Structure of a trimeric bacterial microcompartment shell protein, EtuB, associated with ethanol utilisation in Clostridium kluyveri

Dana Heldt, Stefanie Frank, Arefeh Seyedarabi, Dimitrios Ladakis, Joshua B Parsons, Martin J Warren, Richard W Pickersgill

Biochemical Journal Immediate Publication. Published on 28 Jul 2009 as manuscript BJ20090780

Structure of a trimeric bacterial microcompartment shell protein, EtuB, associated with ethanol utilisation in *Clostridium kluyveri*

Dana Heldt*, Stefanie Frank*, Arefeh Seyedarabi†, Dimitrios Ladikis*, Joshua B Parsons*, Martin J. Warren*, Richard W. Pickersgill†

*** Centre for Molecular Processing, School of Biosciences, University of Kent, Giles Lane, Canterbury, Kent CT2 7NJ, UK**

† School of Biological and Chemical Sciences, Queen Mary University of London, Mile End Road, London E1 4NS, UK

Correspondence should be sent to either:

m.j.warren@kent.ac.uk, tel 00 44 1227 824690, fax 00 22 1227 763912 or

r.w.pickersgill@qmul.ac.uk, tel 00 44 2078828444, fax 00 44 20 8983 0973

Short title: structure of a trimeric microcompartment protein

Key words: bacterial microcompartment, shell protein, metabolosome, organelle, pore, protein sheet

Biochemical Journal Immediate Publication. Published on 28 Jul 2009 as manuscript BJ20090780

SUMMARY

It has been suggested that ethanol metabolism in the strict anaerobe *Clostridium kluyveri* occurs within a metabolosome, a subcellular proteinaceous bacterial microcompartment. Two bacterial microcompartment shell proteins (EtuA and EtuB) are found encoded on the genome clustered with the genes for ethanol utilisation. The function of the bacterial microcompartment is to facilitate fermentation by sequestering the enzymes, substrates, and intermediates. Recent structural studies of bacterial microcompartment proteins have revealed both hexamers and pentamers that assemble to generate the pseudo-icosahedral bacterial microcompartment shell. Some of these shell proteins have pores on their symmetry axes. Here we report the structure of the trimeric bacterial microcompartment protein EtuB, which has a tandem structural repeat within the subunit and pseudo-hexagonal symmetry. The pores in the EtuB trimer are within the subunits rather than between symmetry related subunits. We suggest that the evolutionary advantage of this is that it releases the pore from the rotational symmetry constraint allowing more precise control of the fluxes of asymmetric molecules, such as ethanol, across the pore. We also model EtuA and demonstrate that the two proteins have the potential to interact to generate the casing for a metabolosome.

INTRODUCTION

Bacterial microcompartments are prokaryotic organelles made up of a protein shell that encapsulates a metabolic process [1-5]. These polyhedral structures are typically 100-200 nm in diameter and represent novel bioreactors that are dedicated to a specific cellular pathway. The best characterised of the bacterial microcompartments is the carboxysome, which are structures associated with carbon fixation that are found in cyanobacteria and many chemotrophic bacteria [6-9]. Here, polyhedral shell proteins encase the enzymes carbonic anhydrase and ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCo). The role of the microcompartment is thought to help concentrate carbon dioxide to overcome the inefficiency of RuBisCo, which is the rate-limiting step in the Calvin cycle [10-14].

Carboxysomes are icosahedral structures with an outer shell that contains somewhere around 10,000 protein subunits [3, 5, 9]. In the last few years progress has been made on the structural biology of these interlocking shell proteins [4, 7, 9, 15]. In the carboxysome there appear to be about 5 proteins that make up the outer shell. The proteins have a recognisable sequence that is referred to as the bacterial microcompartment (BMC) motif. Most of the individual shell proteins form cyclical hexamers, which further assemble in a side-by-side fashion to generate large protein sheets. This likely forms the facet of the structure [7, 9, 15]. Some of the shell proteins appear to form a pentameric structure, and these have been proposed to act as the vertices of the icosahedral shell [5, 9]. Within some of the hexameric and pentameric protein assemblies, a central pore is observed and this is thought to serve as the major route for the diffusion of substrates and products [4, 7, 9, 15]. The proteins have a very distinctive surface charge distribution, with many positively charged amino acid residues surrounding the central pore and these again are thought to assist in the diffusion of substrates and products.

Sequencing studies revealed the presence of genes encoding shell proteins in operons associated with metabolic processes other than carbon fixation. Proteins with BMC motifs were found encoded in the operons associated with 1,2-propanediol metabolism and ethanolamine utilisation in *Salmonella enterica* [16-18]. Indeed, growth of *S. enterica* on media containing either 1,2-propanediol or ethanolamine as an energy source resulted in the appearance of microcompartments when the bacteria were analysed by electron microscopy after thin sectioning [19]. Both these metabolic processes require adenosylcobalamin (coenzyme B_{12})-dependent enzymes and both proceed via aldehyde intermediates. It has been suggested that the main purpose of the microcompartment here is to reduce the toxicity of the aldehyde intermediate by sequestering the compound within the confines of the organelle [20] although it has also been proposed that the organelle could conserve the volatility of the intermediate [13]. The aldehyde is then disproportionated by alcohol and aldehyde dehydrogenases within the structure [1, 3, 21].

The microcompartment associated with 1,2-propanediol metabolism has been characterised more than its ethanolamine orthologue. The actual physical structure is not so well defined as the clear icosahedral shape observed with the carboxysome. However, the operon encodes 21 gene products and 9 of these are thought to constitute shell proteins [1, 3, 16, 22]. This larger number of shell proteins may give rise to a more complex rounded structure. In contrast to the carboxysome, only one (non-essential) shell protein has had its structure determined, PduU, which reveals that it adopts a similar fold to the other BMC-containing proteins and forms a hexamer [23]. The microcompartment also contains more metabolic enzymes than are found in the carboxysome as well as a number of reactivation factors for the adenosylcobalamin coenzyme [3, 22, 24-26]. For this

reason the structure has been referred to as a metabolosome [27, 28]. Significantly, it has been shown that the propanediol utilisation operon can be genetically transferred between bacterial species to allow production of a functional structure in the new host [28]. This has significant potential for the design of bespoke bioreactors and metabolic engineering.

More recently, genome sequencing projects have revealed that bacterial microcompartments are more widespread than originally thought [3, 4, 29]. Indeed, it has been estimated that about 25% of prokaryotes have the genetic capacity to form such structures. Significantly, though, only the carboxysome and the ethanolamine and propanediol metabolosomes have been physically characterised and the remaining systems remain hypothetical. In this paper we report on a potentially novel bacterial microcompartment that is found in the anaerobic organism *Clostridium kluyveri*. This bacterium is able to grow on ethanol and acetate as sole carbon sources, and it has been proposed that the organism encloses ethanol and acetaldehyde dehydrogenases within a microcompartment [30]. Indeed, a serendipitous electron micrograph of a thin section of *C. kluyveri* grown on ethanol and acetate reveals the presence of such a polyhedral structure [31]. The genes for ethanol metabolism are encoded within a potential operon, which includes genes for two ethanol dehydrogenases, three aldehyde dehydrogenases and two putative shell proteins (Figure 1A). The enzymes are known to associate to allow the oxidation of ethanol to acetyl CoA (Figure 1B) [30]. The two shell proteins have been termed ethanol utilization shell protein (*etu*) genes, *etuA* and *etuB*. EtuA contains 92 residues and shares 60% sequence identity with the carboxysome protein CsoS1A from *Halothiobacillus neapolitanus* (2ewh) and PduA from the propanediol metabolosome. These proteins have a single BMC motif and, in the case of CsoS1A, form a hexamer [7]. The second *C. kluyveri* protein, EtuB, is unusual in that it is much larger (304 residues), has a tandem sequence repeat of the BMC motif and shares about 60% similarity with PduB from the propanediol metabolosome. In this paper we report on the cloning of *etuA* and *etuB*, the effect of their recombinant production in *E. coli* and the structure determination of the more unusual shell protein, EtuB.

EXPERIMENTAL PROCEDURES

Cloning and protein production

The *etuA* gene was amplified by PCR from *C. kluyveri* genomic DNA using primers CK1 (cttcatatgggacaggaagcattagg) and CK2 (cttggatccgctagcttagccaatgttaggc) and cloned into the *Nde*I and *Bam*HI site of pET3a and pET14b. The *etuB* gene was amplified by PCR from *C. kluyveri* genomic DNA using primers CK3 (cttcatatgaataatgagttaatcg) and CK4 (cttactagtttaaaaagtaggtttagac). The PCR product was cloned into pET14b into the *Nde*I and *Spe*I site. This cloning strategy allowed the encoded proteins to be overproduced with N-terminal hexahistidine tags.

