doi:10.1016/j.jmb.2006.07.069 J. Mol. Biol. (2006) 363, 125–136

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

High Resolution Structure of BipD: An Invasion Protein Associated with the Type III Secretion System of *Burkholderia Pseudomallei*

P. T. Erskine[1], M. J. Knight[1], A. Ruaux[1], H. Mikolajek[1] N. Wong Fat Sang[1], J. Withers[1], R. Gill[1], S. P. Wood[1], M. Wood[2] G. C. Fox[3] and J. B. Cooper[1]*

[1]*School of Biological Sciences University of Southampton Bassett Crescent East Southampton, SO16 7PX, UK*

[2]*Institute of Animal Health Division of Environmental Microbiology, Institute for Animal Health, Compton Laboratory, Berks, RG20 7NN, UK*

[3]*BM16, European Synchrotron Radiation Facility, BP 220 38043 Grenoble, France*

**Corresponding author*

Burkoldheria pseudomallei is a Gram-negative bacterium that possesses a protein secretion system similar to those found in Salmonella and Shigella. Recent work has indicated that the protein encoded by the *BipD* gene of *B. pseudomallei* is an important secreted virulence factor. BipD is similar in sequence to IpaD from Shigella and SipD from Salmonella and is therefore likely to be a translocator protein in the type-III secretion system of *B. pseudomallei*. The crystal structure of BipD has been solved at a resolution of 2.1 Å revealing the detailed tertiary fold of the molecule. The overall structure is appreciably extended and consists of a bundle of antiparallel α-helical segments with two small β-sheet regions. The longest helices of the molecule form a four-helix bundle and most of the remaining secondary structure elements (three helices and two three-stranded β-sheets) are formed by the region linking the last two helices of the four-helix bundle. The structure suggests that the biologically active form of the molecule may be a dimer formed by contacts involving the C-terminal α-helix, which is the most strongly conserved part of the protein. Comparison of the structure of BipD with immunological and other data for IpaD indicates that the C-terminal α-helix is also involved in contacts with other proteins that form the translocon.

Keywords: type-III secretion system; invasion protein; BipD; *Burkholderia pseudomallei*; X-ray structure

Introduction

Burkholderia pseudomallei and *Burkholdheria mallei* are intra-cellular bacteria that cause serious invasive infections in animals and are emerging as major causes of infectious disease in South East Asia.[1] These bacteria cause the disease melioidosis, which is endemic to tropical and subtropical regions particularly Southeast Asia and Northern Australia.[2] Most commonly this disease manifests itself clinically as abscesses, pneumonia and, at worst, as a fatal septicaemia, but has a complex spectrum of clinical manifestations making it difficult to diagnose and treat effectively.[3] *B. pseudomallei* is known to infect almost all types of cell and it appears able to evade, and perhaps actively interfere with, the host's immune system.[2] Awareness of *B. pseudomallei* has increased in recent years since being identified and labelled as a potential biological terrorist weapon.[3] This attention has now resulted in the completion of the genome sequence for *B. pseudomallei*.[4] Unlike most bacteria, the genome comprises two chromosomes of 4.07 megabase-pairs and 3.17 megabase-pairs, showing significant functional partitioning of genes between them. The large chromosome encodes many of the core functions associated with central metabolism and cell growth, whereas the small chromosome carries more accessory functions associated with adaptation and survival in different environments. At 7.24 Mb, the *B. pseudomallei* genome is large in comparison with the typical prokaryotic genome,

Abbreviations used: TTSS, type III protein secretion system; SeMet, selenomethionine.

E-mail address of the corresponding author: J.B.Cooper@soton.ac.uk

perhaps accounting for the bacterium's versatility and adaptability.[4]

Both *B. pseudomallei* and *B. mallei* are Gram-negative bacteria that possess a protein secretion apparatus that is highly similar to those found in Salmonella and Shigella.[5–7] A major function of these secretion systems is to secrete virulence-associated proteins into target cells of the host organism. Recently, it has been reported that disruption of the *bipD* gene, which is within a putative secretion system locus, reduces the ability of *B. pseudomallei* to invade eukaryotic cells and reduces virulence in mice.[5–7] The findings of this work indicate that the protein encoded by the *BipD* gene is an important secreted virulence factor of *B. pseudomallei*.

BipD is similar in sequence and most likely functionally analogous to IpaD from Shigella (26% identity; 36% similarity) and SipD from Salmonella (27% identity; 39% similarity). Thus by analogy with these better studied organisms, the BipD protein is likely to be a component of a type III protein secretion system (TTSS) in *B. pseudomallei*. TTSSs are large assemblies of proteins that span the inner bacterial membrane, the periplasmic space, the peptidoglycan layer, the outer bacterial membrane, the extra-cellular space and the target cell membrane.[8] The function of bacterial protein secretion systems is to transport "effector" and other proteins across the bacterial inner membrane and the outer envelope in an ATP-dependent manner.[9] There are five systems (types I–V), which operate in different ways and vary in complexity. Several involve a hollow tube or needle (the injectisome) through which the secreted proteins travel.[8,10] The TTSS injectisome varies in length between 45 nm and 80 nm depending on bacterial species, is made by the polymerisation of a major subunit (BsaL in *B. pseudomallei*) and has a hollow interior of approximately 25 Å diameter. The ring-like assembly that spans the membrane of the host cell is referred to as the translocon. This is formed by the initial secretion of a small number of proteins into the extra-cellular environment as a result of contact between the bacterium and target cell.[11] The translocator proteins act to transport bacterial proteins across the plasma membrane into the host cell which are then able to subvert or inhibit the cell's processes for the benefit of the bacterium. It appears that the translocator proteins form a pore in the lipid membrane of the target cell through which the effector proteins are able to pass.[12]

It has been found that *B. pseudomallei* contains at least three loci encoding putative type III protein secretion systems, one of which shares homology with the TTSS of *Salmonella typhimurium* and *Shigella flexneri* and has been designated BSA (Burkholderia secretion apparatus).[5,13–15] The BSA effector proteins are termed Bop proteins and the translocators are termed Bip proteins, short for Burkholderia invasion protein. In Salmonella SipB, SipC and SipD have been shown to be required for the injection of effector proteins and the invasion of epithelial cells *in vitro*,[16] and likewise for IpaB, IpaC and IpaD from shigella.[17] Hence, it is thought that the *B. pseudomallei* homologues BipB, BipC and BipD perform a similar function.

Serum from melioidosis victims contains antibodies to BipD and the other putative translocator proteins. It has been shown that *B. pseudomallei* bacteria with mutated BipD show reduced virulence *in vivo* but are still able to cause fatal melioidosis.[5–7] The BipD mutant exhibited impaired invasion of HeLa cells and reduced intra-cellular survival in J774.2 murine macrophage-like cells and a marked reduction in actin tail formation.[5] There is evidence that BipD facilitates the escape of *B. pseudomallei* from endocytic vesicles during the initial infection of a cell and is also involved in the actin polymerisation that allows the subsequent escape of progeny bacteria into surrounding host cells.[7] Observations made during these experiments support the concept that a functional TTSS is required for full virulence of *B. pseudomallei*.

