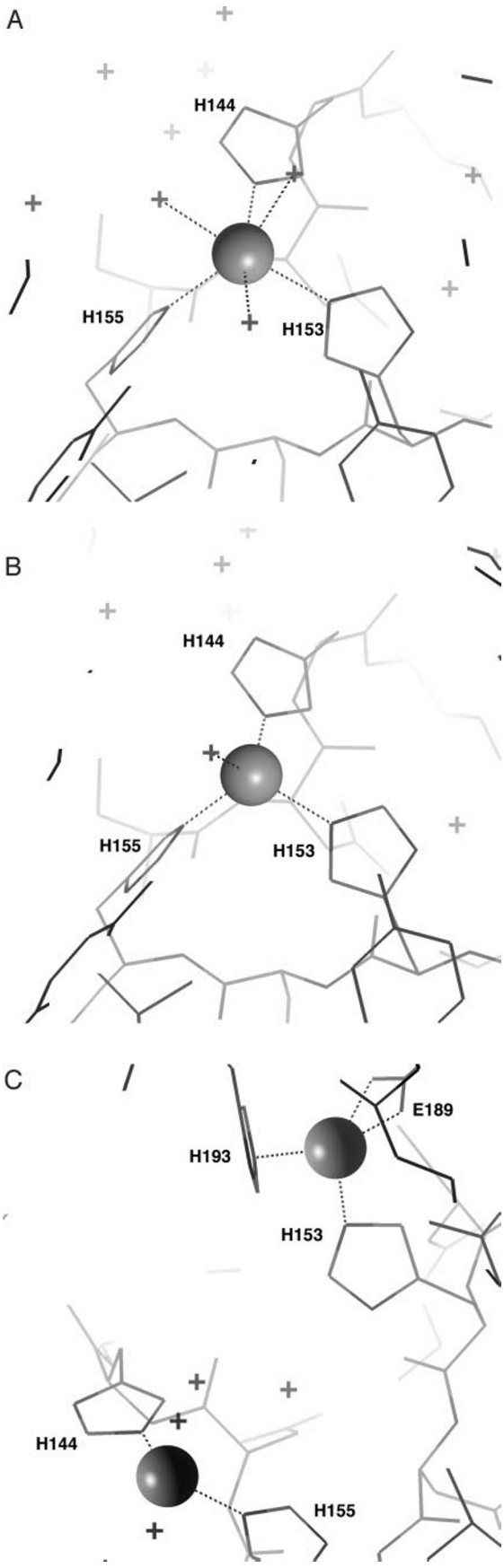

FIG. 4. **Detail of the principal metal-binding site in the cadmium (*A*), the native (*B*), and the zinc (*C*) crystal forms.** *A* and *B* are in the same orientation, and in *C* the coordinates were rotated by ~180° about the vertical direction for clarity. This figure was made with Molscript (57) and Raster-3D (58).