The encoded proteins were overproduced in *E. coli* BL21(DE3)plysS cells with an N-terminal hexahistidine tag. Cells were grown in LB containing 100 mg/liter ampicillin and 34 mg/liter chloramphenicol with aeration at 37° C. Upon reaching an A_{600} of 0.6, the protein overproduction was induced with isopropyl 1-thio-β-D-galactopyranoside (0.4 mM), and growth was continued overnight at 16 °C. The cells were harvested by centrifugation (15 min, 4000 x g) and resuspended in 15 ml of binding buffer (0.5 M NaCl, 5 mM imidazole, 20 mM Tris-HCl, pH 8.0). Cell lysis was achieved by sonication, and the recombinant protein was purified using immobilized metal affinity chromatography. The supernatant was applied to a nickel-charged sepharose column. Unbound protein was washed off with binding buffer (0.5 M NaCl, 5 mM imidazole, Tris-HCl, pH 8.0), wash buffer I (0.5 M NaCl, 50 mM imidazole, Tris-HCl, pH 8.0) and wash buffer II (0.5 M NaCl, 100 mM imidazole, Tris-HCl, pH 8.0). Protein were eluted with buffer containing 0.5 M NaCl, 400 mM imidazole, Tris-HCl, pH 8.0 and finally buffer exchanged into buffer containing 50 mM Tris-HCl, pH 8.0, and 100 mM NaCl.

Prior to crystallisation, EtuB was further purified on a size exclusion column (Superdex 200 Global 10/30) which was connected to an Akta® FPLC chromatography system. Protein was eluted with a flow rate of 0.5 ml/min in 50 mM Tris-HCl pH 8.0 buffer containing 100 mM NaCl.

TEM: Preparation of samples for thin sectioning

Dehydration and embedding. Samples of *E. coli* strains overproducing EtuA, EtuB or PduB were prepared for thin sectioning to visualize the interior of the cells. The *E. coli* strains harbouring the appropriate plasmids were grown in 20 ml cell cultures at 37ºC to an OD_{600} of 1.5. The protein production was induced with 400 mM IPTG at 18 ˚C overnight. The cells were harvested by centrifugation and fixed for 2 hours in 5 ml 2.5 % Glutaraldehyde in PBS. Subsequently, the cells were pelleted and washed twice with PBS to remove traces of the fixing solution. The cells were stained for 2 hour in 1% osmium tetroxide and then washed twice with PBS before dehydration. This was accomplished by subjecting the samples to a solvent gradient: 60 % Industrial Methylated Spirit (IMS) overnight, 90 % IMS for 15 minutes, 100 % IMS for 15 minutes, three times 100 % dried ethanol for 2 hours each. The cells were embedded by incubation in 30 % Low Viscosity Resin (LV, Agar) in dried ethanol overnight followed by three LV resin changes with 100 % resin for 2 hours. This constituted blocks with medium hardness. The Samples were placed in 0.5 ml beem capsules, centrifuged for 5 minutes at 4000 x g to concentrate the cells to the tip and incubated at 60ºC overnight to polymerise.

Sectioning and visualisation of samples

Sections 80 nm thick were cut with an Ultra 450 MX 3768 diamond knife and placed on 300-mesh copper grids. The samples were stained with 4.5 % uranyl acetate in 1 % acetic acid for 30 minutes and with Reynolds lead citrate for 10 minutes at room temperature. Sections were observed and photographed with a JEOL-1230 transmission electron microscope.

Crystallization, Preparation of Heavy Atom Derivatives, and Data Collection

EtuB crystals were grown using the hanging drop method using 4 M sodium formate reservoirs and a protein concentration of 8 mg/ml. The 10 µl drops comprised equal volumes of protein solution and reservoir. Single crystals were harvested in litholoops, transferred through reservoir supplemented with 10% PEG200 as cryoprotectant, and stored in liquid nitrogen prior to data collection. Five heavy atom reagents were screened at two concentrations, the most successful of which was a 5 mM methylmercury chloride soak for 3 hours. The heavy atom derivative was prepared because the diffraction from crystals of Se-Met labelled EtuB was too poor to enable protein phases to be calculated from anomalous differences. Data were collected at 100 K using station ID29 at ESRF and IO4 at Diamond.

Structure Determination and Analysis

EtuB crystallizes in space group $I2_13$ with a = 174.3 Å with a single molecule in the asymmetric unit and a solvent content of 80%. Native data were collected to 3.0 Å and protein phases calculated to 3.3 Å using the monomethyl mercury derivative. The CCP4 program suite was used for structure solution [32]. MOSFLM [33], was used for data processing and SCALA [34] for scaling. Two mercury sites were found using SHELX [35], refined using MLPHARE [36], and the phases improved and extended using DM [37]. ARP/wARP [38] was used for automated model building with some rebuilding completed using COOT [39] and REFMAC [40]. The final model comprises residues 75 to 304 and one water molecule. Data collection statistics and the quality of the final model are presented in Table 1. Comparative modelling used MODELLER [41, 42] and structural comparisons made use of DALI [43].

RESULTS

Cloning of *etuA* and *etuB* and shell protein production

The genes corresponding to the putative *C. kluyveri* ethanol metabolosome shell proteins (*etuA, etuB*; database entry numbers CKL_1072 and CKL_1073 respectively) were amplified from *C. kluyveri* genomic DNA by PCR using primers that allowed the amplified fragments to be cloned into pET14b and pET3a vectors to permit the production of the shell proteins with and without and N-terminal His-tag. When transformed into *E. coli* BL21(DE3)plysS cells, the resulting strains harbouring *etuA* were found to overproduce the protein, but the protein was found largely in the insoluble fraction upon cell lysis. EtuA could be solubilised with chaotrophic agents such as urea but the protein precipitated as the salt was removed. In contrast, cells harbouring EtuB produced large quantities of soluble protein, which was easily purified by immobilised metal affinity chromatography using a Ni^{2+}-column.

Formation of long axial filaments in *E. coli* by EtuA production

Analysis of thin sections of the *E. coli* strain containing *etuA* in pET3a by transmission electron microscopy revealed some interesting features. Cells grown to stationary phase revealed the presence of ordered long axial filaments (Figure 2A). These filaments appear to interfere with cell division, apparently preventing septation, and also take up a significant amount of the cell cytoplasm. Similar, though smaller, structures have previously been reported with the overproduction of PduA from the propanediol metabolosome [28, 44]. It would appear that EtuA has the ability to self assemble into large macromolecular structures. Overproduction of EtuB did not produce any internal structures when cells were analysed by transmission electron microscopy (Figure 2B), although overproduction of the orthologous protein from the *pdu* operon (PduB) did reveal the presence of protein filaments wound around the inside of the cell membrane (Figure 2C)

Purification and crystallisation of the Etu shell proteins

Attempts were made made to purify EtuA but the protein was found to be sparingly soluble. However, as EtuA has 60% sequence identity with the carboxysome protein CsoS1A from *Halothiobacilus neapolitanus* (2ewh) whose structure has previously been determined [15], a model of EtuA (Figure 3A) can be readily built by comparative modelling. This suggests that EtuA is likely to form a hexamer. In contrast to EtuA, EtuB was much more soluble and could be purified by metal affinity and size exclusion column chromatography. The protein was then used in crystallisation trials using Molecular Dimensions crystal screens I and II. Several conditions were found that produced small crystals. One of these was optimised to give large crystals that were used for structural studies.

Structure of EtuB

The crystal structure of EtuB was determined using single isomorphous replacement with anomalous scattering and refined to a resolution of 3.0 Å, with final R and R-free values of 0.18 and 0.20 (see Experimental Procedures). The quality of the final model is good as judged by the stereochemistry and the validation of results. The EtuB molecular trimer sits on the crystallographic three-fold axis of the $I2_13$ cell so that there is a single molecule in the asymmetric unit (Figure 3B). The close and extensive association of three subunits around the crystallographic three-fold axis reveals that this is the biologically authentic

oligomer (Figure 4A).

The EtuB trimer has a concave and a convex face (Figures 4A and 4C, respectively); both faces are characterised by significant negative potential. The convex face has positive potential close to the three-fold axis and on the periphery. The concave face is punctuated by a number of small positive patches. Both the pseudo-hexameric structure and the three-fold symmetry of the EtuB trimer manifest in the electrostatic potential are clear in Figure 3B. In contrast to the acidic faces, the edges of the EtuB trimer, which are anticipated to be involved in interactions forming the shell of the microcompartment (Figure 4B), are more neutral in nature.