Deletion of the first 20 residues of the BipD homologue in Shigella (IpaD) prevents its secretion,[18] suggesting that these residues act as a secretion signal. Mutants of BipD (in Burkholderia), IpaD (in Shigella) and SipD (in Salmonella) confer the respective organism with reduced invasiveness.[5,16,18] IpaD is known to associate with a component of the Shigella translocon (IpaB) in the outer cell membrane,[17] indicating that BipD may likewise associate with BipB in the putative Burkholderia translocon. Two of the protein components of the shigella translocon (IpaB and IpaC) associate less efficiently with the erythrocyte membrane in the presence of IpaD mutants.[18] Thus, by analogy, mutants of BipD that give rise to reduced invasiveness are likely to affect the formation of the translocon in target cell membranes. Since mutation of BipD, SipD and IpaD causes an increase in the secretion of other TTSS effector proteins into the extra-cellular medium,[5,16,18] it is likely that they regulate secretion of the other effector proteins *via* the TTSS. Deletion of the IpaD secretion signal prevents its export and also results in an increase in secretion of the other effector proteins.[18] This suggests that BipD, like IpaD, needs to enter the secretion apparatus in order to regulate the secretion of effector proteins. In Salmonella, if either SipB, SipC or SipD are mutated, this causes an increase in secretion of the other two proteins.[19] However, the amount of each protein reaching the cytoplasm of the target cell decreases, indicating that all three proteins (SipB, SipC or SipD) are required for translocation. It has been suggested that proteins in the same class as BipD (i.e. IpaD and SipD) act as extra-cellular chaperones to help the hydrophobic translocators enter the target cell membrane and might even link the translocon pore with the secretion needle.[10] Since IpaD, as a complex with IpaB and IpaC, can bind the $\alpha_5\beta_1$ integrin, stimulating the phosphorylation of a focal adhesion kinase,[20] it is possible that BipD acts in a similar way to aid actin-mediated insertion of the bacterium into the host cell membrane.

Table 1. Refinement statistics for the BipD structure

R-factor (%)	22.6
R-free (%)	29.9
Number of reflections	28,488
rmsd bond lengths (Å)	0.004
rmsd 1–3 distances (Å)	0.017
rmsd bumps (Å)	0.014
rmsd chiral tetrahedra (Å)	0.024
rmsd planar groups (Å)	0.022
Mean *B*-factor for main chain atoms (Å^2)	43.0
Mean *B*-factor for side chain atoms (Å^2)	46.5
Mean *B*-factor for the whole chain (Å^2)	44.7

All data between 10.0 Å and 2.1 Å resolution with no σ(I) cutoff were used in the refinement except for 5% of the data which were reserved for the *R*-free set. Note that residues in the highly disordered loop between residues 112 and 128 were omitted from the average *B*-factor calculations.

BipD from *B. pseudomallei* consists of 310 amino acid residues and has a molecular mass of 33 kDa. Curiously BipD differs from IpaD and SipD by completely lacking cysteine residues, although IpaD and SipD only possess one cysteine each, in different positions. Here we report the crystal structure of BipD, which has been solved at a resolution of 2.1 Å, revealing the detailed tertiary fold of the molecule. The structure suggests that the biologically active form of the molecule may be a dimer and by comparison with immunological and other data for IpaD, indicates the putative regions of contact with other effector molecules that form the translocon.

Results

Quality of the structure

The final crystal structure consists of two molecules of BipD and 681 water molecules in the asymmetric unit. The model, which has been refined using data between 10.0 Å and 2.1 Å, has an *R*-factor of 22.6% and an *R*-free of 29.9% (Table 1). The final structure has 88.8% of its residues within the "most favoured" regions of the Ramachandran plot by the PROCHECK criteria[21] and 99.6% of residues lie within the so-called "generously allowed" boundary. The remaining 0.4% of residues are in parts of the molecule with poor electron density and hence their conformations are poorly defined. The isotropic temperature factors for the protein are reasonable except for some residues at the N and C termini of both molecules in the asymmetric unit and the loop region between residues 112 and 128. In these regions the electron density remains poor despite extensive efforts to rebuild them throughout the refinement process. However approximately 82% of the molecule is extremely well defined by electron density map; the disorder in the remaining regions may have functional significance and may also account for the slightly high *R*-free obtained in refinement. Both molecules of BipD within the asymmetric unit superimpose closely with an rms deviation of 0.6 Å for 240 C^{α} atoms in the structurally equivalent regions of both monomers.

Tertiary structure

The overall structure is appreciably extended and mainly consists of a bundle of antiparallel α-helical segments with two three-stranded β-sheet regions, one being at one end of the bundle and the other being on the side (Figure 1). The overall topology of the molecule is shown in Figure 2. The crystal structure is consistent with far-UV circular dichroism (CD) spectra for BipD (Figure 3), analysis of which confirmed the marked predominance of α-helix over β-sheet in the tertiary structure. The secondary structure content of the molecule as determined by X-ray analysis (59% α-helix and 19% β-sheet) agrees reasonably well with the

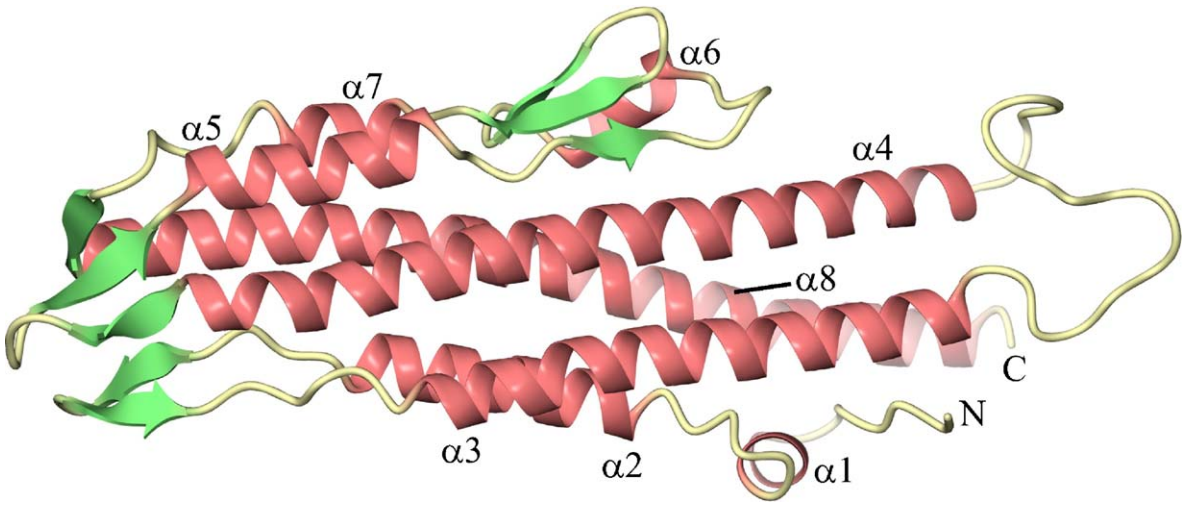

Figure 1. The tertiary structure of BipD determined at 2.1 Å resolution. The helices are shown in pink and strands in green. The α-helices are numbered according to the topology shown in Figure 2. The longest helical segment (α8) is at the rear of the molecule in this view. The strong sequence conservation of this helix, which is at the C-terminal end of the molecule, suggests that it plays an important role in protein–protein interactions in the translocon.

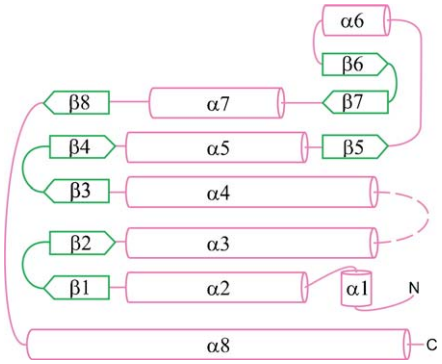

Figure 2. The overall topology of the BipD molecule, with α-helical segment shown in pink and β-strand regions shown in green. The loop region linking helices α3 and α4 is shown dashed since this region is poorly defined by the electron density map. Note that the β-hairpin formed by strands β1 and β2 is not part of the β-sheet formed by strands β3, β4 and β8, although all of these strands are at the same end of this elongated molecule.

estimates of 66% helix and 4% sheet from the CD data. All of the helical segments in the structure are amphiphilic and associate together with a considerable buried hydrophobic interface. The quality of the electron density is exemplified by the region shown in Figure 4, which is invariant in BipD, IpaD and SipD and lies within the strongly conserved C-terminal region of the molecule.