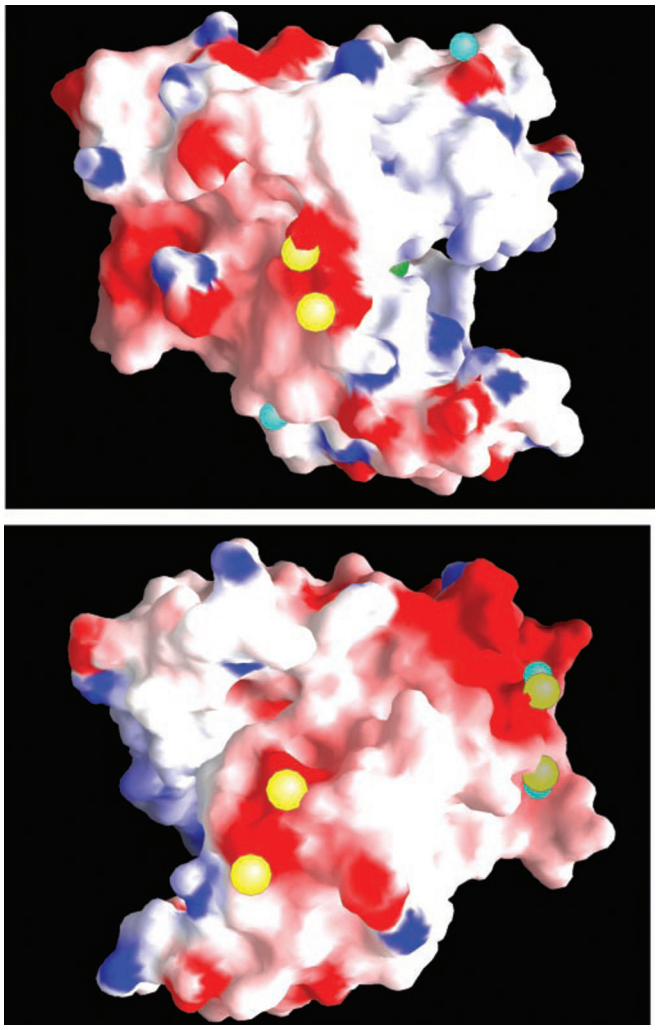

FIG. 5. **The electrostatic surface of YodA with bound metals represented as *green* (nickel), *cyan* (zinc), and *yellow* (cadmium) balls.** The intermolecular metal sites depicted for the cadmium and zinc forms correspond to all ions bound to the protein molecule, including those shared by symmetry-related molecules. This figure was prepared using GRASP (59).

we performed an x-ray fluorescence analysis of the protein solution, as used for crystallization. The major component was nickel, with a negligible signal for copper and iron, whereas zinc give a signal only 0.25 of that of nickel. We therefore conclude that the metal bound in this case is nickel, in an unusual, distorted tetrahedral geometry. It has to be noted that the ligand geometry for nickel in proteins has not been studied very carefully to date. Upon examination of the available data (*via* the Metalloprotein Data Base) (31), we found two other cases of distorted tetrahedral geometry of nickel ion ligands from high resolution protein crystal structure determinations: azurin (1NZR) (32) and a complex of trypsin with ecotin (1SLW) (33).

YodA Expression Results—YodA had been identified in the past in *E. coli* cells under two different sets of conditions. Ferianc *et al.* (1) found an increase of YodA expression when submitting *E. coli* to increased levels of cadmium in the medium, whereas Laurent-Winter *et al.* (34) observed a decrease of YodA expression in *E. coli* mutants defective in the histone-like nucleoid-structuring protein. The latter study would suggest a basal level of YodA expression in *E. coli* cells. Because neither case corresponds to our experimental conditions, we wished to ascertain, if possible, the factor(s) responsible for the overproduction of YodA in our hands.

We initially obtained YodA when in fact attempting to overexpress a heterologous protein in *E. coli* (CyaC from *Bordetella pertussis* (19)). The recombinant protein was co-expressed with the molecular chaperones GroEL-GroES to assist its correct folding. Expression of molecular chaperones was consecutive to a transient temperature shift to 42 °C for 20 min. To obtain sufficient quantities of recombinant protein, the production was carried out in fermentors using a high cell density culture mode in a FedBatch process, where the addition of nutrients and growth ($\mu = 0.25\ h^{-1}$) are automatically controlled.

Several authors have already described that the overexpression of a recombinant gene induces a heat shock-like response that enhances both proteolytic activities and the levels of chaperones (35, 36). Because YodA had been related to cellular stress, we first thought that the overexpression of a heterologous protein could trigger a cellular stress provoking YodA expression in *E. coli*.

Our first results (Fig. 6) had shown that in FedBatch mode YodA is mainly expressed in the soluble cellular fraction, whether or not the plasmid carrying the heterologous protein is present. The presence of YodA seems to be independent of induction; the protein is already expressed before the addition of isopropyl-β-D-thiogalactopyranoside to the medium. This observation also suggests a basal level of YodA expression in *E. coli* cells.