The total solvent accessible surface area which is buried on forming the trimer is 7850 Å^2. Surprisingly, the atomic structure of EtuB lacks the first 74 residues but is otherwise complete. To test the possibility that the N-terminal region of the EtuB structure might not comply with the 3-fold symmetry of the C-terminal domain, the diffraction data were processed in the lower symmetry sub-group $I2_12_12_1$ with a trimer, rather than a single subunit in the asymmetric unit. In the resulting electron density map there was no evidence for the N-terminal 74 residues, even in electron density maps calculated at low resolution. It is likely that the first 74 residues are cleaved during crystallization. SDS-PAGE, Western and Maldi analysis of solubilized protein crystals all support this conclusion (data not shown). The EtuB sequence is closely similar, over 50% sequence identity, to that of the *Citrobacter fruendii* PduB (Figure 5A). The major area of homology aligns between positions 75 and 304 of EtuB i.e. from where EtuB is truncated during the crystallisation process. The N-terminal regions of EtuB (1-74) and PduB (1-43) share a much lower level of similarity. In both *S. enterica* and *C. freundii* PduB is produced in two forms, termed PbuB and PduB', which result from the presence of two translation start sites on the mRNA giving rise to two protein species that differ in the nature of the first 37 amino acid residues [22, 28]. It is not known if EtuB is also produced in more than one form. Nonetheless, the high degree of sequence similarity between EtuB and PduB suggest that they have a very similar structure, a conservation that includes many of the residues that are found around the three pores in the complex. Conversely, it is not known if the first 70 residues of EtuB have any functional significance. In EtuB the presence of the first 70 amino acids appears to make the protein more soluble in comparison to PduB. Indeed, PduB appears to form stable filaments within *E. coli* that form an inner insulating layer (Figure 2C). The filaments are possibly formed through a helical arrangement of trimers within the fibre.

A tandem repeat of the BMC fold

The EtuB subunit is wedge shaped and comprises a four-layer βαβα sandwich (Figure 5B). There is a tandem repeat of the BMC motif within the single polypeptide chain. The domain architecture is most closely similar to that of PduU [23] (Figure 5C) with rmsd of 2.2 Å over 89 equivalenced Cα atoms with sequence identity 19%; DALI server. PduU is described as a circularly permuted BMC domain with the N-terminal region contributing the secondary structure elements β2 and the short first α-helix typically contributed by the C-terminal region. The principal elements of the BMC fold are a four-stranded antiparallel β-sheet with two α-helices on one side. PduU has an additional short β-strand, β6. Compared to PduU, the N-terminal polypeptide chain of the N-terminal BMC repeat of EtuB replaces the C-terminal polypeptide of PduU running antiparallel to β2 and β1 is missing (the strands are labelled according to the PduU assignment). The C-terminal polypeptide of the N-terminal BMC repeat of EtuB then extends as a bent α-helix and

continues to form a short first β-strand (β6') of the second BMC fold (Figure 5B). The topology of this second fold is the same as the first, the C-terminal polypeptide folds down and across the surface of the second BMC repeat (Figure 5B).

Superimposing the two BMC folds of EtuB, reveals they have 21% sequence identity over 94 aligned residues, which have a rmsd of Cα atoms of 2.7 Å (Figure 5D). The conserved β-core with strand-order 2,3,5,4 can be clearly seen. The greatest structural difference between the two BMC domains occurs in the loops between β4 and β5. There are good structural and functional reasons for the differences between these loops. In the N-terminal BMC domain the β4/β5 loop forms one side and contributes His 156, Thr 150 and main chain amides and carbonyls to the pore. In the C-terminal BMC domain the β4/β5 loop packs with the equivalent loops from the two other subunits close to the molecular 3-fold axis. An obvious residue from this loop involved in packing close to the rotation axis is Tyr 259. The duplication of the BMC motif facilitates the specialised roles of these two loops.

Assembly into an icosahedral shell

The EtuB trimer has the appearance of a hexameric disk of side 42 Å with the subunits packing tightly around the 3-fold axis (Figure 3B and 4). The N and C termini are on the convex side of the pseudo-hexameric disk (Figure 4B, C), the other side is concave (Figure 4A). Close to the 3-fold axis the disk is approximately 25 Å thick increasing to 35 Å towards the edge of the disk. The packing in this crystal form gives no indication of how the hexamers may pack to form a flat sheet of molecules required to form an icosahedral face. There is also no indication as to which face of the oligomer is inside, which out, or indeed if there is a preferred inside or outside face. The sloped packing surfaces of the oligomer (Figure 4B) suggests that to form a flat strip these molecules would have to pack alternatively face in followed by face out. These strips could then be staggered to form a sheet. The missing N-terminal 74 residues might also play a part in the assembly of the pseudo-hexamers, either by binding them together or by filling the wedge between oligomers when packing into a sheet in which the oligomers all have the same orientation. The N-termini is located close to the edge of the trimer such that it could have a role in either an embrace between adjacent oligomers. It is correctly positioned such that it could fill the wedge left when the oligomers are packed in a consistent orientation. If this were the case then perhaps it is more likely that the convex side of the oligomer is in. It appears that the EtuB trimer can be accommodated within a sheet of EtuA hexamers and this may be a plausible model for a facet (see discussion).

As mentioned earlier, a structure of EtuA can be readily built by comparative modelling (Figure 3A). On the basis of this model, EtuA is expected to form a hexamer with the packing of subunits as seen in CsoS1A. The β4/β5 loop of EtuA occurs close to the symmetry axis and although the model suggests this loop may pack close to the symmetry axis it is equally probable that this flexible four-residue β-hairpin (Ile-Gly-Ser-Gly) packs alternatively creating a pore close to the symmetry axis. Below the β4/β5 loop there is a ring of lysines (Lys 37) that might contribute to the selectivity of the pore. It is notable that in the *C. kluyveri* ethanol utilization gene cluster there is no gene for encoding for a product that could form a pentamer. Pentamers would occupy each of the vertices of a regular icosahedral shell. There is however a candidate for the pentamer elsewhere in the *C. kluyveri* genome associated with the apparently defunct glycerol utilization microcompartment (database entry number CKL_0849) [30]. This protein has 52%

sequence identity with the carboxysome pentamer CcmL from *Synechocystis* 6893 (2qw7) [9]. More work on isolated microcompartments from *C. kluyveri* is needed to determine their precise composition so that models of the pseudo-icosahedral shell can be constructed.

Subunit pore

The pores in the shell proteins solved to date are located around the five-fold or six-fold rotation axis that relates subunits. EtuB has no such pore, the 12 residue β4/β5 loop and β4'/β5' loops are relatively complex in structure, certainly compared to the β-hairpin in EtuA, and pack tightly around the three-fold axis. Arg 148 and Tyr 259 are two residues that pack around the symmetry axis of the EtuB trimer. EtuB does however have a pore within the subunit, facilitated by the gene duplication that has given rise to the tandem repeat of EtuB. No subunit pore has previously been observed in any BMC shell protein characterised structurally to date. Three histidines (156, 195, and 224) and two glutamates (197 and 262) line the pore (Figure 6A) and four out of these five residues are conserved in the PduB sequence highlighting their significance (Figure 5A). Additionally some hydrophobic residues line the pore including Phe 212, Val 256, and Pro 124. The ethanol utilization operon of *C. kluyveri* contains two aldehyde dehydrogenases and three alcohol dehydrogenase genes (Figure 1A). The size and characteristics of the subunit pore are consistent with ethanol diffusion into the BMC. The pore in the crystal structure has a single water molecule trapped within it (Figure 6B). Clearly, the tandem duplication affords the possibility of evolving an asymmetric pore potentially one with greater selectivity than one that must conform to strict rotational symmetry.

DISCUSSION

Shell proteins for a potential metabolosome within *C. kluyveri*.

It has been proposed that the ethanol utilisation enzymes in C. kluyveri are contained within a BMC. Indeed, the genes for the ethanol and acetaldehyde dehydrogenases are clustered on the genome together with two genes for potential shell proteins of such a proteinaceous organelle. However, although the formation of such a metabolosome is attractive, hard experimental evidence for its presence has not yet been produced and its existence cannot therefore be assumed. Nonetheless, in this paper we report the structure of one of these two shell proteins (EtuB) and model the other (EtuA).

Bacterial shell proteins characterised to date form hexamers and pentamers with the single pore per oligomer centred on the symmetry axis [4, 5]. The EtuB subunit has a domain duplication and the EtuB trimer has pseudohexameric appearance due to the structural similarity of the two BMC domains present in the subunit. The core structure of the domain is restrained to the ancestral hexameric template essential for forming a facet of the microcompartment. In fact, the conservation of the pseudohexameric structure and of residues that may be key in forming contacts with neighbouring molecules provide support for the occurrence of EtuB in the *C. Kluyveri* microcompartment. Both pseudohexamer edges have the strongly conserved lysine at position 23 characteristic of the canonical BMC domain (see 567 Pfam00936 HMM Logo) this corresponds to lysines 135 and 241 in EtuB. This lysine is found at the 2-fold axis between adjacent hexamers in the CcmK1, CcmK2, and CsoS1A layers [15]. The lysine makes a hydrogen bond to the

main chain carbonyl of the lysine in the opposite hexamer and to the side chain of aspartate within the same subunit (position 19 in the canonical BMC domain). This aspartate is also conserved in each BMC domain of EtuB (Asp 131 and Asp 237). The conservation of these surface residues further supports the view that EtuB forms a part of the microcompartment facet.