The first 20–30 residues of the protein were not clearly defined in the electron density map but in one subunit within the asymmetric unit, the residues 36–43 were clearly seen to form a short α-helix (helix 1). In the other molecule in the asymmetric unit, this region is less well defined and adopts a slightly different conformation (Figure 5). The polypeptide then forms a longer helical region (helix 2) from residue 47 to residue 63, which is followed by a somewhat irregular β-hairpin involving residues 64–81 (β1 and β2 in Figure 2). The hairpin region is followed by another α-helix, which extends from residue 82 to residue 111 (helix 3) running antiparallel to helix 2. Most of the next 20 residues were found to be poorly defined in the electron density map, indicating that this region is rather disordered in the crystal. The lack of definition in this region led us to suspect proteolysis had occurred during crystallisation but SDS-PAGE analysis of protein from dissolved crystals confirmed that the molecule was of the correct length. Electron density returns again for residue 128, which is at the N-terminal end of a very long helix extending for almost the full length of the molecule and ending at residue 170 (helix 4). This helix runs antiparallel to helix 3 and leads into a β-hairpin formed by residues 171 to 183 (β3 and β4 in Figure 2). This is followed by a short region of helix formed by residues 184 to 196 (helix 5), which runs anti-parallel to helix 4. Residues 197 to 203 form an extended region (β5) of which the last three residues take part in a three-stranded β-sheet involving a β-hairpin formed by residues 220–230 (β6 and β7). There is a short region of helix formed by residues 209 to 216 (helix 6), which connects the first strand of this sheet (β5) with the start of the hairpin region (β6 and β7). This topology of this region of the molecule has some parallels with a Greek key motif with one of the outer β-strands replaced by a helix (helix 6).

The following residues 233 to 241 then form a short helix (helix 7) that runs antiparallel to helix 5 and leads into a short β-strand region formed by residues 246 to 250 (β8). This strand forms a three-stranded β-sheet with the β-hairpin formed by residues 171–183 (β3 and β4). The final α-helix of the molecule (helix 8) extends for 50 residues from 251 to 301, running the whole length of the molecule. This helix is a remarkably conserved part of the protein family (Figure 6), with sequence identities of 60% with IpaD and 52% with SipD, suggesting that it plays an important role in the function of the secretion apparatus. Whilst use of the DALI[22] and MSDfold[23] servers did reveal structural similarity between BipD and a number of other four-helix bundle proteins, none of these possessed the same β-sheet regions. Hence we believe this structure to be a novel elaboration of the four-helix bundle fold.

Quaternary structure

Gel filtration of purified BipD (~3 mg/ml) on a calibrated Superdex-75 column at pH 7.5 established that it has a V_e/V_o ratio of 1.17, which corresponds to monomeric protein. However, the two BipD molecules in the crystallographic asymmetric unit associate by forming extensive contacts involving the C-terminal end of helix 8 and the N-terminal end of helix 4 from both subunits (Figure 7(a) and (b)).

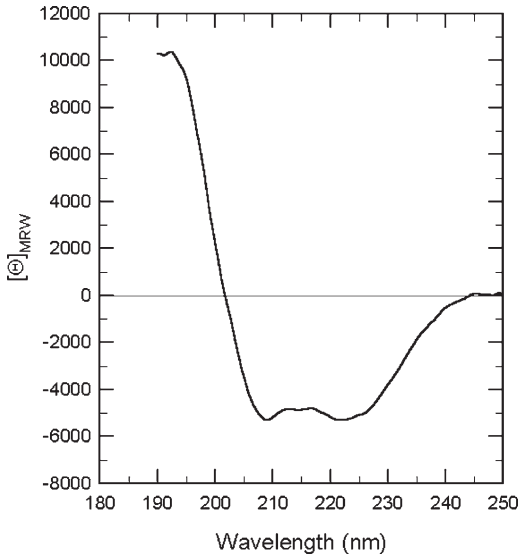

Figure 3. The far-UV CD spectrum of BipD confirming the preponderance of α-helix in the structure. The protein was diluted to a concentration of 0.15 mg/ml in 20 mM phosphate buffer (pH 7.6) to record the spectrum.

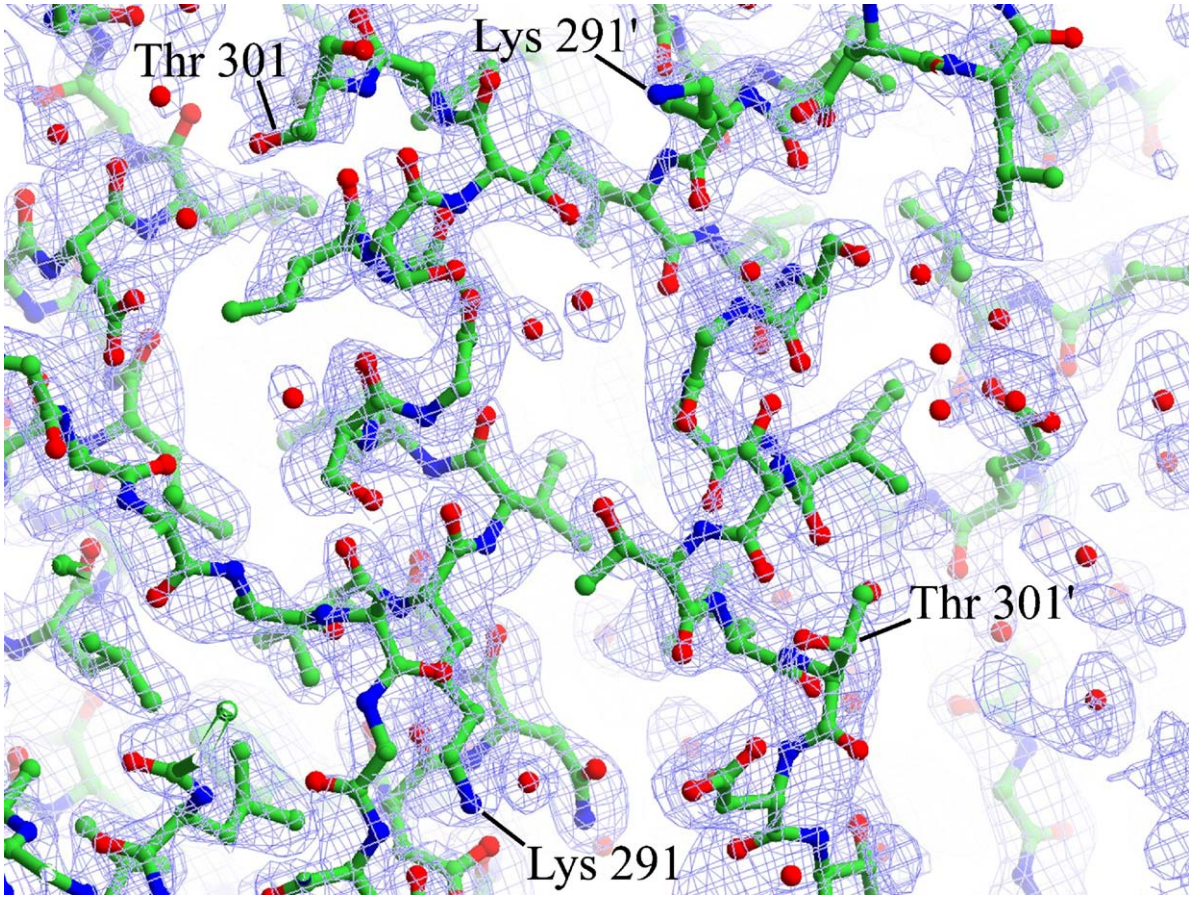

Figure 4. The electron density map for part of helix 8, which interacts in an antiparallel manner with the same helix from the second monomer in the asymmetric unit. Both helices in adjacent monomers are shown running vertically in opposite directions. The part of this helix involved in these contacts is the most strongly conserved region of BipD, IpaD and SipD. The electron density is contoured at 1.0 rms.

