Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of DraD invasin from uropathogenic *Escherichia coli*: a dimer with swapped β-tails

Robert Jędrzejczak,[a] Zbigniew Dauter,[a]* Mirosława Dauter,[b] Rafał Piątek,[c] Beata Zalewska,[c] Marta Mróz,[c] Katarzyna Bury,[c] Bogdan Nowicki[d] and Józef Kur[c]*

[a]Synchrotron Radiation Research Section, MCL, National Cancer Institute, Argonne National Laboratory, Argonne, IL 60439, USA, [b]SAIC-Frederick Inc., Basic Research Program, Argonne National Laboratory, Argonne, IL 60439, USA, [c]Department of Microbiology, Gdańsk University of Technology, ul. Narutowicza 11/12, 80-952 Gdańsk, Poland, and [d]Department of Obstetrics and Gynecology, The University of Texas Medical Branch, Galveston, TX 77555, USA

Correspondence e-mail: dauter@anl.gov, kur@chem.pg.gda.pl

Received 16 September 2005
Accepted 8 November 2005

PDB Reference: DraD invasin, 2axw, r2axwsf.

The *dra* gene cluster of uropathogenic strains of *Escherichia coli* produces proteins involved in bacterial attachment to and invasion of the eukaryotic host tissues. The crystal structure of a construct of *E. coli* DraD possessing an additional C-terminal extension of 13 amino acids, including a His_6 tag, has been solved at a resolution of 1.05 Å. The protein forms symmetric dimers through the exchange of the C-terminal β-strands, which participate in the immunoglobulin-like β-sandwich fold of each subunit. This structure confirms that DraD is able to act as an acceptor in the donor-strand complementation mechanism of fiber formation but, in contrast to DraE adhesin, its native sequence does not have a donor strand; therefore, DraD can only be located at the tip of the fiber.

1. Introduction

The protein products of the *dra* gene cluster of diffusely adhering *Escherichia coli* (DAEC) strains take part in the invasion of eukaryotic cells by these bacteria, in particular in urinary tract infections and diarrheal diseases (Van Loy *et al.*, 2002). The *dra* cluster codes for the following proteins: DraA, DraB, DraC, DraD, DraE and DraP (Nowicki *et al.*, 1989). The DraE adhesin forms homopolymeric fimbriae at the surface of the bacterium, synthesized at the cell surface by the chaperone–usher mechanism, with DraB as the chaperone and DraC as the usher. The oligomerization of DraE occurs by the donor-strand complementation (DSC; Choudhury *et al.*, 1999; Sauer *et al.*, 2002, 2004) mechanism, in which the N-terminal β-strand of one subunit completes the immunoglobulin-like structure of the next subunit, forming a linear polymer. The attachment of consecutive subunits of DraE is accomplished with the cooperation of the chaperone DraB, which shields one subunit of DraE by DSC and is then replaced by the other DraE molecule in the process of donor-strand exchange (DSE; Soto & Hultgren, 1999; Sauer *et al.*, 2004; Piątek *et al.*, 2005). Recent work (Anderson *et al.*, 2004) suggests that the fimbrial fiber formed by DraE is capped at the tip by the molecule of DraD invasin, again joined to the last DraE subunit by DSC. The expression of the fimbrial protein DraE at the cell surface does not require the expression of DraD and, conversely, the surface expression of DraD does not depend on the expression of the adhesin DraE and, in addition, can be independent of the outer-membrane channel DraC (Zalewska *et al.*, 2005).

The molecular mechanism of fimbriae formation has been formulated on the basis of X-ray and NMR elucidation of several three-dimensional structures of DSC-connected adhesins and chaperones related or analogous to the Dra proteins. The structures of the following chaperone–adhesin complexes are available in the PDB (Berman *et al.*, 2000):

© 2006 International Union of Crystallography
Printed in Denmark – all rights reserved

PapD–PapK (from P pili of *E. coli*; PDB code 1pdk; Sauer *et al.*, 1999), PapD–PapE (1n0l, 1n1z; Sauer *et al.*, 2002), FimC–FimH (from type 1 pili of *E. coli*; 1qun, Choudhury *et al.*, 1999; 1kiu, 1klf, Hung *et al.*, 2002; 1ze3, Nishiyama *et al.*, 2005), Caf1M–Caf1 (from F1 antigen of *Yersinia pestis*; 1p5u, 1p5v, 1z9s; Zavialov *et al.*, 2003, 2005). In all of these complexes, the participating molecules interact by the DSC mechanism.

The structures of native or modified adhesins are also available. The fragment of SdrG adhesin has two domains joined by the donor strand (from *Staphylococcus epidermis*; PDB codes 1r17, 1r19; Ponnuraj *et al.*, 2003). The native *E. coli* AfaE-III and DraE adhesins in trimeric forms (PDB codes 1usq, 1usz, 1ut1, 1ut2; Pettigrew *et al.*, 2004) show an unusual association in which only part of the *A* strand of one monomer contributes to the fold of the neighboring subunit, but the N-terminal strand is folded back and participates in the β-sandwich of the original subunit. This is an example of the swapping of a small domain, but not donor-strand complementation; therefore, these molecules were classified as misfolded as a result of their expression in the absence of a chaperone.