Consequences of the novel features of the EtuB pore.

More peripheral structural elements are less restrained by the symmetry of the ancestral protein particularly those structural elements that do not affect the surfaces involved in packing to form the microcompartment facet. In EtuB, the internal β4/β5 loop adopts different conformations in each of the two copies of the BMC domain. These two loops contribute residues that pack around the symmetry axis as well as residues that contribute to the subunit pore. The subunit pore is a novel feature of EtuB, which may have two advantages. The first is that it is present at three-times the concentration of pores that are coincident with the six-fold symmetry axis; this increase in concentration (or density of pores) may be important if the rate of exchange of substrate or product molecules across the microcompartment shell is rate limiting but only if EtuB is abundantly present in the shell. The second advantage is that the architecture of the pore is released from the symmetry constraint which may allow it greater potential to become more selective. The pore is not large but may allow the passage of ethanol.

The asymmetry of the EtuB oligomer, compared to the carboxysome hexamer, might also account for the more pleiomorphic shape of the *C. kluyveri* shell compared to that of the carboxysome. In reality little is known about the proposed ethanol utilisation bacterial microcompartment. EM evidence is suggestive of such bodies [31] and biochemical evidence points to a close association between the alcohol and aldehyde deydrogenases since the various enzymes and isoenzymes are isolated in a large macromolecular complex [31, 45] as would be expected for a metabolosome. This coupled with the organisation of the respective genes into a cluster with the genes for the shell proteins [30] provides compelling evidence that this ethanol utilisation system is incorporated into a bacterial microcompartment. With this in mind, the structure of the microcompartment must also reflect the biochemical role that it plays within the metabolism of the cell. The structure must be able to allow ethanol into the macromolecular assembly together with coenzyme A (CoASH) and NAD^+. The products of the reaction have to be able to leave the complex so acetyl-CoA and NADH have to be able to exit the structure. Access to and entry from the complex is likely to be via the pores in the shell proteins. The pores in EtuB are comparatively small and polar and thus should be able to allow the passage of ethanol. The pores are too small to allow larger more complex molecules through such as the coenzymes or their derivatives. The structure of EtuB is likely to be very similar to PduB, which is involved in propanediol metabolism. Here the pores in PduB could also allow the substrate 1,2-propanediol into the bacterial microcompartment. In both cases, however, the other substrates/coenzymes/cofactors must access by another shell protein, or via specific channels formed between different shell unit components. In the propanediol metabolosome there are seven different shell protein subunits providing a high level of complexity. In the ethanol utilisation system there appear to be only two components. The EtuA shell has been modelled and does not appear to have a central pore (Figure 3A). A conformational change has to be invoked to permit the passage of larger molecules. In the large macromolecular assembly formed between EtuA and EtuB a certain amount of breathing could be envisaged that could also permit the transfer of metabolites from outside to inside and vice versa. Much more detail on the composition of

the ethanol utilisation bacterial microcompartment is required as well as detail on the interaction between EtuA and EtuB in order to advance our current understanding of these structures.

Recent trimeric shell protein structures.

Remarkably, a proof appeared at the time of writing this paper describing the structure of a trimeric carboxysome shell protein CsoS1D [46]. Both the BMC architecture and the tandem fusion of BMC motifs appear to be closely similar to EtuB (CsoS1D coordinates are on hold at PDB). The major difference is that CsoS1D has a large pore, apparently gated, on the three-fold symmetry axis, as opposed to the small subunit pore of EtuB. Duplication of the BMC motif has therefore allowed the evolution of at least two types of novel pore. The crystallisation of another shell protein, EutL, which also contains a duplicated BMC domain has also been reported recently, and this protein displays many structural similarities to EutB, including the presence of three pores [47]. Together, all of this recent information will greatly assist in providing further molecular detail on the structure/function of these remarkable proteinaceous bacterial organelles and may allow, in due course, the construction of bespoke bioreactors.

ACCESSION NUMBERS

Coordinates and structure factor amplitudes for EtuB have been deposited in the protein databank with the accession code XXXX.

ACKNOWLEDGEMENTS

We acknowledge financial support from BBSRC (BB/E01563) and HEFCE. Synchrotron data were collected at ESRF (Grenoble) and Diamond (Oxford).

Table 1. Crystallographic data statistics

	Native	***MMC***
Resolution (Å)[1]	86.71-3.00 (3.17-3.00)	71.43-3.33 (3.51-3.33)
Rmerge[2]	0.151 (0.609)	0.115 (0.434)
Mean((I)/sd(I))	16.8 (3.3)	14.3 (3.5)
Completeness (%)	100.0 (100.0)	100.0 (100.0)
Multiplicity	10.7 (10.9)	7.1 (7.2)
R-factor (R-free)[3]	0.188 (0.204)	
rmsd bond (Å)[4]	0.011	
rmsd angle (Å)[4]	1.815	
Ramachandran	96.4% (allowed)	

[1]The high-resolution range and the parameter values for this range are given in parentheses in this Table. [2] Rmerge = $\Sigma_{hkl} \Sigma_i |I_i - \langle I \rangle| / \Sigma_{hkl} \Sigma_i I_i$, where I_i is the intensity of the i^{th} observation, $\langle I \rangle$ is the mean intensity of the reflection, and the summations extend over all unique reflections (hkl) and all equivalents (i), respectively. [3]R-factor = $\Sigma_{hkl} |F_o - F_c| / \Sigma_{hkl} F_o$, where F_o and F_c represent the observed and calculated structure factors, respectively. The R-Factor is calculated using 95% of the data included in refinement and R-free the 5% excluded. [4]Root mean square deviation (rmsd). MMC is the data collected from a crystal soaked in 5mM mercury methyl chloride soaked for 1 hour.

Biochemical Journal Immediate Publication. Published on 28 Jul 2009 as manuscript BJ20090780

FIGURE LEGENDS

Figure 1. The putative *C. kluyveri* ethanol utilisation bacterial microcompartment.

A. The putative ethanol metabolising operon found in the *C. kluyveri* genome.

B. Diagrammatic representation of the metabolic role of the ethanol utilisation bacterial microcompartment. Ethanol enters the structure via a pore and is metabolised to acetyl CoA by the internalised enzymes. The product is then released back into the cytoplasm.

Figure 2. Thin sections of *E. coli* strains overproducing some shell proteins.

A. Thin section of an *E. coli* strain overproducing *C. kluyveri* EtuA. The protein overproduction results in the appearance of long axial filaments that then interfere with septation and cell division.

B. Thin section of an *E. coli* strain overproducing *C. kluyveri* EtuB. The cells appear as normal with no extra internal structures.

C. Thin section of an *E. coli* strain overproducing *C. freundii* PduB. The cells are found to contain long filaments that are wound around the internal side of the cell membrane.

Figure 3. Cartoon representation of the bacterial microcompartment proteins

The structures of two shell proteins EtuA and EtuB found in the ethanol utilization gene cluster of *C. kluyveri*. The cartoon representations are coloured by chain to demonstrate the relationship between the previously seen hexameric structure and the novel trimeric architecture. It is anticipated that these proteins form the facets of the microcompartment. It is not yet clear if *C. kluyveri* produces pentamers to occupy the vertices of the pseudo-icosahedral shell or indeed if this is necessary given the more pleiomorphic shape of the *C. kluyveri* bacterial microcompartments and the possibility of quasi-equivalent interactions around a pentamer axis.

A. Homology model of EtuA, a hexamer, which is similar to previously solved shell proteins and is modelled on a carboxysome shell protein (see text for further details of both EtuA and EtuB).

B. Crystal structure of EtuB is a novel shell protein, a trimer, with a tandem repeat of the BMC motif.

Figure 4. Electrostatic potential mapped to the surface of the EtuB trimer.