The temperature shift to 42 °C, carried out just before the protein expression analysis in noninduced conditions, could have been a sufficient factor to trigger a "stress" response including the induction of the YodA expression. But surprisingly, the same heat shock used in the Batch process did not induce any YodA production.

The comparison of two cell culture strategies, Batch mode or FedBatch mode, reveals that YodA appears to be expressed only when using a high cell density culture strategy (final biomass concentration is about 25 g/liter; Fig. 6). Thus if YodA is indeed a stress response protein, then the conditions used in automatic fermentor systems might represent a mild stress for *E. coli*.

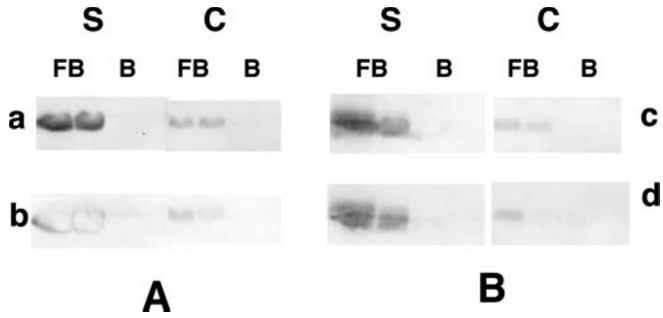

FIG. 6. **Western blot analysis of YodA expression in FedBatch (*FB*) or Batch (*B*) culture for different *E. coli* strains.** The analysis was performed prior to the inducer incorporation into the medium and after the temperature shift (*A*), and 2 h after induction (*B*), for the soluble cellular fractions (*S*) and for the cell debris (*C*). *Panels a* and *c*, BL21λ DE3-Gold strain as control in *lane 1*, the same strain containing the pET28b plasmid vector in *lane 2*. *Panels b* and *d*, *E. coli* BL21λ DE3-Gold expressing CyaC, whose gene had been cloned in the pET28b expression vector with an additional insertion of the two molecular chaperones GroEL/ES genes upstream of the CyaC gene in *lane 1* and the same construct with the CyaC gene only in *lane 2*.

DISCUSSION

The close resemblance of YodA to the calycin family excludes any structural relationship with the copper-binding amine oxidase, as had been suggested in the past (1), because the latter is composed of β-sheet domains with a totally different topology (37). Members of the lipocalin-calycin family of proteins have been found in all organisms from bacteria to man (18), where they play very diverse roles (where these have been identified). Most are secreted proteins, and a large number seem to be involved in the transport of hydrophobic molecules that can be sequestered within the β-barrel. Others, such as a number of bacterial lipocalins, but also apolipoprotein D, contain extra hydrophobic sequences either as loops or at the N terminus, which allow them to interact with membranes. Still others, such as avidin and protease inhibitors (*e.g.* triabin), interact with other proteins, forming tight complexes.

In general, the open end of the calycin β-barrel corresponds to the entry of ligands in the case of calycins involved in the transport of hydrophobic substrates. In the case of YodA, the entrance to the internal cavity is blocked at this end. Furthermore, the cavity itself is completely filled with hydrophobic residues, leaving no space for other molecules. We can therefore conclude that YodA was not designed to carry substrates within its calycin barrel.

Although it is quite likely that YodA was purified on a nickel-nitrilotriacetic acid column with the help of the N-terminal His-rich sequence, this region is disordered in all of the crystal forms we studied and clearly does not bind any metals. The principal, high affinity metal site in YodA is the much more buried site described above. All of the His residues implicated in this site, as well as the C-terminal His that participates in ion binding in our zinc form, are absolutely conserved in the three related proteins, as well as in the C-terminal part of *S. pneumoniae* AdcA. We propose that this is the functional metal-binding site in the entire family of these proteins. It has to be noted, however, that we cannot conclude on the nature of the metal that would bind under physiological conditions, within the *E. coli* cell, from our study. Nevertheless YodA, and by inference the closely related proteins YrpE, pXO1–130, and the C-terminal domain of AdcA, is the first naturally occurring metal-binding lipocalin described to date.

