REVIEW ARTICLE

# Immunoglobulin domains in *Escherichia coli* and other enterobacteria: from pathogenesis to applications in antibody technologies

Gustavo Bodelón, Carmen Palomino & Luis Ángel Fernández

Department of Microbial Biotechnology, Centro Nacional de Biotecnología, Consejo Superior de Investigaciones Científicas (CNB-CSIC), Campus de Cantoblanco UAM, Madrid, Spain

**Correspondence:** Luis Ángel Fernández, Centro Nacional de Biotecnología, CNB-CSIC, Campus UAM, Cantoblanco, Darwin 3, Madrid 28049, Spain. Tel.: +34 91 585 48 54; fax: +34 91 585 45 06; e-mail: lafdez@cnb.csic.es

Received 30 January 2012; revised 7 June 2012; accepted 14 June 2012. Final version published online 23 July 2012.

DOI: 10.1111/j.1574-6976.2012.00347.x

Editor: Mecky Pohlschroder

**Keywords**
immunoglobulin; *E. coli*; antibody; fimbriae; adhesins; protein folding; pathogenesis.

## Abstract

The immunoglobulin (Ig) protein domain is widespread in nature having a well-recognized role in proteins of the immune system. In this review, we describe the proteins containing Ig-like domains in *Escherichia coli* and enterobacteria, reporting their structural and functional properties, protein folding, and diverse biological roles. In addition, we cover the expression of heterologous Ig domains in *E. coli* owing to its biotechnological application for expression and selection of antibody fragments and full-length IgG molecules. Ig-like domains in *E. coli* and enterobacteria are frequently found in cell surface proteins and fimbrial organelles playing important functions during host cell adhesion and invasion of pathogenic strains, being structural components of pilus and nonpilus fimbrial systems and members of the intimin/invasin family of outer membrane (OM) adhesins. Ig-like domains are also found in periplasmic chaperones and OM usher proteins assembling fimbriae, in oxidoreductases and hydrolytic enzymes, ATP-binding cassette transporters, sugar-binding and metal-resistance proteins. The folding of most *E. coli* Ig-like domains is assisted by periplasmic chaperones, peptidyl–prolyl *cis/trans* isomerases and disulfide bond catalysts that also participate in the folding of antibodies expressed in this bacterium. The technologies for expression and selection of recombinant antibodies in *E. coli* are described along with their biotechnological potential.

## Introduction

Since the initial report of the three-dimensional structure of a human antibody fragment (Poljak *et al.*, 1973), the immunoglobulin (Ig) fold has been found in an increasing number of proteins with diverse biological functions and without an apparent sequence identity (Bork *et al.*, 1994; Halaby & Mornon, 1998; Otey *et al.*, 2009). Initially, the Igs and the proteins involved in the immune response sharing the same fold were classified under the term 'immunoglobulin superfamily' (IgSF; Williams, 1984; Williams & Barclay, 1988). Currently, the IgSF is recognized as one of the largest families in vertebrate genomes, and an analysis of the human genome revealed that the Ig-like domain has the widest representation of any protein domain (Lander *et al.*, 2001), accounting for more than 2% of the total human genes and one of most common structural motifs found in more than 80 protein superfamilies (Srinivasan & Roeske, 2005). The Ig-like domain is amply distributed in nature, and it is present in vertebrates and invertebrates, plants, fungi, parasites, bacteria, and viruses (Halaby & Mornon, 1998). While the general functional role of Ig-like domains is related to binding or molecular recognition processes, the specific reactions mediated by these domains vary widely. According to their functional characteristics, members of the IgSF are classified into eight groups (Halaby & Mornon, 1998): molecular transport, morphoregulation, cell phenotype markers, cell adhesion molecules, virus receptors, shape recognition and toxin neutralization, viral and bacterial molecules and other functions including regulation of gene transcription, cell migration or death cell. The







FEMS MICROBIOLOGY REVIEWS

heterogeneity of this classification reflects the ample functional diversity of the IgSF.

In bacteria, the first examples of Ig-like domains were found in the crystal structures of the proteins PapD of *Escherichia coli* (Holmgren & Branden, 1989) and Cyclodextrin glycosyltransferase of *Bacillus circulans* (Hofmann *et al.*, 1989), thus demonstrating the presence of the Ig-like fold in both Gram-negative and Gram-positive bacteria. The success and the prevalence of the Ig-like domain can be explained by its structural and functional properties such as its stability and resistance to proteolysis. Functionally, it can interact through its different faces and across domains to form homodimers and heterodimers or tandem linear arrays of Ig-like domains. Also, the presence of a single exon coding for most IgSF domains provides the genetic basis for duplication and diversification.