The structure of the modified *E. coli* adhesin AfaE-III has been solved by NMR (PDB code 1rxl; Anderson *et al.*, 2004). This construct of AfaE-dsc was engineered by removing the 16-residue N-terminal donor strand and inserting it at the C-terminus. This strand folded back and participated in the formation of a self-complementing monomeric structure. On the basis of this structure, the authors proposed that *in vivo* AfaE and analogous DraE adhesins assemble head-to-tail, forming linear filaments through the DSC mechanism which are capped with a single molecule of invasin at the tip. The hypothesis of the DSC interaction between adhesin AfaE and invasin AfaD was supported by NMR titration of AfaD with the 18-amino-acid peptide corresponding to the N-terminus of AfaE.

No three-dimensional structure of a bacterial DraD-like invasin is known to date. In this report, the high-resolution structure of a construct of the DraD invasin possessing an additional 13-residue extension at the C-terminus, which contains a linker and a 6×His tag, is presented. Fig. 1 shows the sequence of this construct.

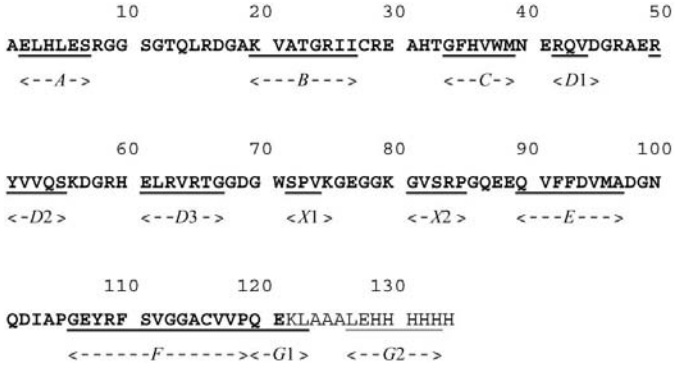

Figure 1
Sequence of DraD. The original protein sequence is in bold, the C-terminal addition is in non-bold letters and β-sheets are underlined.

2. Materials and methods

2.1. Protein expression and purification

10 l LB medium containing 100 μg ml^{-1} kanamycin was inoculated with 100 ml of an overnight culture of *E. coli* BL21 (DE3) cells carrying plasmid pInvD-C-His encoding the full sequence of DraD extended by 13 amino acids, including a C-terminal 6×His tag (Zalewska *et al.*, 2005). Bacteria were grown at 310 K and 250 rev min^{-1} and on reaching an OD_{600} of 0.6 were cooled on ice to 295 K, followed by addition of IPTG to a concentration of 0.3 m*M*. After 2–3 h, the culture was harvested by 10 min centrifugation at 6000*g*, 277 K and immediately suspended in 150 ml buffer containing 20% sucrose, 5 m*M* EDTA, 30 m*M* Tris pH 8.0. After 10 min at room temperature, the mixture was centrifuged and the pellet was resuspended in 150 ml of an ice-cold solution of 10 m*M* $MgCl_2$. After another 10 min on ice, the solution was centrifuged at 6000*g* at 277 K for 20 min. The supernatant containing the periplasmic proteins was collected and conditions were adjusted to the Ni–NTA loading buffer: 0.3 *M* NaCl, 5 m*M* imidazole, 30 m*M* Tris pH 7.6. This protein solution was loaded onto Ni–NTA (Qiagen). The bound proteins were eluted with 200 m*M* imidazole, 0.3 *M* NaCl, 30 m*M* Tris pH 7.6. The fractions containing DraD were precipitated with 0.516 g solid ammonium sulfate per millilitre of protein solution. After 30 min of gentle agitation, the suspension was centrifuged. The pellet was resuspended in 2 ml 0.2 *M* NaCl, 10 m*M* HEPES pH 7.5 and loaded onto a Superdex 200 column (GE Healthcare). The fractions containing DraD were collected and concentrated on Centricon YM-3 (Amicon) to 7 mg ml^{-1}.

2.2. Diffraction data collection, phasing and model building

A sample of the obtained DraD crystallized in a hanging drop from solution consisting of 7 mg ml^{-1} protein and 0.2 *M* NaCl in 10 m*M* HEPES buffer pH 7.5 after mixing in a 2:1 proportion with well solution containing 32% PEG 2000 MME and 15% glycerol in 0.1 *M* MES buffer pH 6.0. Only one large crystal grew after about a week and had approximate dimensions of 0.1 × 0.1 × 1.0 mm. This crystal was broken into several fragments and various diffraction data were collected from separate pieces of this crystal at the SER-CAT sector 22 beamlines of the Advanced Photon Source, Argonne National Laboratory. The well solution served as cryosolution and all data were measured at a temperature of 100 K.

The first set of native data was measured to 1.55 Å resolution and the crystal unit cell was interpreted as orthorhombic, space group $P2_12_12_1$, with approximate unit-cell parameters $a = 33$, $b = 61$, $c = 112$ Å. Unusually, but quite beneficially from the point of view of its orientation with respect to the cryoloop and the spindle axis, the long physical dimension of the crystal corresponded to its longest unit-cell parameter.

The other fragment of the crystal was soaked for about 2 h in well solution containing in addition about 20 m*M* $KAu(CN)_2$. However, diffraction data measured on this crystal with a wavelength of 1.0 Å and quickly processed with *HKL*2000 (Otwinowski & Minor, 1997) did not show any trace

Table 1
Diffraction data and refinement statistics.

Values in parentheses are for the highest resolution shell.