The area of contact between the two monomers involves helix 8 pairing up with the same helix from the other monomer in an antiparallel fashion (Figures 4 and 7(b)), and likewise for helix 4. Whilst gel-filtration of BipD indicated that BipD was monomeric in solution, the fact that the region of

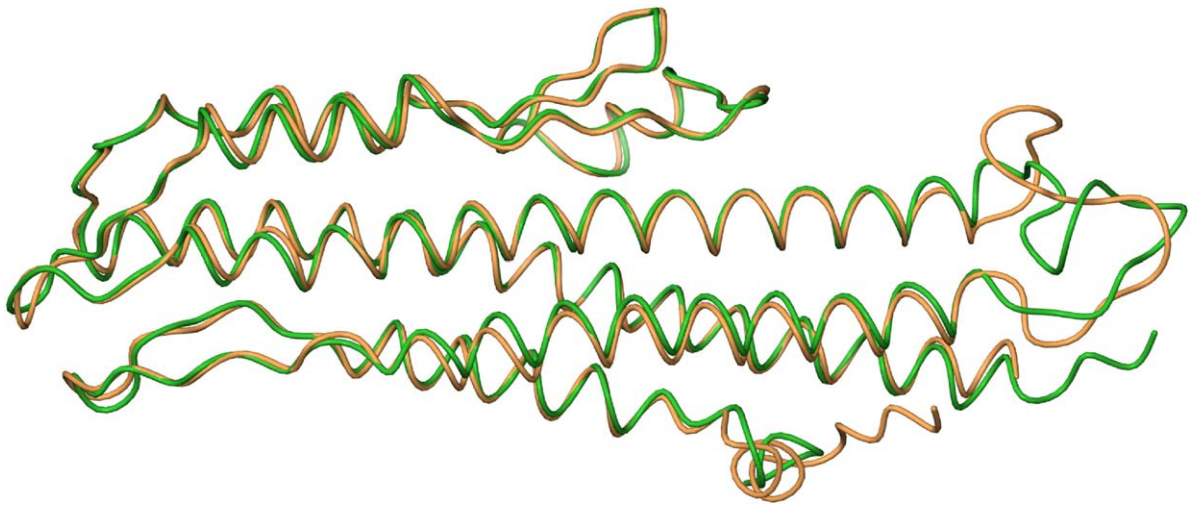

Figure 5. A superposition of the two molecules of BipD in the crystallographic asymmetric unit viewed from the same direction as Figure 1. The structurally conserved regions of both monomers superimpose with an rms C^{α} deviation of 0.6 Å emphasising the similarity of their conformations. The only regions where the structures diverge (N and C termini and the loop between helices 3 and 4) are poorly defined by the electron density.

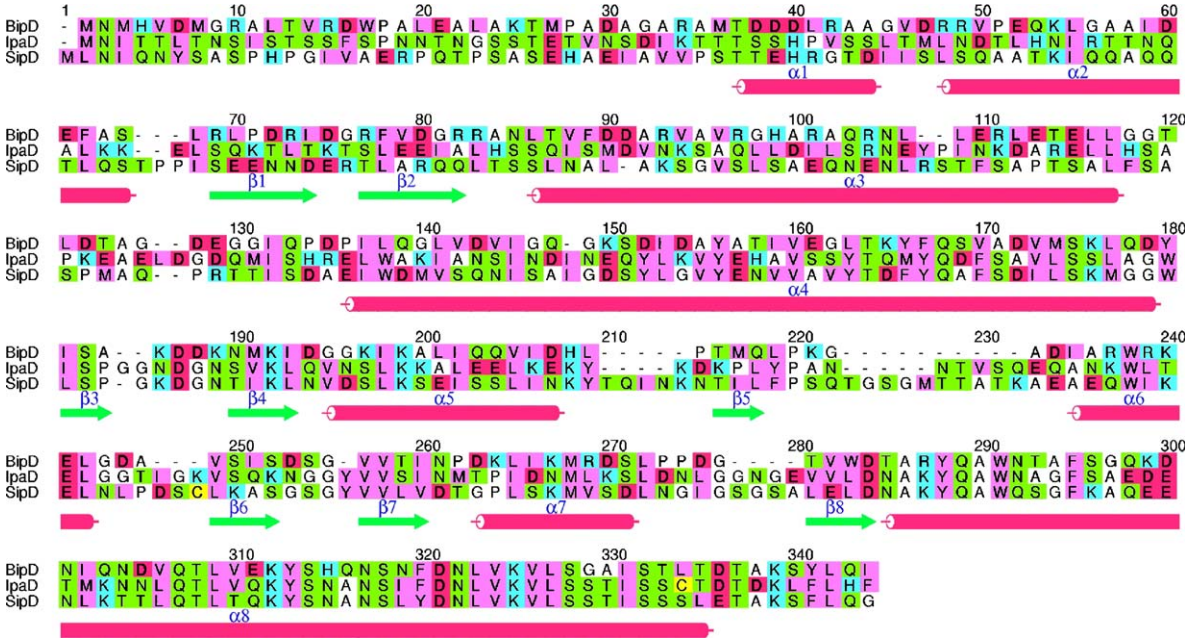

Figure 6. A sequence alignment of BipD with IpaD from *S. flexneri* and SipD from *S. typhimurium*. The BipD sequence has 26% identity and 36% similarity with IpaD and 27% identity and 39% similarity with SipD. The amino acids are coloured according to the scheme: cyan, basic; red, acidic; green, neutral-polar; pink, bulky hydrophobic; white, Gly, Ala and Pro; and yellow, Cys. The secondary structure elements of BipD (numbered according to the scheme in Figures 1 and 2) are indicated below the alignment.

helix 8 forming the interaction is the most highly conserved region of the molecule, suggests that the dimer may be biologically relevant, possibly forming at higher protein concentrations. Both monomers lose 970 Å^2 of solvent accessible area on formation of the dimer, which represents about 7% of the total solvent accessible area of each monomer. The residues of helix 8 taking part in this contact region extend from Asn288 to Leu300 and comprise approximately three turns of the helix. They form an interface dominated by van der Waals contacts between aliphatic side-chains although there are some water-mediated contacts too. Most of the amino acid residues in this region are conserved in BipD, IpaD and SipD and eight consecutive residues from 287 to 294 (sequence: DNLVKVLS) are completely invariant in these three proteins. These findings, combined with the known sensitivity of the molecule to deletions in this region,[18] suggest that the dimer observed in the X-ray structure may be a functionally important form of the molecule, although the interface lacks specific direct interactions such as hydrogen bonds or salt-bridges between the monomers. There are other stretches of invariant sequence in helix 8 that are not involved in dimer formation. Hence it is conceivable that these regions are involved in binding partner proteins in the translocon. In the dimer, the quaternary contacts involving helix 4 from each BipD monomer are of a similar nature (i.e. involve small hydrophobic or neutral polar side-chains) and extend over approximately four turns of the helix between Pro128 and Asp143. The sequence conservation in this region is poorer than for helix 8 and likewise the interface between helix 4 from each monomer lacks specific hydrogen bonds and salt-bridges.