The surface is coloured according to electrostatic potential calculated using macroscopic dielectric constants of 2 and 80 for protein and water, respectively. The molecule is rotated 90° about the y-axis in the successive images A, B, C to reveal the two faces and the side of the pseudohexameric disk. The pseudohexameric structure is seen in the shape of the molecule while the electrostatics clearly reveal the trimeric distribution of charge i.e. there is no hexameric constraint on the charge distribution, but there is on shape. If A is the outer surface; B is the side view (that seen by the other shell proteins);

and C is then the inner surface. The C-terminus is on the C side of the molecule and the N-terminal residue in the structure (residue 77) is closer to the C side (as indicated in panels B and C). The position of one pore is indicated on surfaces A and C, the other pores can be seen following the 3-fold symmetry of the trimer. The two faces A and C are dominated by negative potential, as are the channels in this calculation which includes formal charges only. The edges of the molecule are more balanced in potential. Red is potential below 3kT, blue above 3kT, and white is neutral. Electrostatic calculations were made using DELPHI [48].

Figure 5. Sequence and structure of EtuB.

A. Sequence of *C. kluyveri* EtuB and its similarity to *C. freundi* PduB. Amino acid residue similarity/identity as well secondary structure alignment is highlighted.

B. Stereo view of the structure of the EtuB subunit coloured from blue to red as the polypeptide chain progresses from the N to the C-terminus.

C. A dimer of PduU subunits (green and cyan) to allow comparison to the EtuB subunit. The orientation is similar to allow comparison with the EtuB subunit (B). Note however the difference in position of the N- and C- termini compared to EtuB.

D. Superimposition of the two BMC motifs of EtuB, the N-terminal repeat and C-terminal repeat are cyan and magenta, respectively. The four central β-strands are conserved in position; the most striking difference is in the β4/β5-loop (labelled β4'/β5' in the C-terminal BMC repeat).

Figure 6. The pores in the trimeric structure of EtuB.

A. Cutaway section looking through the pore of EtuB showing the positions of the three histidines and two glutamates that line the pore. The calculation of electrostatic potential was as described in the legend to Figure 4.

B. Electron density map and final structure of EtuB showing a slab of electron density perpendicular to the pore in the EtuB subunit. The electron density σ_A-weighted 2Fobs-Fcalc map is contoured at 1σ. The density of the blob is roughly spherical and has been modelled as a water molecule although the distance to the protein ligands are too long for hydrogen bonds. Residues lining this part of the pore are labelled.

Figure 7. A model of part of the face of the *C. kluyveri* microcompartment.

A. The packing of the EtuA hexamers (blue) is modelled on the packing of CsoS1A hexamers (2EWH) in the crystal. The central EtuA hexamer has been removed and replaced by a EtuB trimer (coloured by chain). The EtuA hexamers, like the CsoS1A hexamers, have a conserved interaction at their interfaces involving the absolutely conserved lysine of the canonical BMC motif (lysine 26 in EtuA). The EtuB trimer fits well and is also capable of making the same intermolecular interactions involving the conserved lysine.

B. The conserved lysines, Lys 135 and Lys 240, of the EtuB subunit form pseudo-equivalent interactions with the conserved lysines of the adjacent EtuA subunits involving side-chain to carbonyl oxygen hydrogen bonds as indicated in pink line. All figures were drawn using PYMOL [49].

REFERENCES

1 Bobik, T. A. (2006) Polyhedral organelles compartmenting bacterial metabolic processes. Applied Microbiology and Biotechnology. **V70**, 517-525

2 Cannon, G. C., Bradburne, C. E., Aldrich, H. C., Baker, S. H., Heinhorst, S. and Shively, J. M. (2001) Microcompartments in prokaryotes: carboxysomes and related polyhedra. Appl Environ Microbiol. **67**, 5351-5361

3 Cheng, S., Liu, Y., Crowley, C. S., Yeates, T. O. and Bobik, T. A. (2008) Bacterial microcompartments: their properties and paradoxes. Bioessays. **30**, 1084-1095

4 Tanaka, S., Sawaya, M. R., Phillips, M. and Yeates, T. O. (2009) Insights from multiple structures of the shell proteins from the beta-carboxysome. Protein Sci. **18**, 108-120

5 Yeates, T. O., Kerfeld, C. A., Heinhorst, S., Cannon, G. C. and Shively, J. M. (2008) Protein-based organelles in bacteria: carboxysomes and related microcompartments. Nat Rev Microbiol. **6**, 681-691

6 Iancu, C. V., Ding, H. J., Morris, D. M., Dias, D. P., Gonzales, A. D., Martino, A. and Jensen, G. J. (2007) The structure of isolated Synechococcus strain WH8102 carboxysomes as revealed by electron cryotomography. J Mol Biol. **372**, 764-773

7 Kerfeld, C. A., Sawaya, M. R., Tanaka, S., Nguyen, C. V., Phillips, M., Beeby, M. and Yeates, T. O. (2005) Protein structures forming the shell of primitive bacterial organelles. Science. **309**, 936-938

8 Schmid, M. F., Paredes, A. M., Khant, H. A., Soyer, F., Aldrich, H. C., Chiu, W. and Shively, J. M. (2006) Structure of Halothiobacillus neapolitanus carboxysomes by cryo-electron tomography. J Mol Biol. **364**, 526-535

9 Tanaka, S., Kerfeld, C. A., Sawaya, M. R., Cai, F., Heinhorst, S., Cannon, G. C. and Yeates, T. O. (2008) Atomic-level models of the bacterial carboxysome shell. Science. **319**, 1083-1086

10 Badger, M. R. and Price, G. D. (2003) CO2 concentrating mechanisms in cyanobacteria: molecular components, their diversity and evolution. J Exp Bot. **54**, 609-622

11 Dou, Z., Heinhorst, S., Williams, E. B., Murin, C. D., Shively, J. M. and Cannon, G. C. (2008) CO2 fixation kinetics of Halothiobacillus neapolitanus mutant carboxysomes lacking carbonic anhydrase suggest the shell acts as a diffusional barrier for CO2. J Biol Chem. **283**, 10377-10384

12 Heinhorst, S., Williams, E. B., Cai, F., Murin, C. D., Shively, J. M. and Cannon, G. C. (2006) Characterization of the carboxysomal carbonic anhydrase CsoSCA from Halothiobacillus neapolitanus. J Bacteriol. **188**, 8087-8094

13 Penrod, J. T. and Roth, J. R. (2006) Conserving a volatile metabolite: a role for carboxysome-like organelles in Salmonella enterica. J Bacteriol. **188**, 2865-2874

14 So, A. K., John-McKay, M. and Espie, G. S. (2002) Characterization of a mutant lacking carboxysomal carbonic anhydrase from the cyanobacterium Synechocystis PCC6803. Planta. **214**, 456-467

15 Tsai, Y., Sawaya, M. R., Cannon, G. C., Cai, F., Williams, E. B., Heinhorst, S., Kerfeld, C. A. and Yeates, T. O. (2007) Structural analysis of CsoS1A and the protein shell of the Halothiobacillus neapolitanus carboxysome. PLoS Biol. **5**, e144

16 Bobik, T. A., Havemann, G. D., Busch, R. J., Williams, D. S. and Aldrich, H. C. (1999) The propanediol utilization (pdu) operon of Salmonella enterica serovar Typhimurium LT2 includes genes necessary for formation of polyhedral organelles involved in coenzyme B(12)-dependent 1, 2-propanediol degradation. J Bacteriol. **181**, 5967-5975.

17 Kofoid, E., Rappleye, C., Stojiljkovic, I. and Roth, J. (1999) The 17-gene

ethanolamine (eut) operon of Salmonella typhimurium encodes five homologues of carboxysome shell proteins. J Bacteriol. **181**, 5317-5329

18 Stojiljkovic, I., Baumler, A. J. and Heffron, F. (1995) Ethanolamine utilization in Salmonella typhimurium: nucleotide sequence, protein expression, and mutational analysis of the cchA cchB eutE eutJ eutG eutH gene cluster. J Bacteriol. **177**, 1357-1366

19 Shively, J. M., Bradburne, C. E., Aldrich, H. C., Bobik, T. A., Mehlman, J. L., Jin, S. and Baker, S. H. (1998) Sequence homologs of the carboxysomal polypeptide CsoS1 of the thiobacilli are present in cyanobacteria and enteric bacteria that form carboxysomes - polyhedral bodies. Can. J. Bot.-Rev. Can. Bot. **76**, 906-916

20 Rondon, M. R., Kazmierczak, R. and Escalante-Semerena, J. C. (1995) Glutathione is required for maximal transcription of the cobalamin biosynthetic and 1,2-propanediol utilization (cob/pdu) regulon and for the catabolism of ethanolamine, 1,2-propanediol, and propionate in Salmonella typhimurium LT2. J Bacteriol. **177**, 5434-5439

21 Sampson, E. M. and Bobik, T. A. (2008) Microcompartments for B12-dependent 1,2-propanediol degradation provide protection from DNA and cellular damage by a reactive metabolic intermediate. J Bacteriol. **190**, 2966-2971

