Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Expression, purification and crystallization of a collagen-binding fragment of *Yersinia* adhesin YadA

Heli Nummelin,[a] Yasmin El Tahir,[b] Pauli Ollikka,[b]† Mikael Skurnik[b] and Adrian Goldman[a]*

[a]Structural Biology and Biophysics, Institute of Biotechnology, University of Helsinki, Viikinkaari 1, PO Box 65, FIN-00014 University of Helsinki, Finland, and [b]Department of Medical Biochemistry and Molecular Biology, Kiinanmyllynkatu 10, University of Turku, FIN-20520 Turku, Finland

† Current address: BioTie Therapies Corporation, BioCity, Tykistökatu 6, FIN-20520 Turku, Finland.

Correspondence e-mail: adrian.goldman@helsinki.fi

Received 8 January 2002
Accepted 21 March 2002

A recombinant form of a collagen-binding fragment of *Yersinia enterocolitica* serotype O:3 adhesin YadA with an N-terminal polyhistidine affinity tag has been produced in *Escherichia coli*, purified and crystallized using the sitting-drop vapour-diffusion technique. Crystals belong to the trigonal space group *R*3, with unit-cell parameters $a = b = 67.05$, $c = 221.95$ Å, and diffract to 1.55 Å resolution on a synchrotron-radiation source.

1. Introduction

Yersiniae are facultative anaerobic Gram-negative bacteria that belong to the family of Enterobacteriaceae. Human pathogenic *Yersiniae* species are *Y. pestis*, which causes bubonic plague, and the enteropathogenic *Y. enterocolitica* and *Y. pseudotuberculosis*, which cause a variety of diseases from diarrhoea to septicaemia as well as other diseases whose symptoms are like that of appendicitis (Cover & Aber, 1989). The usual source of infection is ingestion of contaminated food. Enteropathogenic *Yersinae* cross the mucosal membrane of the small intestine preferentially through the M cells on the Peyer's patches in the ileum. The bacteria then multiply and reside in the Payer's patches and mecentric lymph nodes (Autenrieth & Firsching, 1996). Adhesion to host cells is mediated by several mechanisms, which include the chromosomally encoded proteins invasin and Ail (Isberg & Falkow, 1985; Miller & Falkow, 1988) and the *Yersinia* virulence plasmid (pYV) encoded *Yersinia* adhesin YadA (Heesemann & Grüter, 1987; Paerregaard *et al.*, 1991; Bukholm *et al.*, 1990). The expression of the *yadA* gene is temperature regulated in *Y. enterocolitica* and *Y. pseudotuberculosis* (Skurnik & Toivanen, 1992), but in *Y. pestis* it is not expressed because of a single base-pair deletion in the *yadA* gene (Rosqvist *et al.*, 1988; Skurnik & Wolf-Watz, 1989). YadA is also involved in several virulence-related functions such as auto-agglutination (Balligand *et al.*, 1985), serum resistance (Pilz *et al.*, 1992; China *et al.*, 1993), binding to neutrophils (Ruckdeschel *et al.*, 1996) and binding to the extracellular matrix proteins such as collagen, laminin and fibronectin (Emödy *et al.*, 1989; Tertti *et al.*, 1992; Tamm *et al.*, 1993; Schulze-Koops *et al.*, 1992, 1993). The collagen-binding ability of YadA in *Y. enterocolitica* is directly related to its virulence: loss of collagen-binding ability of YadA leads to avirulence in mice (Tamm *et al.*, 1993; Roggenkamp *et al.*, 1995; El Tahir & Skurnik, 2001).