Discussion

The core four-helix bundle of the molecule is formed by the longest helices 2, 3, 4 and 8. The region formed by residues 171–250 (i.e. helices 5, 6 and 7 and both of the three-stranded β-sheets) effectively is an extensive elaboration of the loop linking helices 4 and 8. It is known that chaperones are required to assist the passage of proteins through the secretion needle probably by partially or fully unfolding them to allow them to enter the needle.[8] However, the elongated shape of the BipD molecule indicates that it may be able to pass through the pore of the needle with minimal unfolding. If the region formed by residues 171–250 were to unfold or swing away from the four-helix bundle, this would reduce the diameter of the molecule to approximately 25 Å, which matches the inner diameter of the needle. Accordingly, epitope mapping studies of IpaD[24] indicate that residues in the 171–250 region of BipD could have considerable flexibility. An alternative possibility is that the N-terminal 130 residues, which are poorly conserved, may instead unfold or swing away to allow the remainder of the molecule to fit into the needle. The advantage of this hypothesis is that the N-terminal 130 residues are connected to the remainder of the molecule by a very flexible linker region, although the interface between the helices is extremely hydrophobic.

In studies of Shigella effector proteins it was found that whilst IpaB and IpaC readily associate with membranes, IpaD (the BipD homologue) was only able to associate with liposomes at high concentrations.[25] However lowering the pH below 5.0 was found to increase the binding of IpaD to liposomes. Whilst pH-dependent denaturation of any soluble protein might be expected to expose hydrophobic regions and increase its binding to membranes, it may be that the effect observed with IpaD is of importance for the infectivity of Shigella, particularly in view of the pH drop that occurs in the endosomes. There may be parallels with hemagglutinin, which is known to undergo a pH-induced conformational change exposing a fusigenic peptide that facilitates entry of the influenza virus into the target cell.[26]

All of the secondary structure elements in BipD have a high density of acidic and basic residues (Figure 7(c) and (d)), suggesting that as the pH is lowered in the endosome, protonation of carboxylate groups may cause repulsion of the positively charged basic groups and hence conformational change. This is particularly true for the smallest exposed helices namely helices 1, 6 and 7, which have a very high abundance of acidic and basic residues. Perhaps these regions are the most vulnerable to pH-induced unfolding in the endosome, which could trigger association of the protein with the membrane. In contrast the longest and most conserved helix (helix 8) has by far the lowest density of acidic or basic residues. The pH of crystallisation of BipD is approximately 5.8 (following mixing of the protein in 50 mM Tris (pH 8.0) with the well solution of 0.1 M cacodylate (pH 5.0–5.5)) suggesting that the structure observed in our X-ray analysis may be very close to that which occurs in the endocytic vesicle as the pH is lowered.

It has been suggested that IpaD acts as a plug in the translocon pore that regulates secretion of the other Shigella effector proteins, most notably IpaB and IpaC, which form the translocon complex and trigger host cell responses. In contrast IpaD does not appear to trigger host cell responses itself. Deletion of the first 20 residues of IpaD prevents its secretion, suggesting that these residues act as a secretion signal.[18] Unfortunately this region of BipD was not defined by the electron density map. Removal of residues between 20 and 120 of IpaD reduced contact-mediated haemolysis of red blood cells but had no effect on invasion.[18] Since deletions beyond residue 120 eliminate both haemolysis and invasiveness this suggested that the C-terminal region of the molecule is most important for its function. All deletions (except of the first 20 residues) resulted in an increase in IpaD secretion. Deletion of just a few residues at the C terminus also eliminated both haemolysis and invasiveness and also results in even greater levels of IpaD secretion, suggesting that this region is also involved in transport regulation. The equivalent part of BipD was not defined in the electron density map in spite of it being a strongly conserved region. However we have shown that the immediately preceding region (the C-terminal end of helix 8) is involved in dimer formation *via* residues that show an extraordinary level of sequence conservation in BipD, IpaD and SipD. The extreme C-terminal region of the protein may only become ordered in the presence of the partner proteins BipB and BipC. There is evidence that IpaD regulates the relative rates of IpaB and IpaC transport and the reduction in haemolytic activity of IpaD mutants correlates with the reduced incorporation of IpaB and IpaC into erythrocyte membranes.[18] Thus BipD is likely to play a similar role in regulating transport of BipB and BipC as well as their final integration into the target cell membrane.

IpaD is weakly associated with the surface of Shigella. Use of a series of overlapping synthetic peptides mimicking the entire length of IpaD to screen for reactive antibodies in serum produced against purified IpaD[24] allowed the main epitopes of the protein to be defined. Use of an affinity column with virulent *S. flexneri* bound to it allowed those antibodies that react with epitopes exposed on the surface of the bacteria to be purified. Use of these antibodies to screen the overlapping peptides allowed the epitopes of IpaD that are exposed on the surface of *S. flexneri* to be defined. These studies showed that all IpaD epitopes exposed on the surface of the bacteria are in the N-terminal 120 residues of the protein, which is a region of poor sequence conservation and the most tolerant to deletions.[18] In all cases the epitopes correspond to exposed regions of the BipD molecule e.g. helix 1, the C-terminal end of helix 2 and the β-hairpin linking helices 2 and 3. The dominant epitopes correspond to residues 1–30 and residues 105–120 of IpaD, which are poorly defined in the electron density map for BipD, presumably due to their flexibility. The C-terminal region of the IpaD protein beyond residue 120 is inaccessible to antibodies in intact Shigella and it is this region that is the least tolerant to deletions.[18] Nevertheless the epitopes of purified IpaD that have been mapped in this region of the molecule include residues in the loop preceding helix 6 of BipD, the following β-hairpin and the extreme C-terminal end of the molecule, which is poorly defined by the electron density map. However, by far the most dominant linear epitopes in the C-terminal domain are in the vicinity of helix 7, which is part of the region that was discussed above in terms of being an appendage of the basic four-helix bundle fold. This region clearly has considerable flexibility and is very accessible to antibodies *in vitro*. However, in intact Shigella, the amino acids beyond residue 120 of IpaD are not accessible to antibodies. This region of the molecule has been implicated in the haemolytic, invasion and transport regulation activities of the protein and is the most strongly conserved region of IpaD, SipD and BipD. This suggests that this region of BipD may associate with other effector

proteins (probably BipB) to regulate the secretion of effector proteins by the TTSS. Our analysis also indicates that residues in the functionally important domain of BipD may be involved in dimer formation.

Based on the results of our structural investigation reported here, we are proceeding with mutagenesis of BipD to define the roles of conserved and structurally important residues identified thus far. In addition to further studies of the oligomerisation state of BipD, we will analyse pH-dependent conformational changes in the wild-type and mutant BipD protein by use of CD spectroscopy. In parallel with ultra-structural studies of the Burkholderia type III secretion apparatus and other detailed studies of its protein components (e.g. the needle