22 Havemann, G. D. and Bobik, T. A. (2003) Protein content of polyhedral organelles involved in coenzyme B12-dependent degradation of 1,2-propanediol in *Salmonella enterica* serovar Typhimurium LT2. J Bacteriol. **185**, 5086-5095

23 Crowley, C. S., Sawaya, M. R., Bobik, T. A. and Yeates, T. O. (2008) Structure of the PduU shell protein from the Pdu microcompartment of Salmonella. Structure. **16**, 1324-1332

24 Johnson, C. L., Buszko, M. L. and Bobik, T. A. (2004) Purification and initial characterization of the Salmonella enterica PduO ATP:Cob(I)alamin adenosyltransferase. J Bacteriol. **186**, 7881-7887

25 Johnson, C. L., Pechonick, E., Park, S. D., Havemann, G. D., Leal, N. A. and Bobik, T. A. (2001) Functional genomic, biochemical, and genetic characterization of the Salmonella pduO gene, an ATP:cob(I)alamin adenosyltransferase gene. J Bacteriol. **183**, 1577-1584

26 Sampson, E. M., Johnson, C. L. and Bobik, T. A. (2005) Biochemical evidence that the pduS gene encodes a bifunctional cobalamin reductase. Microbiology. **151**, 1169-1177

27 Brinsmade, S. R., Paldon, T. and Escalante-Semerena, J. C. (2005) Minimal functions and physiological conditions required for growth of salmonella enterica on ethanolamine in the absence of the metabolosome. J Bacteriol. **187**, 8039-8046

28 Parsons, J. B., Dinesh, S. D., Deery, E., Leech, H. K., Brindley, A. A., Heldt, D., Frank, S., Smales, C. M., Lunsdorf, H., Rambach, A., Gass, M. H., Bleloch, A., McClean, K. J., Munro, A. W., Rigby, S. E. J., Warren, M. J. and Prentice, M. B. (2008) Biochemical and structural insights into bacterial organelle form and biogenesis. Journal of Biological Chemistry. **283**, 14366-14375

29 Beeby, M., Bobik, T. A. and Yeates, T. O. (2009) Exploiting genomic patterns to discover new supramolecular protein assemblies. Protein Sci. **18**, 69-79

30 Seedorf, H., Fricke, W. F., Veith, B., Bruggemann, H., Liesegang, H., Strittmatter, A., Miethke, M., Buckel, W., Hinderberger, J., Li, F., Hagemeier, C., Thauer, R. K. and Gottschalk, G. (2008) The genome of Clostridium kluyveri, a strict anaerobe with unique metabolic features. Proc Natl Acad Sci U S A. **105**, 2128-2133

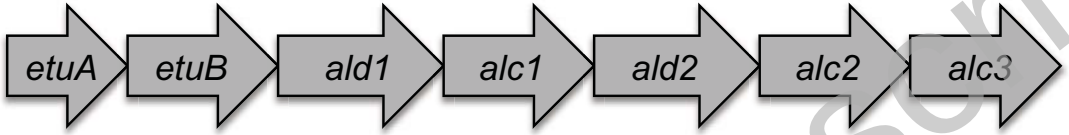
31 Lurz, R., Mayer, F. and Gottschalk, G. (1979) Electron-Microscopic Study on the Quaternary Structure of the Isolated Particulate Alcohol-Acetaldehyde Dehydrogenase Complex and on Its Identity with the Polygonal Bodies of

Clostridium-Kluyveri. Archives of Microbiology. **120**, 255-262
32 Bailey, S. (1994) The Ccp4 Suite - Programs for Protein Crystallography. Acta Crystallographica Section D-Biological Crystallography. **50**, 760-763
33 Leslie, A. G. W. (1992) Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography. **26**
34 Evans, P. (2006) Scaling and assessment of data quality. Acta Crystallographica Section D-Biological Crystallography. **62**, 72-82
35 Sheldrick, G. M. (1998) SHELX applications to macromolecules. Direct Methods for Solving Macromolecular Structures. **507**, 401-411
526
36 Otwinowski, Z. (1991) Isomorphous replacement and Anomalous Scattering Daresbury, UK.
37 Cowtan, K. D. and Zhang, K. Y. (1999) Density modification for macromolecular phase improvement. Prog Biophys Mol Biol. **72**, 245-270
38 Lamzin, V. S. and Wilson, K. S. (1997) Automated refinement for protein crystallography. Macromolecular Crystallography, Pt B. **277**, 269-305
39 Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. Acta Crystallographica Section D-Biological Crystallography. **60**, 2126-2132
40 Murshudov, G. N., Vagin, A. A. and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr. **53**, 240-255
41 Pieper, U., Eswar, N., Webb, B. M., Eramian, D., Kelly, L., Barkan, D. T., Carter, H., Mankoo, P., Karchin, R., Marti-Renom, M. A., Davis, F. P. and Sali, A. (2009) MODBASE, a database of annotated comparative protein structure models and associated resources. Nucleic Acids Research. **37**, D347-D354
42 Sali, A. and Blundell, T. L. (1993) Comparative Protein Modeling by Satisfaction of Spatial Restraints. Journal of Molecular Biology. **234**, 779-815
43 Holm, L., Kaariainen, S., Rosenstrom, P. and Schenkel, A. (2008) Searching protein structure databases with DaliLite v.3. Bioinformatics. **24**, 2780-2781
44 Havemann, G. D., Sampson, E. M. and Bobik, T. A. (2002) PduA is a shell protein of polyhedral organelles involved in coenzyme B(12)-dependent degradation of 1,2-propanediol in *Salmonella enterica* serovar typhimurium LT2. J Bacteriol. **184**, 1253-1261
45 Hillmer, P. and Gottschalk, G. (1972) Particulate nature of enzymes involved in the fermentation of ethanol and acetate by Clostridium kluyveri. FEBS Lett. **21**, 351-354
46 Klein, M. G., Zwart, P., Bagby, S. C., Cai, F., Chisholm, S. W., Heinhorst, S., Cannon, G. C. and Kerfeld, C. A. (2009) Identification and Structural Analysis of a Novel Carboxysome Shell Protein with Implications for Metabolite Transport. J Mol Biol (in press)
47 Sagermann, M., Ohtaki, A. and Nikolakakis, K. (2009) Crystal structure of the EutL shell protein of the ethanolamine ammonia lyase microcompartment. Proc Natl Acad Sci U S A. **106**, 8883-8887
48 Oron, A., Wolfson, H., Gunasekaran, K. and Nussinov, R. (2003) Using DelPhi to compute electrostatic potentials and assess their contribution to interactions. Curr Protoc Bioinformatics. **Chapter 8**, Unit 8 4
49 DeLano, W. L. The PyMOL Molecular Graphics System. DeLano Scientific LLC, San Carlos, CA, USA, http://www.pymol.org.

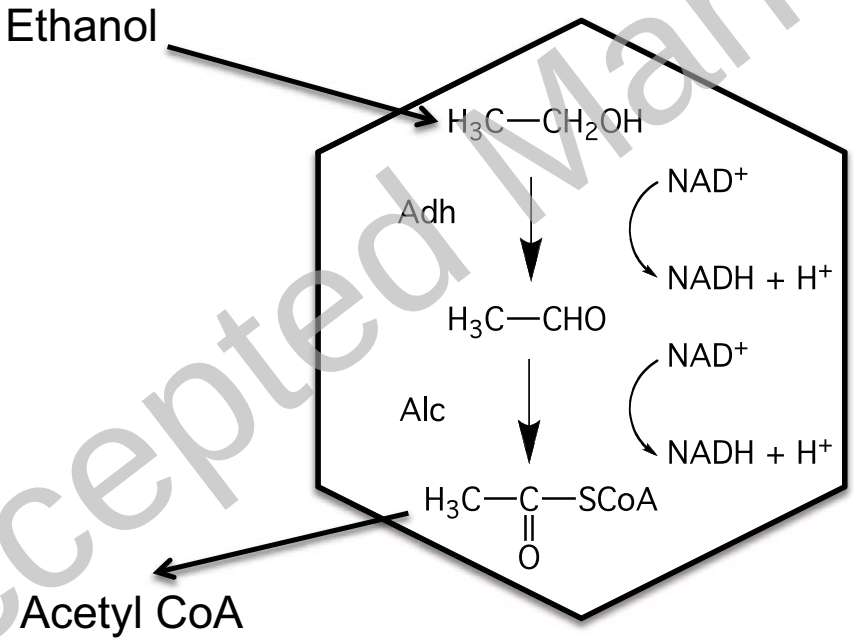
Biochemical Journal Immediate Publication. Published on 28 Jul 2009 as manuscript BJ20090780

Figure 1

(A)