Wild-type YadA from *Y. enterocolitica* serotype O:3 is a 430 amino-acid outer membrane protein, translated with a 25 amino-acid signal peptide that is cleaved off during transport to the outer membrane. Based on the number of amino acids, the predicted molecular mass of the protein was 44 kDa (Skurnik & Wolf-Watz, 1989), but on SDS–PAGE YadA forms stable aggregates of molecular weight 160–240 kDa, suggesting oligomerization of three monomers (Gripenberg-Lerche *et al.*, 1995; El Tahir & Skurnik, 2001). Recent electron micrographs with amino-acid sequence analysis show that YadA adopts a lollipop-shaped form which consists of three domains: a C-terminal membrane anchor domain, a coiled-coil stalk domain and an oval N-terminal head domain (Hoiczyk *et al.*, 2000). The head domain consists of amino acids 26–224 and it includes seven of eight NSVAIG*xx*S motifs, which are required for YadA-mediated collagen binding (El Tahir *et al.*, 2000). Deletion mutations have also shown that the collagen binding is abolished by deletion of amino acids 83–104 or by point mutations H156Y and H158Y (Tamm *et al.*, 1993; Roggenkamp *et al.*, 1995).

The production, purification, crystallization and preliminary crystallographic analysis of the collagen-binding fragment of recombinant *Y. enterocolitica* YadA reported here are steps towards determining the three-dimensional structure of this virulence factor, which may help us to better understand its biochemical functions and bacterial collagen binding in

general and may also facilitate discovery of an anti-adhesion therapy against enteropathogenic *Yersinia*.

2. Materials and methods

2.1. Expression and purification of YadA

Because of the inherent tendency of YadA to aggregate, the shortest possible collagen-binding fragment of YadA was chosen for crystallographic studies, which meant the head domain with one extra repeat from the coiled-coil stalk domain. pOP-1 was constructed by cloning the $yada_{24–378}$ coding region as an *Nsi*I–*Hin*dIII fragment of pSS91 (Tamm *et al.*, 1993) into *Pst*I/*Hin*dIII-digested pQE30 (Qiagen). An expression plasmid for the production of recombinant N-terminally His_6-tagged $YadA_{26–241}$ was produced using the PCR method of Byrappa *et al.* (1995) in two stages. Firstly, pOP-1 was used as a template and the primers 5′-TAG CTG AGC TTG GAC TCC TG-3′ and 5′-GTC TGC ATA CGC ATT AGCG-3′ were used to shorten the *yada* gene, yielding plasmid pOP-2 carrying $yada_{24–241}$. Then, pOP-2 was used as a template and the primers 5′-GAT GAT TAC GAC GGT ATT CCT AAT TTG ACA-3′ and 5′-GTG ATG GTG ATG GTG ATG CGA TCC TC-3′ were used to remove unnecessary amino acids between the mature N-terminus of the *yada* gene and the hexahistidine tag, including amino acids 24–25 which originated from native signal sequence. Self-ligation of the latter PCR product yielded plasmid pHN-1, which was transformed into *Escherichia coli* expression strain M15(pREP4) (Qiagen).

For selenomethione-labelled protein production, I130M and I157M mutations were introduced by PCR using the primers 5′-AT**G** GGT GCG AGA GCA TCA ACT TCA GAT ACT-3′ and 5′-**C**AT GCG AAC AGA GTT TTT TGC ATC AGC TTT-3′, respectively, yielding plasmid pHN-3 (mutated bases are in bold). Both mutated isoleucines were located in NSVAIG*xx*S motifs (El Tahir *et al.*, 2000), which are known to interact with natural ligand collagen and thus were supposed to be on the surface of the protein so that the folding of the protein would not be disturbed. The mutations were confirmed by sequencing.