The three-histidine motif has been found in several metalloproteases, lyases, and oxygenases. In some bi-nuclear metalloenzymes, such as superoxide dismutase, we find a fourth histidine participating in the coordination of the two neighboring ions. We searched several structural data bases, superimposing the metal-binding sites, to investigate the possibility of an enzymatic activity for YodA (*e.g.* the PINTS server) (38). Apart from the immediate metal ligands, there are no further similarities discernable with any type of structurally described enzyme active site. Further studies need to be carried out to verify whether YodA does possess any enzymatic activity.

A three-histidine motif is fairly common in Zn(II)-binding proteins and has been used to design artificial metal-binding proteins. Because the lipocalin structural motif is a rather stable and well defined structure, it had been used as a scaffold for grafting a zinc-binding motif, thus creating a bifunctional lipocalin (39). The authors had chosen two sites on the side of the retinol-binding protein calyx, making use of two neighboring β-strands for each one. Nature has indeed employed the three-histidine motif in YodA, positioning it at the side of the calyx as in the engineered protein but on the opposite side of the β-sheet calyx.

Up to now, YodA induction has been observed under particular expression conditions, such as cadmium stress (12). Our results showing the induction of YodA under the conditions of fermentor culture of bacteria do therefore raise some questions. Cell physiology of microorganisms growing in excess of all required nutrients (Batch phase) differ considerably from cells that are exposed, in high cell density cultures, to transient substrate and energy limitation, to metabolite (acetic acid) and nutrient accumulation, or to changes in oxygen availability. Growing cells to high densities aerobically includes the use of

severely energy-limited processes, the physiological consequences and the cellular stress response of which are not yet well described.

We can suggest several hypotheses for YodA induction. First, in our cell culture conditions, YodA expression could be induced because of a possible accumulation of residual traces of Zn^{2+}. However, this nutrient could be accumulated in the culture medium at a maximum concentration of 0.07 mM, whereas earlier studies had shown (12) that a 10 times higher concentration (0.7 mM) was not sufficient to trigger YodA expression.

Second, the conditions of oxygenation used in a fermentor could cause oxidative stress. Dissolved O_2 is maintained above 20% oxygen saturation by an aeration rate of 1 volume of air/1 volume of culture/min and a stirring rate of over 1200 rpm.

Third, the use of a nutrient-limited growth strategy (Fed-Batch culture) could cause a cellular response. Andersson *et al.* (40) have already reported the relation between cellular physiology in FedBatch cultures and the stringent response. The transition between the Batch and the FedBatch mode immediately increases the content of the stringent response signal substance, the alarmone (p)ppGpp, to stabilize at a higher value than in a Batch phase. These observations have been reproduced with plasmid free and noninduced plasmid bearing cultures. Furthermore, the level of the transient response (σ^S factor and (p)ppGpp) is related to the culture conditions applied (a rapid or progressive change) and to the specific rate of growth imposed (41). Finally, De Lisa *et al.* (42) have reported higher levels of transcript chaperones, proteases, and lysis genes associated with the high cell density mode.

More generally, cellular stress response such as heat shock, oxidative damage, heavy metal accumulation, nutrient starvation, severe energy/carbon limitation, low growth, acidification, and recombinant protein expression are correlated with the stringent response. This results in a transcriptional shutdown of protein synthesis and in the induction of universal stress proteins (*usp*) and in particular the iron response regulator (*fur*), implicated in the protection against oxidative damage (43). We know that cadmium induces the (p)ppGpp response (44) as well as the expression of YodA. The alarmone (p)ppGpp is a prerequisite, even if it is not sufficient for its expression (12). The transcription of YodA is also dependent on *fur* and on *soxS*, which mediates the protection against the reactive oxygen species.