	LiBr-soaked	Native
Data statistics		
Beamline	SER-CAT 22BM	SER-CAT 22ID
Wavelength (Å)	0.9184	1.000
Space group	$P2_12_12_1$	$P2_12_12_1$
Unit-cell parameters (Å)		
a	32.89	33.59
b	60.89	61.82
c	113.12	112.20
Resolution (Å)	30.0–1.76 (1.82–1.76)	30.0–1.05 (1.09–1.05)
Measured reflections	165371 (12494)†	612277 (48872)
Unique reflections	43534 (4215)†	109569 (10677)
$I/\sigma(I)$	13.7 (2.9)†	21.4 (2.9)
Completeness (%)	99.7 (98.1)†	99.0 (97.6)
$R_{\rm merge}$	0.079 (0.332)†	0.066 (0.388)
$R_{\rm anom}$	0.047 (0.178)	
Wilson *B* factor (Å)	13.7	9.1
Matthews coefficient (Å^3 Da^{-1})	1.91	1.91
Solvent content (%)	35	35
Phasing statistics		
FOM	0.651	
Contrast	0.404	
Continuity	0.867	
Refinement statistics		
Resolution (Å)	30–1.76	30–1.05
R factor	0.181	0.151
$R_{\rm free}$	0.233	0.168
No. of protein non-H atom sites	1950	2114
No. of solvent water molecules	258	291
R.m.s. deviation from ideality		
Bonds (Å)	0.016	0.015
Angles (°)	1.73	1.66
Ramachandran plot regions (%)		
Most favored	94.8	95.4
Additionally allowed	5.2	4.6

† Friedel mates treated as independent reflections.

of anomalous signal. The next fragment was soaked in well solution supplemented with 1 *M* LiBr for about 4 s, according to the short cryosoaking with halides approach (Dauter *et al.*, 2000); diffraction data were measured with a wavelength of 0.9184 Å, corresponding to 'near-remote high-energy' X-rays, 100 eV above the absorption edge of Br. A small amount of anomalous signal was present in the diffraction data. Analogous data were then measured from the fragment soaked in the same LiBr solution for 12 s, resulting in a greater anomalous signal. The third fragment was soaked in LiBr for 40 s and produced a diffraction data set with a very significant amount of anomalous signal; these data were used for structure solution by the single-wavelength anomalous dispersion (SAD) phasing approach. Using a MAR 225 CCD detector, all these data were measured at the bending-magnet beamline BM22.

The remaining, largest fragment (0.1 × 0.1 × 0.25 mm) of the original crystal was used to collect native data at the undulator beamline ID22 with a MAR 300 CCD detector. This crystal diffracted to 1.05 Å resolution and data were measured in three passes with different exposure times and beam attenuation to ensure adequate estimation of the weakest and strongest reflections. The statistics of the diffraction data sets

Table 2
Unit-cell parameters of various data sets, with r.m.s.d.s in parentheses.

Data set	*a* (Å)	*b* (Å)	*c* (Å)	Volume (Å^3)
Native, medium resolution	34.01 (9)	60.85 (8)	111.44 (8)	230600
Native, high resolution	33.59 (3)	61.82 (2)	112.20 (7)	233000
LiBr soaked for 4 s	33.46 (9)	61.62 (12)	112.64 (12)	232200
LiBr soaked for 12 s	32.85 (9)	60.56 (4)	112.72 (30)	224200
LiBr soaked for 40 s	32.89 (9)	60.89 (15)	113.12 (17)	226500
$KAu(CN)_2$ soaked	33.69 (5)	61.61 (13)	112.30 (12)	233100

used for structure solution and refinement are shown in Table 1. Even though all fragments used for data collection originated from the same crystal, the data sets have unit-cell parameters that vary considerably, by more than 1 Å in the shortest dimension. The brominated and native data are therefore significantly non-isomorphous.

The 1.75 Å resolution data measured from the crystal soaked in LiBr for 40 s were used for structure solution by the SAD technique. The Bijvoet differences were extracted by *XPREP* (Sheldrick, 2005) and submitted to *SHELXD* (Schneider & Sheldrick, 2002), run at 2.2 Å resolution, using the 1500 largest *E* values and searching for 12 Br sites. All 100% phase trials of *SHELXD* gave correct anomalous substructure solutions, equivalent by enantiomorph or different cell origin, characterized by a correlation coefficient between all/weak reflections of about 36/21%. A *SHELXE* (Sheldrick, 2002) run based on the 12 identified bromide sites was used to calculate the protein phases, which were submitted to *ARP*/*wARP* (Perrakis *et al.*, 1999) in the *warpNtrace* mode. This procedure automatically built 154 residues in several chains, but the resulting electron-density map permitted the manual building of an almost complete model of two molecules of DraD in the asymmetric unit using the graphics program *QUANTA* (Accelrys Inc, San Diego, California, USA). *REFMAC*5 (Collaborative Computational Project, Number 4, 1994; Murshudov *et al.*, 1997) was used to refine this model and *ARP* (Lamzin & Wilson, 1997) served for modeling the solvent waters.

The dimer was then positioned in the native cell with *AMoRe* (Collaborative Computational Project, Number 4, 1994; Navaza, 1994) and refined at first isotropically and subsequently anisotropically at 1.05 Å resolution with *REFMAC*5 using the 'riding' model of H atoms. Default (Engh & Huber, 1991) values of geometrical restraints were used during all refinement cycles.

3. Results and discussion

3.1. Structure solution

The crystal structure of DraD was solved and refined based on diffraction data collected from parts of a single crystal of this protein. This specimen was broken into fragments and each part was used to collect a different type of data: two native sets of medium and atomic resolution, one unsuccessful gold derivative and three sets of bromine-derivative data containing varying amounts of anomalous dispersion signal.