M15(pREP4/pHN-1) cells were grown at 310 K in Terrific Broth media supplemented with 100 µg ml^{-1} ampicillin and 25 µg ml^{-1} kanamycin. When the optical density at 600 nm reached a value of 0.9, the cells were induced with 1 m*M* isopropyl *β*-thiogalactopyranoside (IPTG) and grown for a further 3 h. Cells were harvested by centrifugation at 5000*g* for 30 min. The cell pellet was resuspended in 50 m*M* sodium phosphate pH 7.5 buffer containing 300 m*M* NaCl and stored at 203 K. After thawing on ice, cells were disrupted by sonication and centrifuged at 18 000 rev min^{-1} for 20 min at 277 K. His_6-$YadA_{26–241}$ was purified from the supernatant using metal-affinity chromatography. The protein was eluted from the column with the resuspension buffer containing 200 m*M* imidazole. The protein was further purified by size-exclusion chromatography and was eluted with 10 m*M* Tris pH 7.5 containing 300 m*M* NaCl. The fractions containing soluble YadA were pooled, concentrated and stored at 285 K for crystallization.

To produce selenomethionyl-labelled $YadA_{26–241}$, *E. coli* M15(pREP4/pHN-3) cells were grown in M9 minimal medium supplemented with ampicillin, kanamycin, thiamine and vitamins. The culture was supplemented 15 min before induction with selenomethionine and other amino acids known to inhibit methionine biosynthesis (Van Duyne *et al.*, 1993). The cells were harvested and selenomethionyl-labelled $YadA_{26–241}$ purified as for the native protein, though an additional 5 m*M* *β*-mercaptoethanol was added to all buffers to prevent oxidation of the Se atoms.

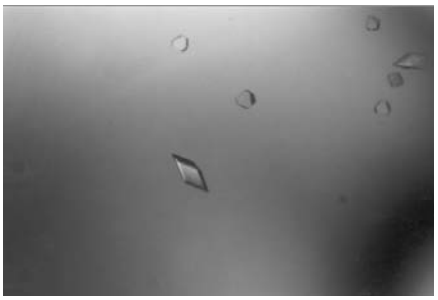

Figure 1
Single $YadA_{26–241}$ crystals photographed under a polarizing microscope. The maximum dimension of the crystals was about 200 µm.

2.2. Crystallization of YadA and data collection

Crystal screening was performed with a sparse-matrix system (Jancarik & Kim, 1991) using Hampton Research Crystal Screens. All experiments were performed using a sitting-drop vapour-phase equilibration method at room temperature. 4 µl drops containing a 1:1 ratio of protein to well solution were placed into the depression of sitting-drop plates (Hampton Research), sealed with clear tape and left at room temperature. The drop was equilibrated against 0.5 ml of reservoir solution.

Table 1
Essential crystallographic data.

Values for the highest resolution shell (1.61–1.55 Å) are given in parentheses.

Space group	*R*3
Unit-cell parameters (Å)	$a = 67.05$, $b = 67.05$, $c = 221.95$
Diffraction range (Å)	20.0–1.55
No. of unique reflections	53657 (5293)
Redundancy	5.8
Completeness (%)	99.7 (98.6)
$I/\sigma(I)$	20.2 (3.1)
R_{merge} (%)	5.2 (39.2)

Crystals were flash-frozen and X-ray data were collected at the DESY EMBL Hamburg X11 beamline using a fixed wavelength of 0.9102 Å and a MAR CCD detector. Space-group assignment and unit-cell parameters were determined using *DENZO* (Otwinowski & Minor, 1995).

3. Results

Several crystallization trials using purified $YadA_{24–378}$ were unsuccessful. Dynamic light-scattering and native-gel analysis indicated that $YadA_{24–378}$ was aggregated in solution (data not shown). Deletion of the coiled-coil stalk yielded the more soluble fragment $YadA_{26–241}$, which was still able to bind collagen (data not shown). In the size-exclusion step, aggregated forms were separated from the soluble $YadA_{26–241}$, which was critical for crystallization. The amount of 95% pure protein (as estimated from SDS–PAGE) obtained from a 1 l culture was about 10 mg.

Initial crystal screening was performed using a low concentration of protein (1 mg ml^{-1}) and several conditions yielded crystals overnight. However, the largest, diamond-shaped, crystals were obtained using 15% polyethylene glycol 8000 (PEG 8000) as a precipitant at pH 6.5 with 0.1 *M* sodium acetate (Fig. 1). By increasing the protein concentration to 8 mg ml^{-1}, we were able to produce crystals of dimensions 200 × 75 × 20 µm in 12% PEG 8000 within a week.