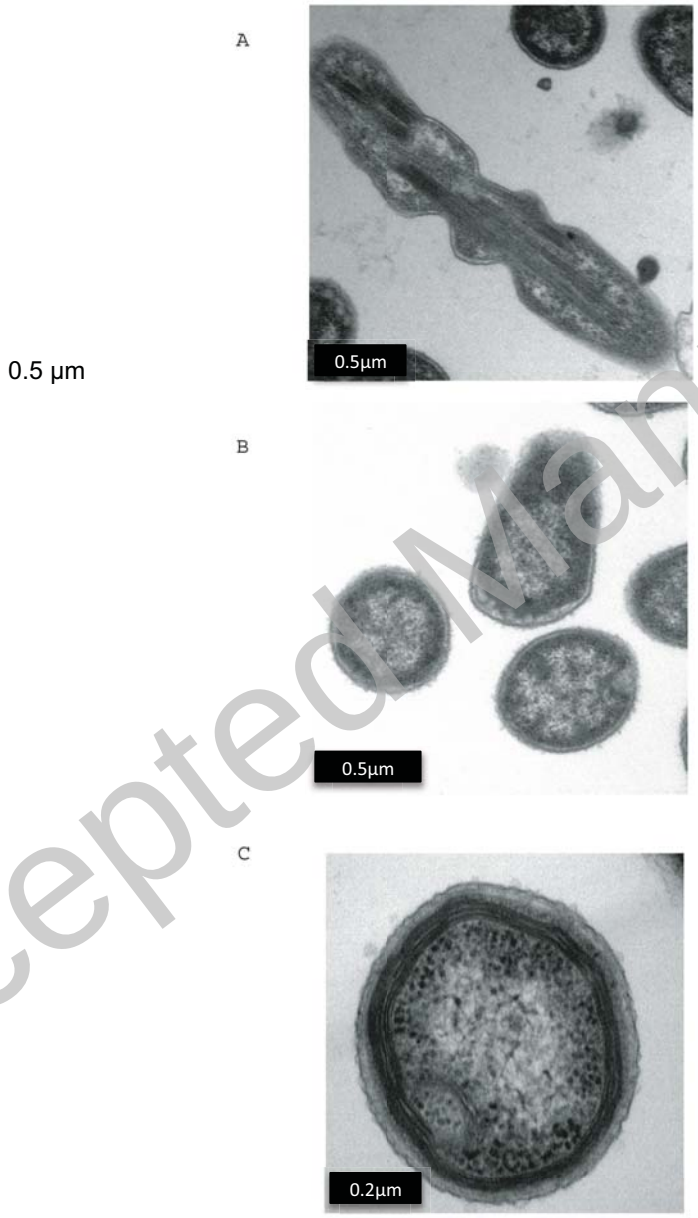


(B)



Biochemical Journal Immediate Publication. Published on 28 Jul 2009 as manuscript BJ20090780

Figure 2



0.5 µm

Figure 3

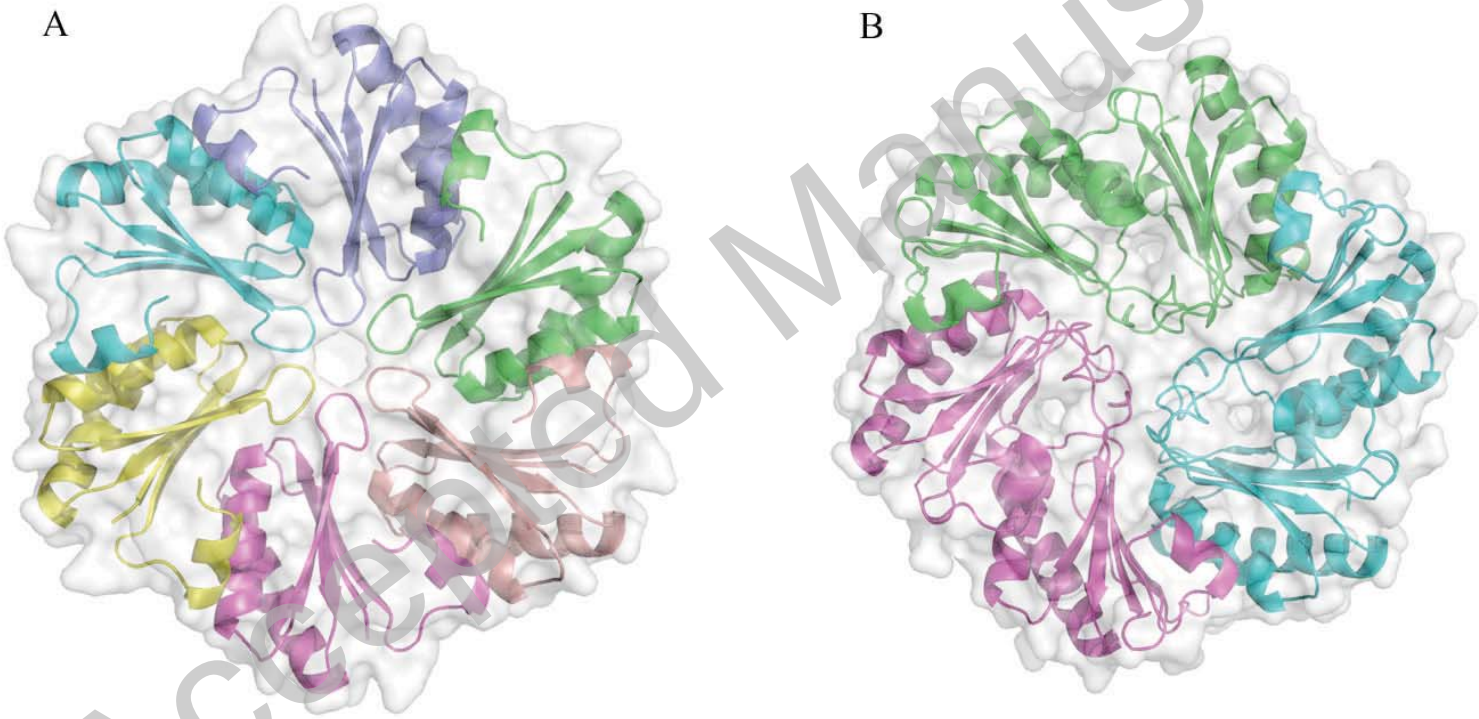
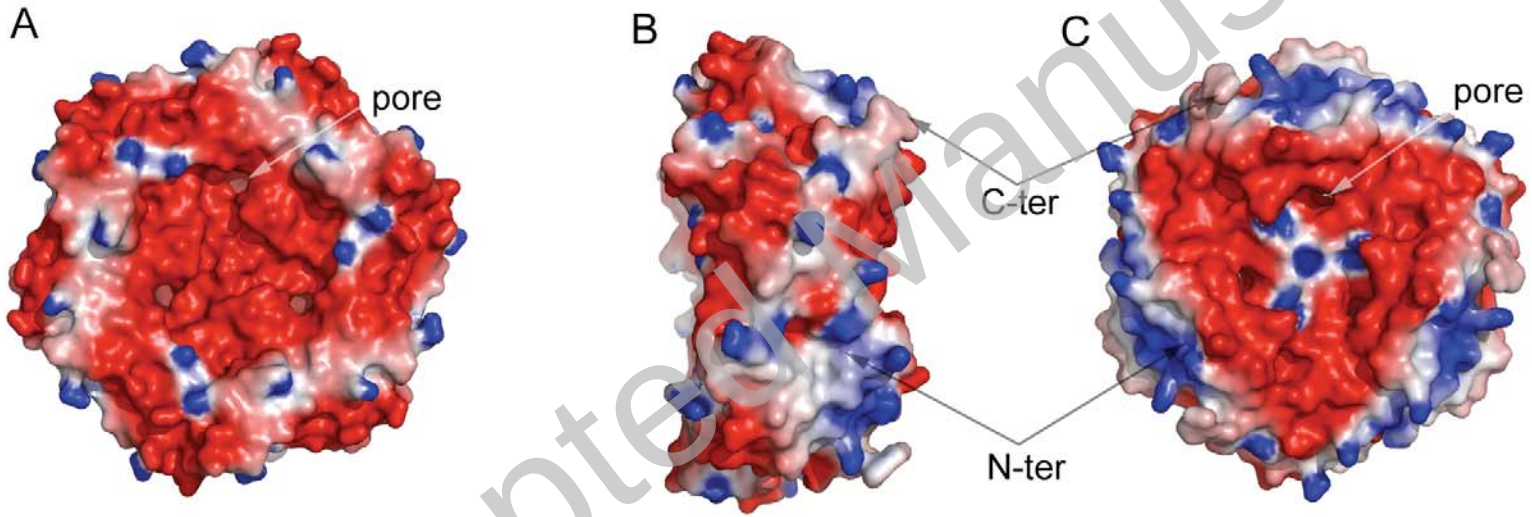