Two of these data sets were ultimately used: the 1.75 Å resolution data from a crystal soaked in 1 *M* LiBr for 40 s for structure solution by SAD and the 1.05 Å resolution native data for final model refinement.

It is noteworthy that despite originating from the same specimen, the individual fragments had unit-cell parameters that differed significantly from each other (Table 2). The r.m.s. deviations included in the table are estimated from the variation of unit-cell parameters refined for consecutive triples of diffraction images during their integration by *HKL*2000. Soaking in LiBr caused the shrinkage of the shortest and medium unit-cell parameters by 2 and 1.5% and elongation of the longest one by 0.8% with respect to the high-resolution native data, which may be a manifestation of structural changes in the DraD molecules and/or their packing in the crystal (see below). The two native data sets differ by 1.2, 1.6 and 0.7% in the three unit-cell parameters, which suggests that either various parts of the same crystal may differ significantly or the non-isomorphism results as a consequence of small variations in the crystal-handling procedures, such as the soaking time in the cryosolution or the speed of transfer from the drop to the cold stream at the goniostat.

The structure of DraD was solved by SAD and its model was built and refined at 1.75 Å resolution using data from the crystal soaked in LiBr for 40 s. This structure will be denoted DraD-Br. The process of initial phasing was straightforward and rapid up to the point of interpretation of the initial electron-density map. The automatic *warpNtrace* procedure built a little more than half the expected structure, in the form of 24 separate main-chain fragments where only two fragments (36 residues) had side chains correctly built according to the sequence. Construction of the complete model required manual effort at the graphics display of the map. Part of the reason for these difficulties lies in the low solvent content of the crystals, 35%, which diminishes the power of density-modification procedures.

The DraD-Br structure was refined to an *R* factor of 18.1% and an $R_{\rm free}$ of 23.3%. In a few places the electron density was not interpretable and the main chain of two residues (Arg8*A*–Gly9*A*) and side chains of 21 residues could not be reliably built. According to *PROCHECK* (Laskowski *et al.*, 1993) all residues in the DraD-Br model have conformations lying in the allowed regions of the Ramachandran plot.

The dimer was then positioned in the native cell and refined against the native data, first isotropically at 1.5 Å resolution and eventually anisotropically at 1.05 Å with *SHELXL*. The final *R* factor was 15.1% and $R_{\rm free}$ was 16.8%. Four residues in the *A* chain (Gly9*A*, Gly10*A*, Ser11*A* and Gly12*A*) and two in the *B* chain (Gly10*B* and Ser11*B*) have weak and unclear electron density, suggesting partial disorder. However, all amino acids were modeled with side chains, some in rather weak density, mostly the charged ones, and exposed to the bulk solvent. Several occurances of alternative conformations of the side chains were evident. However, almost all the main chain and a majority of side chains lie in well defined electron density (Figs. 2*a* and 2*b*).

3.2. Description of the structure

The structure consists of a symmetric dimer, with the long C-terminal extensions swapped between two participating monomers in the form of the long β-strands (Figs. 3 and 4). The 13 C-terminal residues do not belong to the original sequence of DraD, but correspond to the linker and 6×His tag added for the purpose of protein isolation and purification. The monomer of DraD has a somewhat modified immunoglobulin (Ig) fold (Bork *et al.*, 1994), forming a β-sandwich with the participation of the C-terminus of the second monomer as one of the strands. This interaction is analogous to the donor-strand complementation (DSC) mechanism described for the formation of fimbriae (Choudhury *et al.*, 1999; Sauer *et al.*, 1999, 2002).

The topology of DraD is presented in Fig. 3. It differs somewhat from the canonical Ig fold. The β-barrel is 'unzipped', since the *G* strand donated by the C-terminus of the second subunit forms a β-sheet only with the *F* strand, but does not make any main-chain hydrogen bonds with the *A* strand. The *F* and *G* strands of both subunits form a single