Recently, genomics studies (45) have also detected an increase of YodA level in *E. coli* consecutive to an acid-induced change in the culture medium (pH 7 to 5.8). This induction is correlated with the induction of genes implicated in the amino acid metabolism and of the EF-Ts factor, normally induced by the stringent response.

We now show that the FedBatch cell culture strategy is also responsible for YodA induction. Nevertheless, further physiological studies need to be carried out to understand the exact functional role of YodA in bacteria grown in nutrient limited conditions.

Members of the lipocalin family have recently been identified in bacteria (46, 47). The authors have shown that in the case of the *E. coli*, the lipocalin Blc is expressed optimally under conditions of high osmolarity, and the expression is governed by a subclass of the σ^S-dependent promoter. It was therefore suggested that Blc contributes to the adaptation of cells to certain types of stress. Interestingly, several lipocalins have been identified in plants, where they participate in the activation of pigments involved in the light-dark cycle (48). The expression of these lipocalins increases in response to excess light absorption, thus protecting plants from photo-induced oxidative damage. Although YodA does differ from *E. coli* Blc in its N-terminal sequence and in its cellular localization (Blc is an outer membrane lipoprotein), it could be a novel member of a family of bacterial stress response lipocalins.

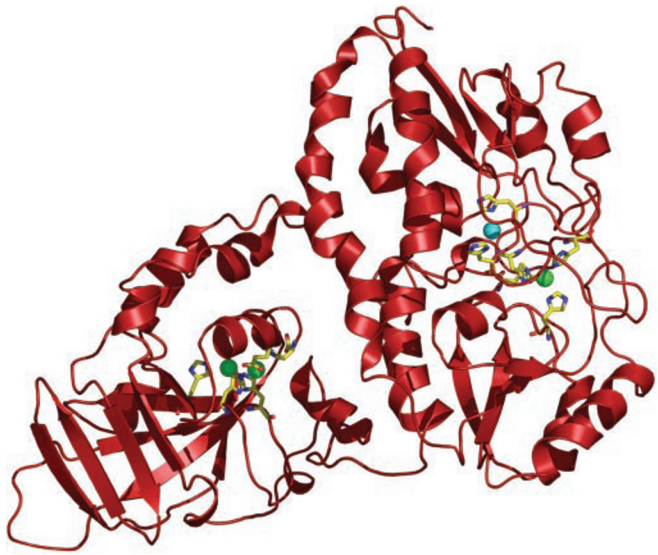

FIG. 7. **A model of AdcA obtained using PsaA (Protein Data Bank code 1PSZ) and YodA coordinates as templates.** The metal-binding sites are depicted in ball-and-stick representation. The *green spheres* represent ions observed in the two template structures, and the *light blue sphere* is a possible further ion binding to the His-rich loop. Figs. 1, 3, and 6 were made using PyMOL (60).

YodA, YrpE, and pXO1–130 show close sequence similarity to the C-terminal half of the AdcA protein from *S. pneumoniae*. The latter is the periplasmic metal-binding receptor of an ATP-binding cassette (ABC) metal-transport system, or permease, that has been shown to bind zinc (49). In fact, within a recently defined new family of ABC manganese and zinc permeases, three members possess a metal-binding receptor that has a C-terminal extension homologous to YodA: ZnuA from *Enterococcus faecalis*, AdcA from *S. pneumoniae*, and Spy9 from *Streptococcus pyogenes* (50). The sequence of the N-terminal domain of AdcA from *S. pneumoniae* is highly homologous to that of PsaA from the same species, whose x-ray crystal structure has been determined recently (51). Together with the YodA structure we have determined, the entire sequence of AdcA can be matched to known structural templates.

We have used the program Modeler (52) to model the AdcA structure using the PsaA (Protein Data Bank code 1PSZ) and YodA coordinates as templates for the N-terminal and C-terminal domains respectively. Fig. 7 gives a ribbon representation of the optimal model. The complete AdcA molecule contains two distinct metal-binding sites, one in each domain. Both metal-binding sites contain three His residues, with the N-terminal site also containing a Glu side chain as a ligand. AdcA has an insertion containing a His-rich sequence close to the N-terminal metal-binding site. In our model these residues lie in a loop that forms a cleft leading toward the metal-binding site, and they could easily bind further metal ions. The C-terminal metal-binding site could also accomodate a further ion, as we see in the YodA zinc structure.