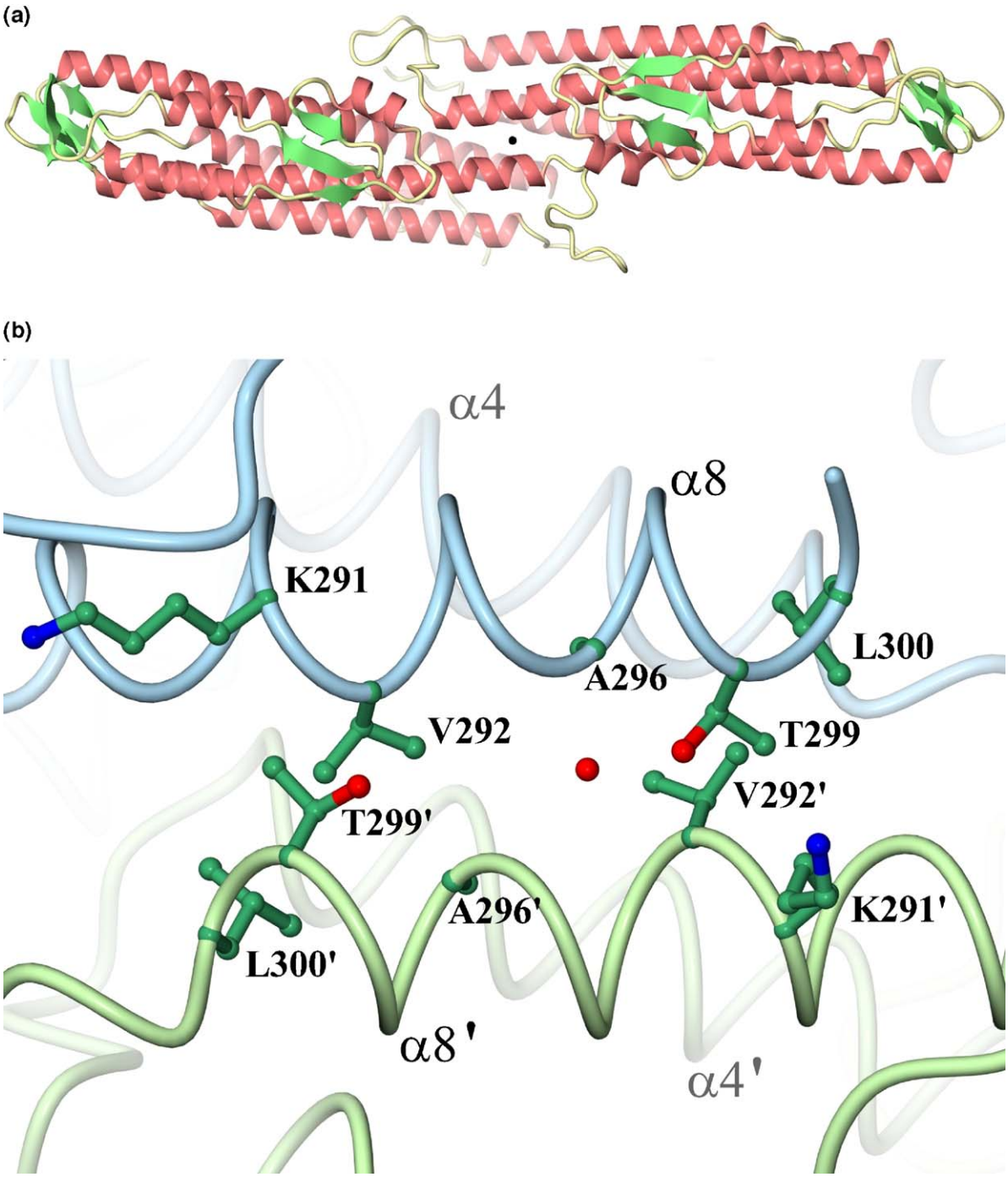

Figure 7. (a) The two monomers of BipD within the crystallographic asymmetric unit. The central black sphere indicates where both monomers associate with approximate 2-fold symmetry by contacts involving helices 4 (front) and 8 (back). (b) Residues of helix 8 forming the most extensive contacts between the two BipD monomers (coloured pale blue and green); helix 4 can be seen in the background. Residues at the C-terminal end of helix 8 are strongly conserved in BipD, IpaD and SipD, suggesting that the dimer may be biologically relevant. (c) The electrostatic surface of the dimer viewed from the same direction as (a). (d) The surface viewed from the opposite direction.

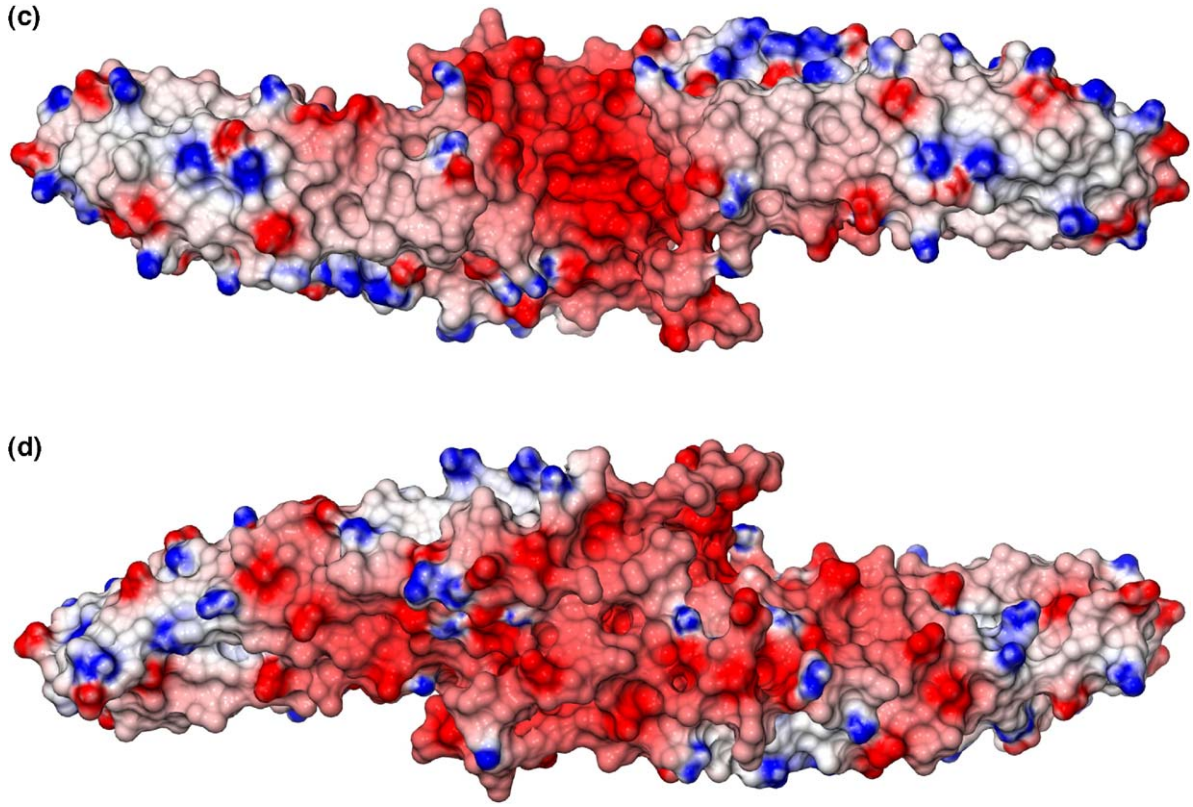

Figure 7 (*legend on previous page*)

protein BsaL recently analysed by NMR[27]), the structure reported here will contribute towards a more detailed understanding of this remarkable molecular machine.

Methods and Materials

Over-expression, purification and crystallisation of selenomethionyl BipD

Expression involved use of a pGEX-4T (Pharmacia) plasmid construct encoding a fusion of BipD with glutathione-*S*-transferase (GST) to facilitate expression in *Escherichia coli* BL21 (DE3) cells and purification by use of a GSTrap affinity column (GE Healthcare). Due to the cloning strategy, the first residue of the expressed protein corresponds to amino acid number 8 in the gene sequence of BipD. CD spectra of native BipD were recorded using a JASCO J720 spectrophotometer and the data were analysed using the DICHROWEB software.[28] Crystals of native BipD were obtained using the hanging-drop method with 5 μl of protein being mixed with 5 μl of each well solution on siliconised glass cover slips. The optimum well solution for crystallisation consisted of 25% (w/v) PEG 4000, 5 mM nickel chloride, 0.1 M glycine, 0.1 M sodium cacodylate (pH 6.5). Crystals appeared after approximately nine weeks at either 4 °C or 22 °C.

Initial attempts to express selenomethionyl-BipD (SeMet-BipD) involved growth of the cells on the 2 l scale to mid-log phase in LB medium (including 50 μg/ml ampicillin) followed by centrifugation to pellet them. The cells were re-suspended in M9 minimal medium (containing ampicillin at the same concentration and 0.4% (w/v) glucose) and grown for another 45 min to use up remaining amino acids. The cultures were then inoculated with vitamins: riboflavin, niacinamide, pyridoxine and thiamine each at final concentrations of 1 mg/l and a mixture of all 20 amino acids, except for methionine, which was replaced by selenomethionine, each at final concentrations of 40 mg/l. The cells were then shaken at 37 °C for another 20 min prior to overnight induction with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG). However, whilst with trial cultures, SDS-PAGE had shown that all of the native protein expressed by the construct is soluble, we found that virtually all of the SeMet-BipD was insoluble.