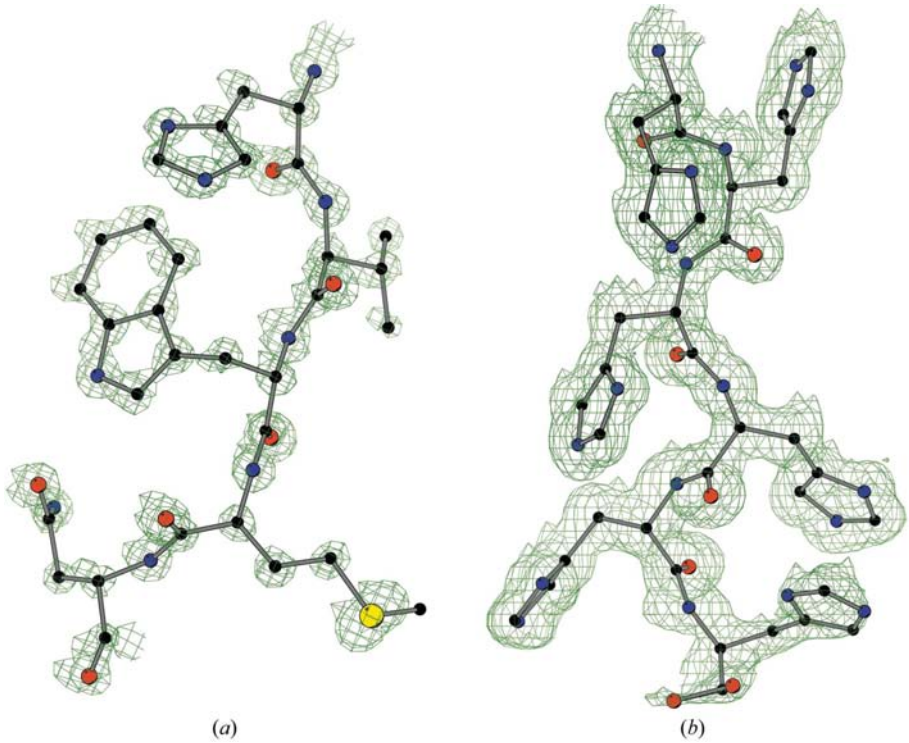

Figure 2
(*a*) A fragment of β-strand 36*B*–40*B* with superimposed $2F_{\rm o} - F_{\rm c}$ electron density at the 4σ level. (*b*) The C-terminal His_6 tag 129*A*–134*A* with $2F_{\rm o} - F_{\rm c}$ electron density at the 1σ level.

long ladder spanning the length of the dimer. The *G* strand in both subunits has a kink in half of its length at three consecutive alanine residues and can be formally divided into the *G*1 and *G*2 strands. The long *FG*:*GF* ladder is therefore not perfect and only partially complements the β-sandwich. The presence of a kink and mostly polar residues at the end (including the last six histidines) prevent the *G* strand from effective completion of the hydrophobic interior of the sandwich and from formation of the parallel ladder with the *A* strand.

In addition, the positions of the minor strands differ from typical variations of the Ig fold (Bork *et al.*, 1994); there is a single *C* strand in the back sheet, three short strands classified as *D*1, *D*2 and *D*3 in the front sheet and two additional short strands in the back sheet marked *X*1 and *X*2. All ladders in DraD are antiparallel.

Nevertheless, such variations of the main Ig theme are typical for different fimbrial proteins from various sources and in general for Ig-like proteins (Halaby *et al.*, 1999). Several X-ray and NMR structures of such proteins are available in the PDB and most of them incorporate the *G* strand from another identical or different molecule, with the *G* strand forming a β-ladder on one side only to the *F* strand. In one of the complexes, the *G* strand is formed by a separate peptide and only then is the β-barrel closed on both sides of this strand. Such a DSC mechanism is utilized in forming long fibers of adhesins, with the cooperation of specific chaperones which prevent the formation of closed oligomers. In the present case, the DraD construct possessed the long C-terminal extension which serendipitously played the role of the donor *G* strand accepted by the second molecule of DraD in a mutual fashion and forming a symmetric dimer. The original DraD sequence does not have the *G* strand but is capable of accepting it from its natural partner, the fiber-forming DraE, capping the fiber. The *G* strand of the DraE subunit is located at its N-terminus.

The only disulfide bridge in DraD joins Cys28 at the end of strand *B* and Cys116 within strand *F*, thus coupling both β-sheets of the sandwich. The S—S bridge between the *B* and *F* strands is typical of Ig domains (Halaby *et al.*, 1999).

The two subunits within the dimer are very similar, except that one loop around Gly58 has a different conformation in the two chains owing to differences in packing contacts. The

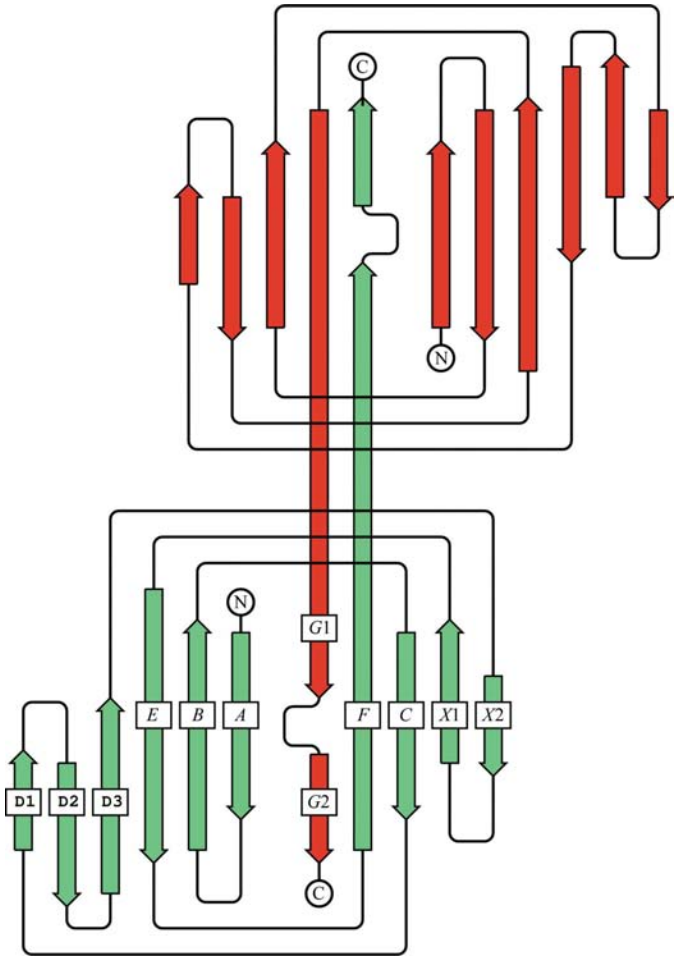

Figure 3
Topology of the DraD dimer.