The metal specificity of these sites cannot be defined from the ligand geometry alone; indeed, YodA structure shows that the same site can accomodate different divalent ions quite easily. The presence of a carboxyl group in the N-terminal site of AdcA might indicate a higher affinity binding in this site. As for the other sites, have they been designed to bind and release metal ions fairly readily, depending on their relative concentration in the periplasmic space of the bacterium? One can imagine that although the principle metal-binding site in the

C-terminal domain is structural, the binding of a second ion in that domain could trigger a conformational change that could be transmitted to the N-terminal domain, thus affecting transport across the membrane.

Bacterial ABC transporters are generally complexes of several protein chains: one or two transmembrane protein subunits, one or two peripheral ATP-binding proteins on the cytoplasmic side of the membrane, and a high affinity solute binding protein (53). In Gram-negative bacteria this protein is a soluble periplasmic protein, whereas in Gram-positive bacteria it is bound to the external side of the membrane *via* a N-terminal sequence. The binding of a solute molecule triggers a conformational change in the solute-binding protein, which increases its affinity for the transmembrane protein. This in turn will activate the transport of the solute across the membrane (53). Several classes of these solute-binding components of bacterial ABC transporters have been identified, each class usually specific for a particular solute. Thus the AdcA protein from *S. pneumoniae* has been defined as a member of a zinc and manganese transport class of permeases (50). The majority of metal-binding proteins in this family contain a single metal site within a mainly α-helical protein of around 300 residues. Only a small group, mentioned above, seem to have arisen by gene fusion with a YodA-like protein to give a much larger two-domain protein with two metal-binding sites. We therefore propose that YodA represents the periplasmic partner of an as yet unknown ABC transporter in *E. coli*, and because its expression is induced under conditions of cadmium or oxidative stress, its role could be to enhance or to block the transport activity of this putative permease under these stress conditions.

Acknowledgments—We thank Dr. Françoise Russo-Marie for permitting us to prepare anti-YodA antibodies, Dr. Stéphane Réty and Dr. Peter Ferianc for very helpful discussions, and Thomas Millerand for the Western blots. We gratefully acknowledge the help of staff at the synchrotron facilities used and in particular Dr. E. Duke (SRS, Daresbury, UK) and Dr. R. Kahn and Dr. J.-L. Ferrer (ESRF, Grenoble, France). P. Zwart (EMBL, Hamburg, Germany) and A. Stark (EMBL, Heidelberg, Germany) very kindly helped with the use of latest versions of ARP/wARP and the PINTS server, respectively, and James D. Watson (EBI, Cambridge, UK) performed enzyme active site data base searches. Dr. Pierre Chevalier at LURE kindly performed the x-ray fluorescence experiment.