The crystals were flash-frozen after a short soak in a solution containing 12% PEG 8000, 100 m*M* sodium cacodylate pH 6.5, 0.1 *M* sodium acetate and 20% ethylene glycol. The crystals belonged to the trigonal space group *R*3, with unit-cell parameters $a = b = 67.05$, $c = 221.95$ Å (Table 1). A high-resolution data set consisting of a total of 180 0.5° oscillation frames (Fig. 2) and a low-resolution native data set consisting of 30 3° oscillation frames were collected from a single flash-frozen crystal at the synchrotron

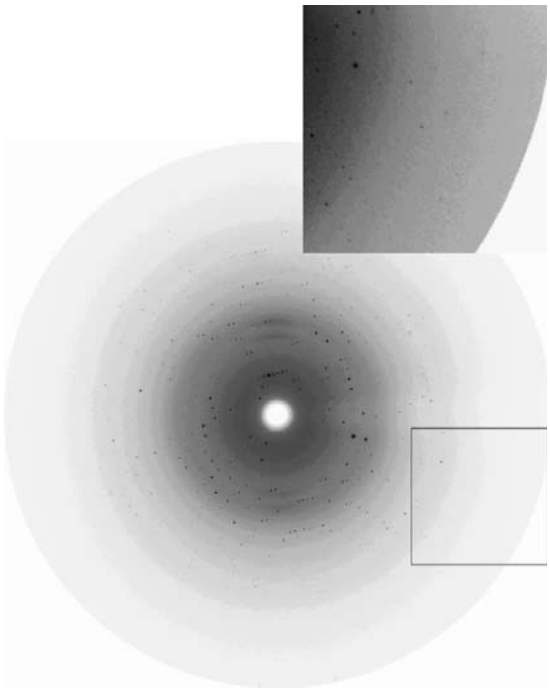

Figure 2
Diffraction pattern of the crystal. The close-up of the outer part of the picture is shown at a higher saturation in order to show the diffraction spots extending below 1.5 Å.

in EMBL, Hamburg and merged together. Based on the Matthews coefficients, there can either be one (4.00 $Å^3\ Da^{-1}$) or two (2.00 $Å^3\ Da^{-1}$) molecules in an asymmetric unit, corresponding to calculated solvent contents of 69 or 38% by volume of the unit cell, respectively (Matthews, 1968). However, given that the crystals diffract to high resolution, the lower solvent content seems more likely.

Expression of the selenomethionyl protein led to crystals under similar conditions to native crystals, using 11% PEG 8000 and 0.2 *M* sodium acetate but Tris–HCl pH 6.5 as precipitant solution and a protein concentration of 2.5 mg ml^{-1}. The buffer was changed from sodium cacodylate which was used for native $YadA_{26–241}$ crystallizations, because the arsenic in cacodylate has an anomalous adsorption edge close to that of selenium and might cause problems in MAD data collection. Both Tris–HCl and MES buffers were tried, but crystals were only obtained using Tris–HCl; the best crystals so far appeared at pH 6.5, although the buffering capacity is low for Tris at this pH. Diamond-shaped crystals grew in a week but were smaller in size, with a maximum dimension of 80 µm. We are presently optimizing crystallization conditions for SeMet crystals in order to obtain crystals suitable for MAD data collection.

We thank Dr Veli-Matti Leppänen, Dr Michael Merckel and Lari Lehtiö MSc for help collecting data and Dr Tuomo Glumoff for dynamic light-scattering measurements. We also thank Elaine Connor, Maria Rehn, Katja Mattsson, Merja Nissinen and Raija Andersen for technical help. This work was supported by National Graduate School in Informational and Structural Biology and the Academy of Finland (Nos. 68155 to AG and 42105 to MS).