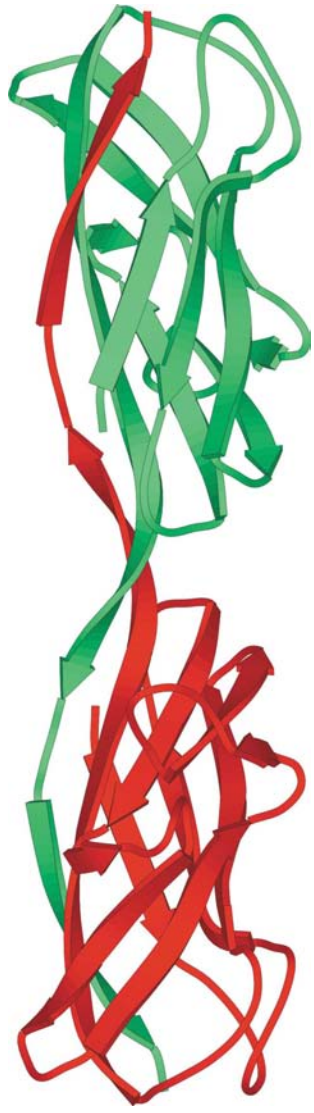

Figure 4
The DraD dimer positioned with its local twofold axis horizontal.

subunits superimpose with an r.m.s.d. of 0.69 Å between 127 pairs of CA atoms (omitting the poorly defined residues 9–12 in both subunits and the three loop residues 57–59). For this superposition individual compact 'structural' subunits were defined encompassing the first 119 residues of one chain and residues 120–134 (*G* strand) from the other chain.

3.3. Packing

The DraD subunits have an elongated shape and dimensions of about 25 × 25 × 40 Å. In the dimer they are positioned head-to-head along their longest dimension, swapping their C-terminal *G* strands (Fig. 4). The dimer is therefore about 85 Å long, oriented in the *bc* plane of the cell approximately along the diagonal direction between the *b* and *c* axes (Fig. 5). Each dimer is in crystal contact with 12 other dimers (Table 3). It is surrounded by two parallel dimers related by positive and negative translations along the shortest crystal *a* axis and by four antiparallel dimers related by the 2_1 axes parallel to the *x* direction. The layer formed by these parallel and antiparallel dimers transformed by any of the remaining 2_1 axes, parallel to the *b* or *c* directions, results in another layer, with the dimers oriented perpendicularly to the previous layer packing in a zigzag fashion along the *c* axis. Each dimer is in contact with six such 'perpendicular' dimers. The resulting packing of DraD is relatively tight, with about 40% of the surface shielded by symmetry-equivalent molecules and a Matthews parameter of 1.91 Å^3 Da^{-1}, corresponding to about 35% solvent content in the crystal (Matthews, 1968). The tight packing of this modest-sized protein is in part responsible for the very high resolution diffraction of the crystals. However, some parts of the structure are very flexible, including the fragment of the main chain residues 9–12 in both subunits and several long and charged side chains.

Table 3
Accessible surface area of DraD (Å^2).

Dimer isolated	14339
Subunit *A* (with 120*B*–134*B*)	7377
Subunit *B* (with 120*A*–134*A*)	7362
Dimer in the crystal	8667
Surface shielded by symmetric molecules	
$1 + x, y, z$	429
$-1 + x, y, z$	412
$1/2 - x, 1 - y, 1/2 + z$	549
$-1/2 - x, 1 - y, 1/2 + z$	293
$1/2 - x, 1 - y, -1/2 + z$	559
$-1/2 - x, 1 - y, -1/2 + z$	328
$1/2 + x, 1/2 - y, -z$	568
$-1/2 + x, 1/2 - y, -z$	567
$1/2 + x, 3/2 - y, -z$	401
$-1/2 + x, 3/2 - y, -z$	465
$-x, 1/2 + y, -1/2 - z$	546
$-x, -1/2 + y, -1/2 - z$	555
Total	5672

In spite of the low solvent content, soaking of the DraD crystal in cryosolution containing 1 *M* LiBr caused the bromide ions to penetrate into the ordered solvent regions around the protein surface. In addition, in both the native and