REFERENCES

1. Ferianc, P., Farewell, A., and Nystrom, T. (1998) *Microbiology* **144,** 1045–1050
2. Puskarova, A., Janecek, S., Ferianc, P., and Polek, B. (2001) *Biologia* **56,** 337–339
3. Huang, K. F., Chiou, S. H., Ko, J. M., and Wang, A. H. J. (2002) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **58,** 1118–1128
4. Petering, D. H., Huang, M., Moteki, S., and Shaw, C. F. R. (2000) *Mar. Environ. Res.* **50,** 89–92
5. Vanysek, P. (1997) in *CRC Handbook of Chemistry and Physics* (Lide, D. R., ed) Vol. 77, pp. 8, 20–28, and 30, CRC Press, New York
6. Beard, S. J., Hughes, M. N., and Poole, R. K. (1995) *FEMS Microbiol. Lett.* **131,** 205–210
7. Babai, R., and Ron, E. Z. (1998) *FEMS Microbiol. Lett.* **167,** 107–111
8. Inbar, O., and Ron, E. Z. (1993) *FEMS Microbiol. Lett.* **113,** 197–200
9. Gomes, D. S., Fragoso, L. C., Riger, C. J., Panek, A. D., and Eleutherio, E. C. A. (2002) *Biochim. Biophys. Acta* **1573,** 21–25
10. Vido, K., Spector, D., Lagniel, G., Lopez, S., Toledano, M. B., and Labarre, J. (2001) *J. Biol. Chem.* **276,** 8469–8474
11. Ferianc, P., Puskarova, A., Godocikova, J., Polek, B., and Toth, D. (2000) *Biologia* **55,** 653–659
12. Puskarova, A., Ferianc, P., Kormanec, J., Homerova, D., Farewell, A., and Nystrom, T. (2002) *Microbiolog y* **148,** 3801–3811
13. Rosen, B. P. (1999) *Essays Biochem.* **34,** 1–15
14. Binet, M. R., and Poole, R. K. (2000) *FEBS Lett.* **473,** 67–70
15. Vasak, M., and Hasler, D. W. (2000) *Curr. Opin. Chem. Biol.* **4,** 177–183
16. Blindauer, C. A., Harrison, M. D., Robinson, A. K., Parkinson, J. A., Bowness, P. W., Sadler, P. J., and Robinson, N. J. (2002) *Mol. Microbiol.* **45,** 1421–1432
17. Flower, D. R., North, A. C. T., and Sansom, C. E. (2000) *Biochim. Biophys. Acta* **1482,** 9–24
18. Åkerstrom, B., Flower, D. R., and Salier, J. P. (2000) *Biochim. Biophys. Acta* **1482,** 1–8
19. David, G., Blondeau, K., Renouard, M., and Lewit-Bentley, A. (2002) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **58,** 1243–1245
20. De la Fortelle, E., and Bricogne, G. (1997) *Methods Enzymol.* **276,** 472–494
21. Collaborative Computational Project, Number 4 (1994) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **50,** 760–764
22. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. Sect. A* **47,** 110–119
23. Lamzin, V. S., and Wilson, K. S. (1997) *Methods Enzymol.* **277,** 269–305
24. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **53,** 240–255
25. Brissaud, I., Wang, J. X., and Chevallier, P. (1989) *J. Radioanal. Nucl. Chem.* **131,** 399–413
26. Bollag, D. M., Rozycki, M. D., and Edelstein, S. J. (1996) *Protein Methods*, 2nd Ed., pp. 195–227, Wiley-Liss, New York
27. Holm, L., and Sander, C. (1996) *Science* **273,** 595–602
28. Newcomer, M. E., Pappas, R. S., and Ong, D. E. (1993) *Proc. Natl. Acad. Sci. U. S. A.* **90,** 9223–9227
29. Lu, G. (1996) *Protein Data Bank Quarterly Newsletter* **78,** 10–11
30. Fuentes-Prior, P., Noeske-Jungblut, C., Donner, P., Schleuning, W. D., Huber, R., and Bode, W. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94,** 11845–11850
31. Castagnetto, J. M., Hennessy, S. W., Roberts, V. A., Getzoff, E. D., Tainer, J. A., and Pique, M. E. (2002) *Nucleic Acids Res.* **30,** 379–382