At this point small trial cultures were used to experiment with minimising the formation of inclusion bodies. This involved giving the cells a 25 min heat shock at 42 °C followed by cooling on ice for 5 min prior to induction with a low concentration of IPTG (0.3 mM). To adapt this method for expression of SeMet protein we followed the protocol for selenomethionine incorporation up to the point where the cells would have otherwise been induced and instead they were subjected to the above heat-shock, cooling and induction. The cells were then grown at a low temperature (16 °C) for two days and, as a precaution against oxidation of the SeMet-BipD, dithiothreitol (DTT) was included in the buffer used for subsequent sonication of the cells. SDS-PAGE gels showed that most of the protein was present in the soluble fraction, thus confirming the success of the method. The SeMet-protein was then expressed using 2×1 l cultures for purification of larger quantities for crystallisation trials.

A crude initial purification of the SeMet-BipD by fractional ammonium sulphate precipitation was undertaken in which the ammonium sulphate concentration was raised to 30% saturation to precipitate impurities and

Table 2. Data collection and processing statistics for selenomethionyl-BipD from *Burkholderia pseudomallei*

Dataset	Peak	Remote	Inflection	SAD 2.1 Å dataset
ESRF beam line	BM16	BM16	BM16	ID23-1
λ (Å)	0.9794	0.9077	0.9796	0.9794
Resolution (Å)	40.0–3.0 (3.2–3.0)	40.0–3.0 (3.2–3.0)	40.0–3.0 (3.2–3.0)	38.7–2.1 (2.2–2.1)
R_{merge} (%)	8.5 (16.2)	8.3 (17.8)	8.4 (21.2)	8.6 (41.1)
Completeness (%)	99.4 (100.0)	99.6 (100.0)	99.7 (100.0)	97.5 (93.4)
Average $I/\sigma(I)$	25.0 (11.3)	26.7 (12.1)	26.9 (11.1)	10.4 (2.8)
Multiplicity	9.9 (10.1)	9.8 (10.0)	9.8 (10.0)	3.7 (3.5)

Values for the outer resolution shell of each dataset are shown in parentheses.
$R_{merge}=\sum_h | (I_{hi}-I_h) | / \sum_i (I_{hi})$ where I_h is the mean intensity of the scaled observations I_{hi}.

then raised to 60% saturation to precipitate the fusion protein. The re-suspended pellet was then dialysed against 50 mM Tris buffer (pH 8.0) (1 mM DTT) prior to loading onto a GSTrap FF 5 ml affinity column *via* an AKTA prime low pressure system. The column was washed with phosphate buffered saline (PBS) (pH 8.0) and 1 mM DTT to remove impurities and the SeMet-BipD was cleaved from its fusion partner (GST) by addition of 1.6 mg of thrombin to the column and incubation for 20 h. The column was then washed with PBS buffer to remove the SeMet-BipD and was regenerated by washing with reduced glutathione to remove the GST fusion partner. The thrombin was removed from the SeMet-BipD by passage through a 1 ml benzamidine column (GE Healthcare). At all stages, the purity of the protein and success of the cleavage reaction was confirmed by SDS-PAGE. The final yield of SeMet-protein was approximately 30 mg/l of culture.

For crystallisation experiments, SeMet-BipD was concentrated to 6 mg/ml using a Vivaspin ultrafiltration spin column. Screening conditions were based on those used previously to obtain crystals of the native enzyme and were set up by the hanging-drop method. Within two to three weeks, promising crystals of SeMet-BipD were obtained and following optimisation of the crystallisation conditions, significantly improved crystals were obtained under the following conditions: [SeMet-BipD] = 4 mg/ml, 20–25% PEG 4000, 5 mM nickel chloride, 0.1 M glycine, 0.1 M sodium cacodylate (pH 5.0–5.5). Crystals of up to 1 mm in length grew within a period of four weeks and were cryo-protected with 30% (v/v) glycerol and mounted in mohair loops or Litholoops (Molecular Dimensions) for freezing in liquid ethane and storage under liquid nitrogen.

X-ray diffraction analysis of BipD

The selenomethionyl-BipD crystals were taken to the European Synchrotron Radiation Facility (ESRF, Grenoble) where they were found to diffract to a resolution of 2.1 Å using the ID23-1 beamline. A dataset was collected at the selenium peak wavelength and was processed using MOSFLM,[29] SCALA[30] and other programs in the CCP4 program suite.[31] This indicated that the crystals belong to the monoclinic space group $P2_1$ and had cell dimensions of $a = 53.5$ Å, $b = 56.2$ Å, $c = 84.2$ Å, $\beta = 94.5°$. Statistics for the processed 2.1 Å resolution selenomethionyl-BipD SAD dataset are shown in Table 2. Use of the MC program in CCP4[31] suggested that the crystals have two monomers per asymmetric unit with a solvent content of 35%.

Since attempts at SAD phasing using the above data were unsuccessful, multi-wavelength anomalous dispersion (MAD) X-ray data were collected from another crystal using beam line BM16. Data to a resolution of 3.0 Å were collected at the selenium fluorescence peak wavelength (0.9794 Å), followed by a remote dataset (0.9077 Å) and lastly an inflection point dataset (0.9796 Å). The data were processed as above and the relevant statistics are shown in Table 2. Six out of the 12 expected selenium sites were located by use of the program SOLVE/RESOLVE,[32] which produced an excellent electron density map with extended regions of α-helix. Segments of polypeptide were fitted to the map using the automatic map-fitting program MAID[33] and were rebuilt using TURBO-FRODO (Biographics, Marseille) and WinCOOT.[34] Thus one monomer with the correct sequence of BipD was built into the map and a copy of this model was then fitted by visual inspection to the electron density for the second subunit in the asymmetric unit. A further three selenium sites were located from an anomalous difference Fourier and used in phasing. Model building was assisted by the use of the density modification program DM[35] to extend the experimental MAD phases to higher resolution. The model was refined using the 2.1 Å resolution dataset from the other crystal by use of the programs SHELX-97[36] and CNS,[37] interspersed with rounds of manual rebuilding and solvent fitting. The *R*-free reflection set was chosen using SHELXPRO[36] in thin resolution shells to avoid bias between the working and test sets due to the presence of non-crystallographic symmetry. Except for the initial round of refinement in which rigid body constraints were applied to both monomers in the asymmetric unit, non-crystallographic symmetry (NCS) restraints were not used. The refinement statistics of the final model are shown in Table 1. Figures showing molecular structures were prepared using CueMol†. The coloured sequence alignment was prepared using Malign[38] and Alscript.[39] The electrostatic surface potential of the protein was calculated using DELPHI[40] and displayed using CueMol.

Protein Data Bank accession numbers

The coordinates and structure factors have been deposited with the RCSB Protein Data Bank and assigned accession numbers: 2izp for the coordinates and r2izpsf for the structure factors.

Acknowledgements

We thank the Wellcome Trust for infrastructure support and a vacation scholarship to M.J.K. We

† http://www.cuemol.org/en/

thank A. Cossins and Professor M. G. Gore for CD measurements as well as Drs S. Findlow and A. Coker for general assistance. We are indebted to the ESRF (Grenoble) for beam time and travel support.