This review summarizes the structural and functional properties of the proteins and organelles containing Ig-like domains in *E. coli* and related enterobacterial species, highlighting their role during cell host and tissue adhesion, invasion, or other steps of the infection process. The folding and assembly of these *E. coli* Ig-like proteins and organelles is discussed, along with the chaperones and protein machineries involved. In addition, we review the expression of heterologous Ig domains in *E. coli*, focusing on the current technologies for expression and selection of small antibody fragments and full-length Ig molecules in this bacterium, given their biotechnological significance for the development of therapeutic antibodies.

## General structural properties and classification of Ig and Ig-like domains

The functional properties conferred by the Ig-like domain rely on the structural characteristics of the molecule that allow a high degree of interaction specificity and diversity (Amit *et al.*, 1986). The Ig-like domain has a three-dimensional structure based on what is called the Ig fold, which is composed of about 70–100 amino acid residues in anti-parallel β-strands designated by letters A, B, C, D, E, F, G in order of appearance in the sequence and organized in two β-sheets that are packed against each other in a β-sandwich (Fig. 1). The β-strands are composed by an alternation of hydrophobic and hydrophilic amino acids with the hydrophobic side chains pointing toward the interior of the molecule. Traditionally, two conserved cysteine residues separated by 55–75 amino acids, and the so-called 'invariant' tryptophan residue located within 10–15 residues C-terminal to the first cysteine were two hallmarks that allowed a putative identification of an Ig-like domain at the primary sequence level (Williams & Barclay, 1988). However, some IgSF domains lack these important features but still adopt an Ig-like fold (Vaughn & Bjorkman, 1996).

Several variants of the Ig-like fold general architecture are found in nature and are classified structurally according to the number of strands, the presence of conserved sequence signatures and the length of the loops interconnecting the β-strands. However, and despite that the Ig-like domains are defined by a common topology and connectivity similar to those of Igs, their amino acid sequences can share as low as 10% sequence identity (Halaby *et al.*, 1999). As an example, the Ig-like structures of the N-terminal domain of the bacterial protein PapD (pdb code: 3dpa), the Fibronectin type-III (FnIII) receptor (pdb code: 1ten), the human growth hormone receptor (pdb code: 3hhr) and the Ig constant (C) 2 domain (pdb code: 7fab) share the same β-sheet domain topology, confirmed by their crystallographic data, although they do not have significant sequence similarity (Vaughn & Bjorkman, 1996). This enormous sequence variability results in a wide range of structural variants responsible for the diversity of function exhibited by members of the IgSF (Chattopadhyay *et al.*, 2009).

The classification of Ig and Ig-like domains has evolved since the first description of the IgSF. Originally, the domains containing an Ig fold were divided into three different topological subtypes V, C1, and C2 (Williams & Barclay, 1988). The structures grouped in the V subtype that includes the variable (V) domain of Igs are generally composed of a back sheet formed by the G, F, C, C′, and C″ strands, and a front sheet formed by the B, E, and D strands. The A strand is shared between the two sheets. The CB, C′C″, and FG strands are connected by loops which correspond to complementarity-determining regions (CDRs) of the V domains of Igs (Fig. 1). The C1 subtype, which includes the constant (C) domain of Igs, contains structures with a back sheet formed by G, F, and C strands and a front sheet formed by A, B, E, and D strands. Protein domains were assigned to the V and C1 subtypes based on the topology of known structures. However, the structural characteristics of the C2 subtype were based on structural predictions. Subsequent structure determinations redefined the C2 domain into the C2 and I subtypes (Fig. 1; Harpaz & Chothia, 1994), and revealed new S (switched) and H (hybrid) subtypes (Bork *et al.*, 1994). Finally, analyzing the sequence and structure of 52 3D structures, Halaby and coworkers classified the Ig-like domains into the C1, C2, C3, C4, V, I, H, and FnIII subtypes (Halaby *et al.*, 1999).