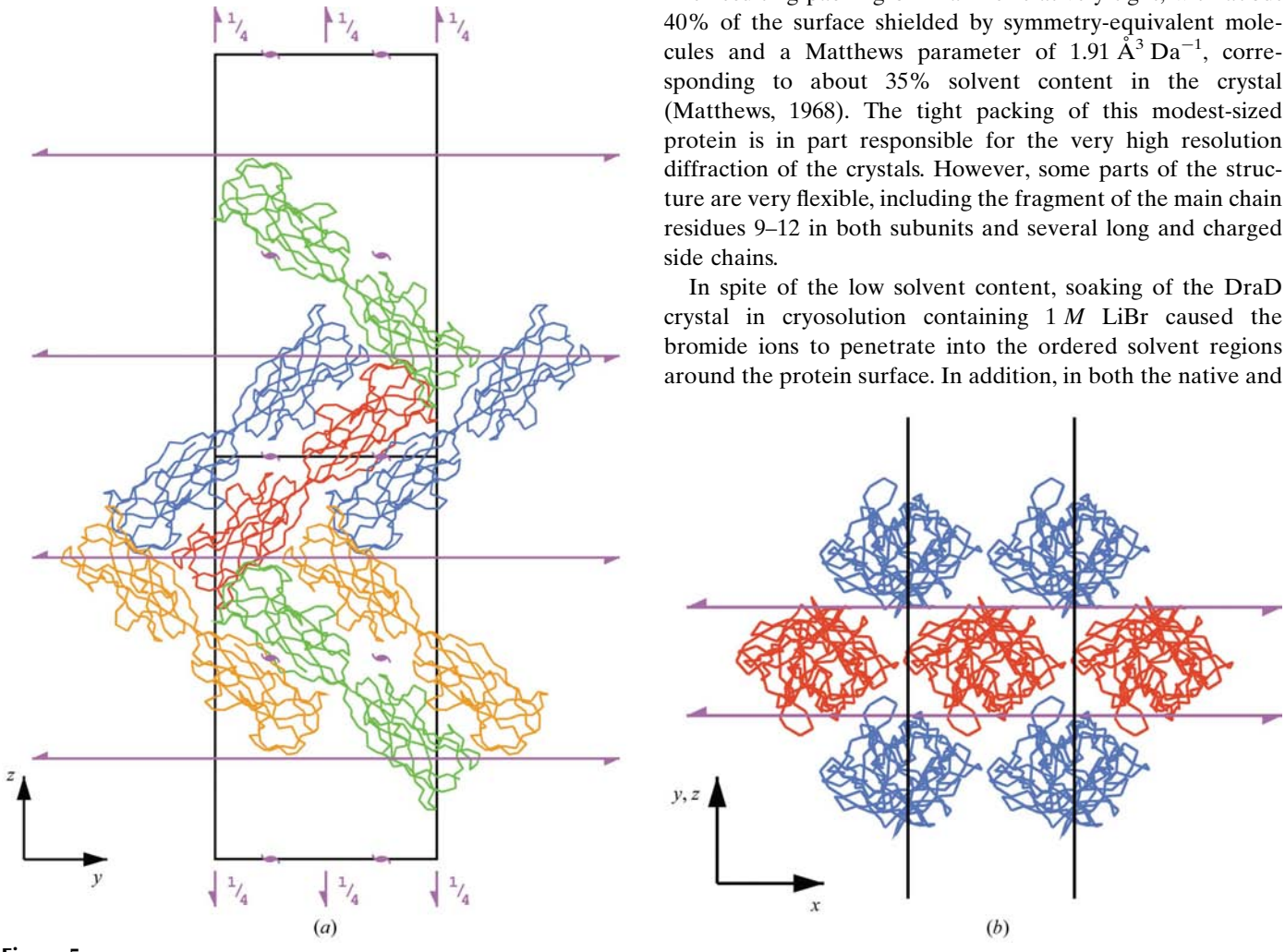

Figure 5
Packing of DraD dimers in the crystal unit cell. (*a*) Projection along the *x* axis, (*b*) view along the length of the dimer. The red molecules are related by translation along *x*, the blue molecules are related to red ones by the screw axis parallel to *x*, the yellow molecules are related by the screw axis along *y* and the green molecules are related by the screw axis parallel to the *z* axis.

brominated DraD structures one fully occupied glycerol molecule is present, hydrogen bonded to three protein molecules. In the native structure one chloride ion was identified, interacting with two arginine guanidinium groups, Arg50*A* and Arg65*A*, and hydrogen bonded to the Asn100*B* peptide amide NH of another molecule.

However, diffusion of bromides resulted in some rearrangements of protein molecules in the crystal. The crystal soaked in LiBr solution for 4 s showed the poorest quality diffraction data and somewhat diffuse reflection profiles, whereas crystals soaked for a longer time produced data of higher quality. Various rearrangements caused by soaking in halide salts have been observed previously; for example, a change of the lattice from orthorhombic to hexagonal symmetry (Dauter *et al.*, 2001) or a transformation from the twinned monoclinic lattice to non-twinned tetragonal form (Declercq & Evrard, 2001).

The non-isomorphism between the native and brominated crystals is manifested not only by the differences in unit-cell parameters (Table 2), but also by a variation in the mutual disposition of the two subunits within the dimer. In both structures the dimer is not ideally symmetric, but in the native structure optimal superposition of the two subunits involves rotation by 170°, whereas in the brominated DraD it requires rotation by 165°. The angular difference of 5° is equivalent to a shift of 3.5 Å at the end of the 40 Å long subunit.

3.4. Comparison with related structures

No other structure of a DraD-like bacterial invasin is known, but a similar Ig-type fold and the presence of the external DSC strand have been observed in several X-ray and NMR structures of adhesins present in the PDB.

In spite of very low sequence identity of below 15%, the structure of DraD shows common features with each of the above-mentioned adhesin molecules. The overall fold of DraD differs in the disposition of minor strands, but it preserves the general immunoglobulin topology with two antiparallel β-sheets forming a sandwich in which one of the sheets contains the externally donated *G* strand. However, the individual strands have a somewhat different twist than in adhesins and it is not possible to effectively superimpose the DraD molecule on any of the adhesin models listed above.

4. Conclusions

The high-resolution structure of DraD-His_6 confirms the ability of DraD to aggregate by accepting the external complementary donor *G* strand. The additional C-terminal extension of 13 amino acids serendipitously performed the role of such a donor strand, which permitted the molecules to form a symmetric dimer. The sequence of additional residues was not selected for its role as the *G* strand and there is a kink in this strand which optimizes the fit of hydrophobic side chains in the interior of the β-sandwich. In contrast to DraE-like adhesins, the native sequence of DraD does not have such a strand; therefore, it can only act as an acceptor, not as a donor, in the DSC mechanism. This confirms the notion that DraD is located at the end of the fibrillium, capping its tip.