References

Autenrieth, I. B. & Firsching, R. (1996). *J. Med. Microbiol.* **44**, 285–294.
Balligand, G., Laroche, Y. & Cornelis, G. (1985). *Infect. Immun.* **48**, 782–786.
Bukholm, G., Kapperud, G. & Skurnik, M. (1990). *Infect. Immun.* **58**, 2245–2251.
Byrappa, S., Gavin, D. K. & Gupta, K. C. (1995). *Genome Res.* **5**, 404–407.
China, B., Sory, M. P., Nguyen, B. T., Debruyere, M. & Cornelis, G. R. (1993). *Infect. Immun.* **61**, 3129–3136.
Cover, T. L. & Aber, R. C. (1989). *N. Engl. J. Med.* **321**, 16–24.
El Tahir, Y., Kuusela, P. & Skurnik, M. (2000). *Mol. Microbiol.* **37**, 192–206.
El Tahir, Y. & Skurnik, M. (2001). *Int. J. Med. Microbiol.* **291**, 209–218.
Emödy, L., Heesemann, J., Wolf-Watz, H., Skurnik, M., Kapperud, G., O'Toole, P. & Wadström, T. (1989). *J. Bacteriol.* **171**, 6674–6679.
Gripenberg-Lerche, C., Skurnik, M. & Toivanen, P. (1995). *Infect. Immun.* **63**, 3222–3226.
Heesemann, J. & Grüter, L. (1987). *FEMS Microbiol Lett.* **40**, 37–41.
Hoiczyk, E., Roggenkamp, A., Reichenbecher, M., Lupas, A. & Heesemann, J. (2000). *EMBO J.* **19**, 5989–5999.
Isberg, R. R. & Falkow, S. (1985). *Nature (London)*, **317**, 262–264.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Miller, V. L. & Falkow, S. (1988). *Infect. Immun.* **56**, 1242–1248.
Otwinowski, A. & Minor, W. (1995). *DENZO.* Yale University, New Haven, Connecticut, USA.
Paerregaard, A., Espersen, F. & Skurnik, M. (1991). *APMIS*, **99**, 226–232.
Pilz, D., Johns, M. A. & McCabe, W. R. (1992). *Infect. Immun.* **60**, 189–195.
Roggenkamp, A., Neuberger, H.-R., Flügel, A., Schmoll, T. & Heesemann, J. (1995). *Mol. Microbiol.* **16**, 1207–1219.
Rosqvist, R., Skurnik, M. & Wolf-Watz, H. (1988). *Nature (London)*, **334**, 522–525.
Ruckdeschel, K., Roggenkamp, A., Schubert, S. & Heesemann, J. (1996). *Infect. Immun.* **64**, 724–733.
Schulze-Koops, H., Burkhardt, H., Heeseman, J., von der Mark, K. & Emmrich, F. (1992). *Infect. Immun.* **60**, 2153–2159.
Schulze-Koops, H., Burkhardt, H., Heeseman, J., Kirsch, T., Swoboda, B., Bull, C., Goodman, S. & Emmirch, F. (1993). *Infect. Immun.* **61**, 2513–2519.
Skurnik, M. & Toivanen, P. (1992). *J. Bacteriol.* **158**, 1033–1036.
Skurnik, M. & Wolf-Watz, H. (1989). *Mol. Microbiol.* **3**, 517–529.
Tamm, A., Tarkkanen, A. M., Korhonen, T., Kuusela, P., Toivanen, P. & Skurnik, M. (1993). *Mol. Microbiol.* **10**, 995–1011.
Tertti, R., Skurnik, M., Vartio, T. & Kuusela, P. (1992). *Infect. Immun.* **60**, 3021–3024.
Van Duyne, G. D., Standaert, R. F., Karplus, P. A., Schreber, S. L. & Clardy, J. (1993). *J. Mol. Biol.* **229**, 105–124.