32. Nar, H., Messerschmidt, A., Huber, R., Van De Kamp, M., and Canters, G. W. (1991) *J. Mol. Biol.* **218,** 427
33. McGrath, M. E., Erpel, T., Bystroff, C., and Fletterick, R. J. (1994) *EMBO J.* **13,** 1502–1507
34. Laurent-Winter, C., Ngo, S., Danchin, A., and Bertin, P. (1997) *Eur. J. Biochem.* **244,** 767–773
35. Neubauer, P., and Winter, J. (2001) in *Recombinant Protein Production with Prokaryotic and Eukaryotic Cells* (Merten, O. W., Mattanovich, D., Lang, C., Larsson, G., Neubauer, P., Porro, D., Postma, P., Teixeira de Mattos, J., and Cole, J. A., eds) pp. 195–258, Kluwer Academic Publishers, The Netherlands
36. Dürrschmid, K., Marzban, G., Dürrschmid, E., Striedner, G., Clementschitsch, F., Cserjan-Puschmann, M., and Bayer, K. (2003) *Electrophoresis* **24,** 303–310
37. Parsons, M. R., Convery, M. A., Wilmot, C. M., Yadav, K. D. S., Blakely, V., Corner, A. S., Phillips, S. E. V., McPherson, M. J., and Knowles, P. F. (1005) *Structure* **3,** 1171–1184
38. Stark, A., Sunyaev, S., and Russell, R. B. (2003) *J. Mol. Biol.* **326,** 1307–1316
39. Skerra, A. (2000) *Biochim. Biophys. Acta* **1482,** 337–350
40. Andersson, L., Yang, S., Neubauer, P., and Enfors, S. O. (1996) *J. Biotechnol.* **46,** 255–263
41. Teich, A., Meyer, S., Lin, H. Y., Andersson, L., Enfors, S. O., and Neubauer, P. (1999) *Biotechnol. Progr.* **15,** 123–129
42. De Lisa, M. P., Gill, R. T., and Bentley, W. E. (2001) in *Recombinant Protein Production with Prokaryotic and Eukaryotic Cells* (Merten, O. W., Mattanovich, D., Lang, C., Larsson, G., Neubauer, P., Porro, D., Postma, P., Teixeira de Mattos, J., and Cole, J. A., eds) pp. 43–55, Kluwer Academic Publishers, The Netherlands
43. Nyström, T. (2003) *Mol. Microbiol.* **48,** 17–23
44. Vanbogelen, R. A., Kelley, P. M., and Neidhart, F. C. (1987) *J. Bacteriol.* **169,** 26–32
45. Birch, R. M., O'Byrne, C., Booth, I. R., and Cash, P. (2003) *Proteomics* **3,** 764–776
46. Bishop, R. E., Penfold, S. S., Frost, L. S., Höltje, J.-V., and Weiner, J. H. (1995) *J. Biol. Chem.* **270,** 23097–23103
47. Bishop, R. E. (2000) *Biochim. Biophys. Acta* **1482,** 73–83
48. Hieber, A. D., Bugos, R. C., and Yamamoto, H. Y. (2000) *Biochim. Biophys. Acta* **1482,** 84–91
49. Dintilhac, A., Alloing, G., Granadel, C., and Claverys, J. P. (1997) *Mol. Microbiol.* **25,** 727–739
50. Claverys, J. P. (2001) *Res. Microbiol.* **152,** 231–243
51. Lawrence, M. C., Pilling, P. A., Epa, V. C., Berry, A. M., Ogunniyi, A. D., and Paton, J. C. (1998) *Structure* **6,** 1553–1561
52. Sali, A., and Blundell, T. L. (1993) *J. Mol. Biol.* **234,** 779–815
53. Tam, R., and Saier, M. H. J. (1993) *Microbiol. Rev.* **57,** 320–346
54. Brünger, A. T. (1992) *Nature* **355,** 472–474
55. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr.* **26,** 283–291
56. Cowan, S. W., Newcomer, M. E., and Jones, T. A. (1990) *Proteins Struct. Funct. Genet.* **8,** 44–61
57. Kraulis, P. J. (1991) *J. Appl. Crystallogr.* **24,** 946–950
58. Merritt, E. A., and Murphy, M. E. P. (1994) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **50**, 869–873
59. Nicholls, A., Sharp, K., and Honig, B. (1991) *Proteins Struct. Funct. Genet.* **11,** 281–296
60. DeLano, W. L. (2002) *The PyMOL User's Manual*, DeLano Scientific, San Carlos, CA