This work was supported in part by the Polish State Committee for Scientific Research, project No. 2P04A 039 27 to JK and 2P05A 119 27 to BZ, in part by Federal funds from the National Cancer Institute, National Institutes of Health under contract No. NO1-CO-12400 and in part by the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research. Use of the Advanced Photon Source X-ray facility was supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No. W-31-109-Eng-38. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products or organizations imply endorsement by the US or any other Government.

References

Anderson, K. L., Billington, J., Pettigrew, D., Cota, E., Simpson, P., Roversi, P., Chen, H. A., Urvil, P., du Merle, L., Barlow, P. N., Medof, M. E., Smith, R. A. G., Nowicki, B., Le Bouguénec, C., Lea, S. M. & Matthews, S. (2004). *Mol. Cell*, **15**, 647–657.

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. (2000). *Nucleic Acids Res.* **28**, 235–242.

Bork, P., Holm, L. & Sander, C. (1994). *J. Mol. Biol.* **242**, 309–320.

Choudhury, D., Thompson, A., Stojanoff, V., Langermann, S., Pinkner, J., Hultgren, S. J. & Knight, S. D. (1999). *Science*, **285**, 1061–1066.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Dauter, Z., Dauter, M. & Rajashankar, K. R. (2000). *Acta Cryst.* D**56**, 232–237.

Dauter, Z., Li, M. & Wlodawer, A. (2001). *Acta Cryst.* D**57**, 239–249.

Declercq, J. P. & Evrard, C. (2001). *Acta Cryst.* D**57**, 1829–1835.

Engh, R. A. & Huber, R. (1991). *Acta Cryst.* A**47**, 392–400.

Halaby, D. M., Poupon, A. & Mornon, J.-P. (1999). *Protein Eng.* **12**, 563–571.

Hung, C. S., Bouckaert, J., Hung, D., Pinker, J., Widberg, C., DeFusco, A., Auguste, C. G., Strouse, R., Langermann, S., Waksman, G. & Hultgren, S. J. (2002). *Mol. Microbiol.* **44**, 903–915.

Lamzin, V. S. & Wilson, K. S. (1997). *Methods Enzymol.* **277**, 269–305.

Laskowski, R. A., MacArthur, R. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.

Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.

Nishiyama, M., Horst, R., Eidam, O., Herrmann, T., Ignatov, O., Vetsch, M., Bettendorff, P., Jelesarov, I., Grütter, M. G., Wüttrich, K., Glockshuber, R. & Capitani, G. (2005). *EMBO J.* **24**, 2075–2086.

Nowicki, B., Svanborg-Edèn, C., Hull, R. & Hull, S. (1989). *Infect. Immun.* **57**, 446–451.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Perrakis, A., Morris, R. & Lamzin, V. S. (1999). *Nature Struct. Biol.* **6**, 458–463.

Pettigrew, D., Anderson, K. L., Billington, J., Cota, E., Simpson, P., Urvil, P., Rabuzin, F., Roversi, P., Nowicki, B., du Merle, L., Le Bouguénec, C., Matthews, S. & Lea, S. M. (2004). *J. Biol. Chem.* **279**, 46851–46857.

Piątek, R., Zalewska, B., Kolaj, O., Ferens, M., Nowicki, B. & Kur, J. (2005). *Infect. Immun.* **73**, 135–145.
Ponnuraj, K., Bowden, M. G., Davis, S., Gurusiddappa, S., Moore, D., Choe, D., Xu, Y., Hook, M. & Narayana, S. V. (2003). *Cell*, **115**, 217–228.
Sauer, F. G., Fütterer, K., Pinkner, J. S., Dodson, K. W., Hultgren, S. J. & Waksman, G. (1999). *Science*, **285**, 1058–1061.
Sauer, F. G., Pinkner, J. S., Waksman, G. & Hultgren, S. J. (2002). *Cell*, **111**, 543–551.
Sauer, F. G., Remaut, H., Hultgren, S. J. & Waksman, G. (2004). *Biochim. Biophys. Acta*, **1694**, 259–267.
Schneider, T. R. & Sheldrick, G. M. (2002). *Acta Cryst.* D**58**, 1772–1779.
Sheldrick, G. M. (2005). *XPREP* Program v.2005/1. Bruker–Nonius Inc., Madison, Wisconsin, USA.
Sheldrick, G. M. (2002). *Z. Kristallogr.* **217**, 644–650.
Soto, G. E. & Hultgren, S. J. (1999). *J. Bacteriol.* **181**, 1059–1071.
Van Loy, C. P., Sokurenko, E. V. & Moseley, S. L. (2002). *Infect. Immun.* **70**, 1694–1702.
Zalewska, B., Piatek, R., Bury, K., Samet, A., Nowicki, B., Nowicki, S. & Kur, J. (2005). *Microbiology*, **151**, 2477–2486.
Zavialov, A. V., Berglund, J., Pudney, A. F., Fooks, L. J., Ibrahim, T. M., MacIntyre, S. & Knight, S. D. (2003). *Cell*, **113**, 587–595.
Zavialov, A. V., Tischenko, V. M., Fooks, L. J., Brandsdal, B. O., Åqvist, J., Zav'yalov, V. P., MacIntyre, S. & Knight, S. D. (2005). *Biochem. J.* **389**, 685–694.