References

1. Dance, D. A. (2002). Melioidosis. *Curr. Opin. Infect. Dis.* **15**, 127–132.
2. Gan, Y.-H. (2005). Interaction between *Burkholderia pseudomallei* and the host immune response: sleeping with the enemy? *J. Infect. Dis.* **192**, 1845–1850.
3. Aldhous, P. (2005). Melioidosis? Never heard of it. *Nature*, **434**, 692–693.
4. Holden, M. T. G., Titball, R. W., Peacock, S. J., Cerdeno-Tarraga, A. M., Atkins, T., Crossman, L. C. *et al.* (2004). Genomic plasticity of the causative agent of melioidosis, *Burkholderia pseudomallei*. *Proc. Natl Acad. Sci. USA*, **101**, 14240–14245.
5. Stevens, M. P., Wood, M. W., Taylor, L. A., Monaghan, P., Hawes, P., Jones, P. W., Wallis, T. S. & Galyov, E. E. (2002). An Inv/Mxi-Spa-like type III protein secretion system in *Burkholderia pseudomallei* modulates intracellular behaviour of the pathogen. *Mol. Microbiol.* **46**, 649–659.
6. Stevens, M. P., Haque, A., Atkins, T., Hill, J., Wood, M. W., Easton, A. *et al.* (2004). Attenuated virulence and protective efficacy of a *Burkholderia pseudomallei* bsa type III secretion mutant in murine models of melioidosis. *Microbiology*, **150**, 2669–2676.
7. Steven, M. P. & Galyov, E. E. (2004). Exploitation of host cells by *Burkholderia pseudomallei*. *Int. J. Med. Microbiol.* **293**, 549–555.
8. Yip, C. K. & Strynadka, N. C. J. (2006). New structural insights into the bacterial type III secretion system. *TIBS*, **31**, 223–230.
9. Mecsas, J. & Strauss, E. J. (1996). Molecular mechanisms of bacterial virulence: type III secretion and pathogenicity islands. *Emerging Infect. Dis.* **2**, 271–288.
10. Mota, L. J., Sorg, I. & Cornelis, G. R. (2005). Type III secretion: the bacteria-eukaryotic cell express. *FEMS Microbiol. Letters*, **252**, 1–10.
11. Pettersson, J., Nordfelth, R., Dubinina, E., Bergman, T., Gustafsson, M., Magnusson, K. E. & Wolf-Watz, H. (1996). Modulation of virulence factor expression by pathogen target cell contact. *Science*, **273**, 1231–1233.
12. Blocker, A., Holden, D. & Cornelis, G. (2000). Type III secretion systems: what is the translocator and what is translocated? *Cell Microbiol.* **2**, 387–390.
13. Rainbow, L., Hart, C. A. & Winstanley, C. (2002). Distribution of type III secretion gene clusters in *Burkholderia pseudomallei*, *B. thailandensis* and *B. mallei*. *J. Med. Microbiol.* **51**, 374–384.
14. Attree, O. & Attree, I. (2001). A second type III secretion system in *Burkholderia pseudomallei*: who is the real culprit? *Microbiology*, **147**, 3197–3199.
15. Hueck, C. J. (1998). Type III protein secretion systems in bacterial pathogens of animals and plants. *Microbiol. Mol. Biol. Rev.* **62**, 379–433.
16. Kaniga, K., Trollinger, D. & Galan, J. E. (1995). Identification of 2 targets of the type-III protein secretion system encoded by the *inv* and *spa* loci of *Salmonella typhimurium* that have homology to the Shigella IpaD and IpaA proteins. *J. Bacteriol.* **177**, 7078–7085.
17. Menard, R., Sansonetti, P. & Parsot, C. (1994). The secretion of the *Shigella flexneri* Ipa invasins is activated by epithelial-cells and controlled by IpaB and IpaD. *EMBO J.* **13**, 5293–5302.
18. Picking, W. L., Nishioka, H., Hearn, P. D., Baxter, M. A., Harrington, A. T., Blocker, A. & Picking, W. D. (2005). IpaD of *Shigella flexneri* is independently required for regulation of Ipa protein secretion and efficient insertion of IpaB and IpaC into host membranes. *Infect. Immun.* **73**, 1432–1440.
19. Collazo, C. M. & Galan, J. E. (1997). The invasion-associated type-III system of *Salmonella typhimurium* directs the translocation of Sip proteins into the host cell. *Mol. Microbiol.* **24**, 747–756.
20. Watarai, M., Funato, S. & Sasakawa, C. (1996). Interaction of Ipa proteins of *Shigella flexneri* with $\alpha_5\beta_1$ integrin promotes entry of the bacteria into mammalian cells. *J. Exp. Med.* **183**, 991–999.
21. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallog.* **26**, 283–291.
22. Holm, L. & Sander, C. (1996). Mapping the protein universe. *Science*, **273**, 595–603.
23. Krissinel, E. & Henrick, K. (2004). Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallog. sect. D*, **60**, 2256–2268.
24. Turbyfill, K. R., Mertz, M. A., Mallett, C. P. & Oaks, E. V. (1998). Identification of epitope and surface-exposed domains of *Shigella flexneri* invasion plasmid antigen D (IpaD). *Infect. Immun.* **66**, 1999–2006.
25. De Geyter, C., Wattiez, R., Sansonetti, P., Falmagne, P., Ruysschaert, J. M., Parsot, C. & Cabiaux, V. (2002). Characterisation of the interaction of IpaB and IpaD, proteins required for entry of *Shigella flexneri* into epithelial cells, with a lipid membrane. *Eur. J. Biochem.* **267**, 5769–5776.
26. Bullough, P. A., Hughson, F. M., Skehel, J. J. & Wiley, D. C. (1994). Structure of influenza hemagglutinin at the pH of membrane-fusion. *Nature*, **371**, 37–43.
27. Zhang, L., Wang, Y., Picking, W. L., Picking, W. D. & De Guzman, R. N. (2006). Solution structure of monomeric BsaL, the type III secretion needle protein of *Burkholderia pseudomallei*. *J. Mol. Biol.* **359**, 322–330.
28. Lobley, A. & Wallace, B. A. (2001). Dichroweb: a website for analysis of protein secondary structure from circular dichroism spectra. *Biophys. J.* **80**, 373a.
29. Leslie, A. G. W. (2006). The integration of macromolecular diffraction data. *Acta Crystallog. sect. D*, **62**, 48–57.
30. Evans, P. R. (2006). Scaling and assessment of data quality. *Acta Crystallog. sect. D*, **62**, 72–82.
31. CCP4 (1994). The CCP4 Suite: programs for protein crystallography. *Acta Crystallog. sect. D*, **50**, 760–763.
32. Terwilliger, T. C. & Berendzen, J. (1999). Automated MAD and MIR structure solution. *Acta Crystallog. sect. D*, **55**, 849–861.
33. Levitt, D. G. (2001). A new software routine that automates the fitting of protein X-ray crystallographic electron density maps. *Acta Crystallog. sect. D*, **57**, 1013–1019.
34. Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallog. sect. D*, **60**, 2126–2132.
35. Cowtan, K. (1994). DM: an automated procedure for phase improvement by density modification. *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography*, **31**, 34–38.

36. Sheldrick, G. M. & Schneider, T. R. (1997). SHELXL: high-resolution refinement. *Methods Enzymol.* **277**, 319–343.
37. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W. *et al.* (1998). Crystallography and NMR system: a new software suite for macromolecular structure determination. *Acta Crystallog. sect. D*, **54**, 905–921.
38. Johnson, M. S., Overington, J. P. & Blundell, T. L. (1993). Alignment and searching for common protein folds using a data bank of structural templates. *J. Mol. Biol.* **231**, 735–752.
39. Barton, G. J. (1993). ALSCRIPT- A tool to format multiple sequence alignments. *Protein Eng.* **6**, 37–40.
40. Rocchia, W., Alexov, E. & Honig, B. (2001). Extending the applicability of the nonlinear Poisson-Boltzmann equation: multiple dielectric constants and multivalent ions. *J. Phys. Chem. B*, **105**, 6507–6514.

Edited by R. Huber

(Received 29 May 2006; received in revised form 26 July 2006; accepted 27 July 2006)
Available online 1 August 2006