Figure 4

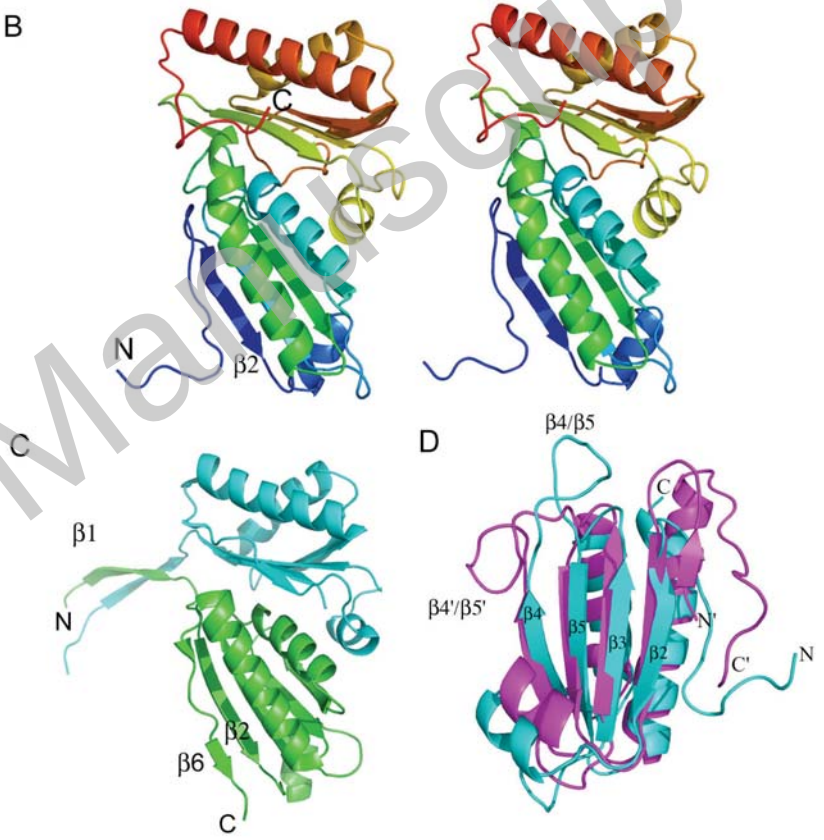


Biochemical Journal Immediate Publication. Published on 28 Jul 2009 as manuscript BJ20090780

Figure 5

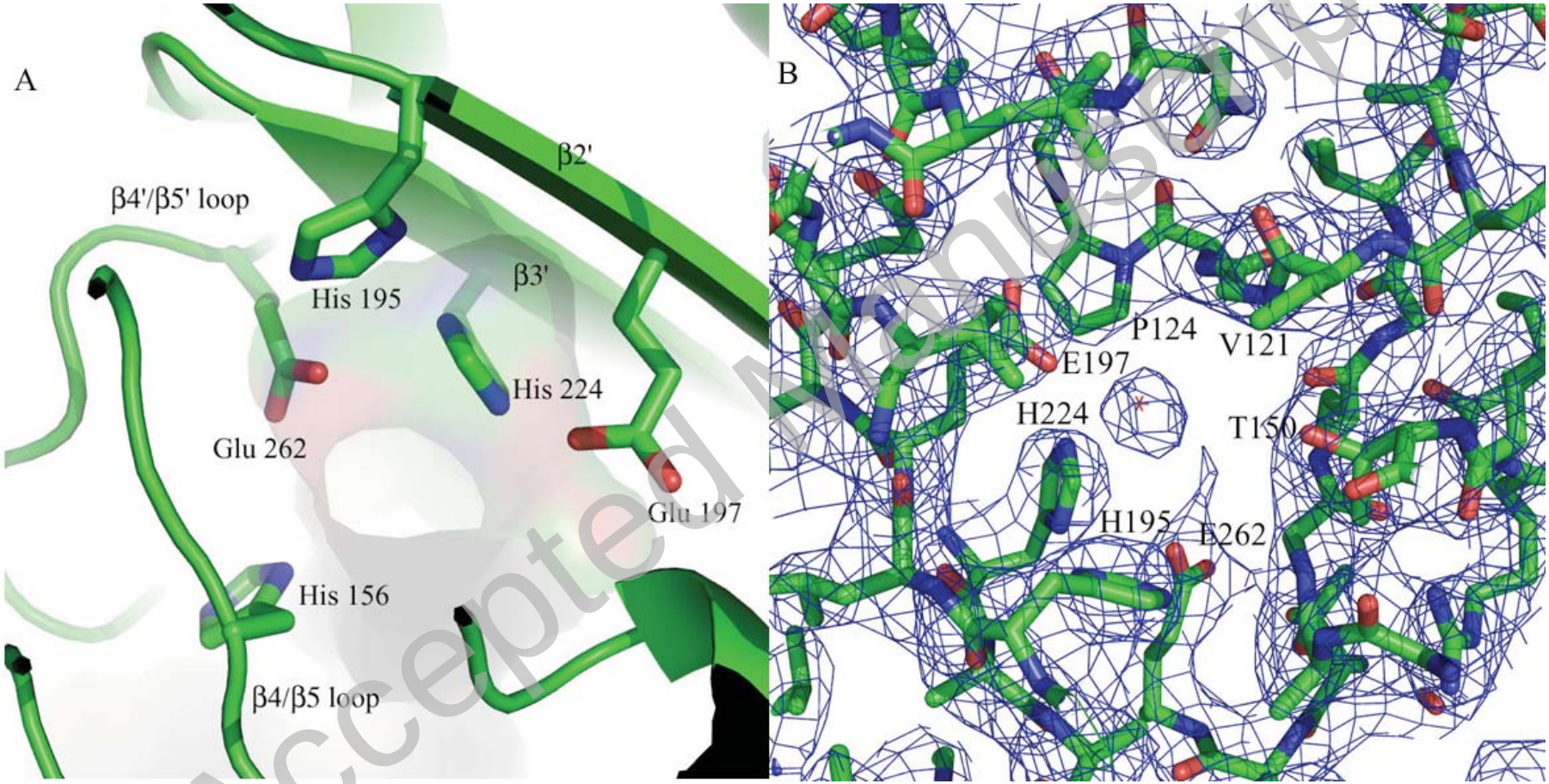
A

```
                                                                    β2
                                                                 EEEEEE    HH
EspB KPDVKVAEESKQAVVEQKANDVKTAPTMTEFVGTAGGDTVGLVIANVDSL 100
PduB ----------------------MAEKSCSLTEFVGTAIGDTLGLVIANVDTA 30
                           .:.: ::******** ***:********:

      α1          β3              α2          β4
     HHHH        EEEEEEE     HHHHHHHHHHHHH   EEEEEE
EspB LHKHLGLDNTCRSIGIISARVGAPAQMMAADEAVKGTNTEVATIELPRDT 150
PduB LLDAMKLEKRYRSIGILGARTGAGPHIMAADEAVKATNTEVVSIELPRDT 80
     * . : *::  *****:.**.** .::********.*****.:*******

            β5              α3          β6′       β2′
          EEEEEEE  HHHHHHHHHHHHHHHHHHHHHEEE    EEEEEE
EspB KGGAGHGIFIVLKAADVSDARRAVEIALKQTDKYLGNVYLCDAGHLEVQY 200
PduB KGGAGHGSLIILGGNDVSDVKRGIEVALKELDRTFGDVYGNEAGHIELQY 130
     ******* :*:* . ****.:*.:*:***: *: :*:**  :***:*:**

           α1′          β3′       α2′              β4′
     E    HHHHHHH       EEEEEE HHHHHHHHHHHHHH    EEEEEE
EspB TARASLIFEKAFGAPSGQAFGIMHAAPAGVGMIVADTALKTADVKLITYG 250
PduB TARASYALEKAFGAPIGRACGIIVGAPASVGVLMADTALKSANVEVVAYS 180
     *****  :******* *:* **: .***.**:::***** :*:*:::*.

                  β5′             α3′
                EEEEEE  HHHHHHHHHHHHHHHHHHH
EspB SPTNGVLSYTNEILITISGDSGAVLQSLTAARKAGLSILRSMGQDPVSMS 300
PduB SPAHGTSFSN-EAILVISGDSGAVRQAVTSAREIGKTVLATLGSEPKNDR 229
     **::*.   . * ::.******** *:*:**: * :* ::*.:* .

EspB KPTF 304
PduB PSYI 233
      . :
```



Biochemical Journal Immediate Publication. Published on 28 Jul 2009 as manuscript BJ20090780

Figure 6



Biochemical Journal Immediate Publication. Published on 28 Jul 2009 as manuscript BJ20090780

Figure 7