An important structural feature of the classical Ig fold is the presence of a conserved intra-domain disulfide bond connecting the β-strands B and F of opposite β-sheets (Lesk & Chothia, 1982; Fig. 1), which in the case of the Ig domains of the heavy (H) and light (L) chains

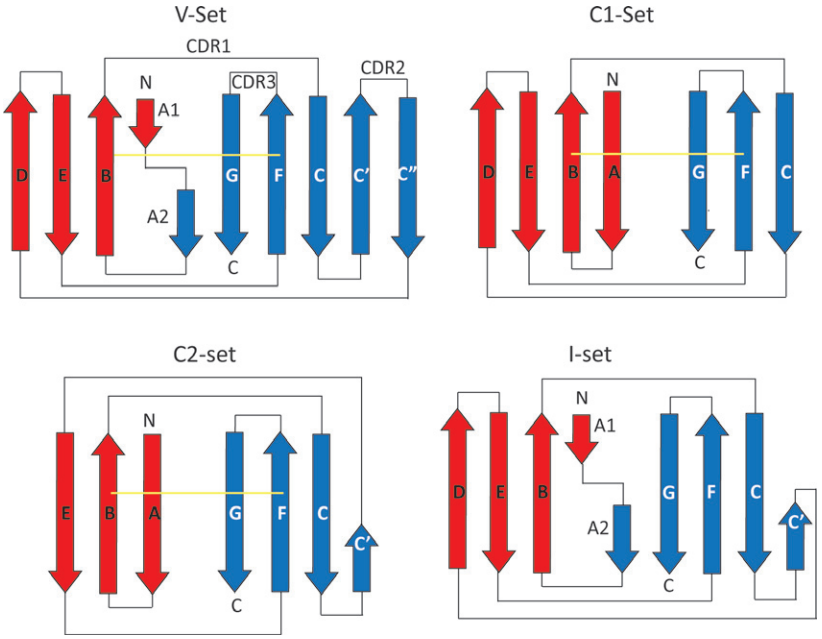


**Fig. 1.** Topology of Ig domains. Topology diagrams and strand nomenclature for the β-sheets of four variants of the Ig fold: variable (V set), constant-1 (C1 set), constant-2 (C2 set) and intermediate (I set). The front and back sheets are depicted in red and blue, respectively, and the presence of a disulfide bond between B and F strands is indicated as a yellow line. The complementarity-determining regions (CDRs) in the V set are also indicated.

of antibodies is formed by the conserved Cys residues at framework positions H22/H92 and L23/L88, respectively (Williams & Barclay, 1988; Worn & Pluckthun, 2001). Whereas this disulfide bond is remarkably well conserved in the V domain of the germline genes of all antibodies (Proba *et al.*, 1998), the number and the location of disulfide bridges in Ig-like structures vary, and when they exist, may connect two strands in the same sheet, a strand and a loop or two loops (Halaby *et al.*, 1999). Both the intra-domain disulfide bond (Proba *et al.*, 1998; Worn & Pluckthun, 1998; Hagihara *et al.*, 2007), and a cluster of hydrophobic amino acid residues that is formed by the packing of the four B, C, E, and F β-strands of the common hydrophobic core (Williams & Barclay, 1988; Bork *et al.*, 1994; Fowler & Clarke, 2001) are responsible for the stability and the folding of Ig and Ig-like domains.

## Origin and evolution of Ig domains

The existence of the Ig-like fold in functionally distinct and phylogenetically distant molecules, such as enterobacterial fimbriae, plant cytochromes or vertebrate antibodies, raises important questions as to when and how the IgSF arose. From an evolutionary standpoint, it is not clear whether the IgSF domains evolved from a common ancestor by vertical or horizontal gene transfer or are the consequence of the drift of independent sequences toward a favorable folding topology (Halaby & Mornon, 1998; Barclay, 2003). It has been suggested that the robust framework of the Ig-like fold is simply the result of an energetically favorable folding (Shapiro *et al.*, 1995). Indeed, the number of possible stable configurations of globular proteins may be limited, and thus, a considerable number of known proteins adopt one of 10 favorable 'superfold' configurations (Orengo *et al.*, 1994, 1997), one of which is the Ig fold (Steiner, 1996).

The fact that many of the IgSF members do not show significant sequence similarity to one another makes very difficult to distinguish between convergent or divergent mechanisms of evolution at such low levels of sequence identity (Klein & Nikolaidis, 2005). Despite of this, it is generally believed that the IgSF derives from a single common ancestor, and examples of putative evolutionary relationships by vertical descendent between bacterial and eukaryotic Ig-like domains can be found in the literature (Bateman *et al.*, 1996; Stevens, 2008).

Interestingly, the structure of the FnIII Ig-like domain of certain carbohydratases (e.g. chitinases) from soil bacteria and that of PapD chaperone is so similar to the eukaryotic FnIII sequences that it has been proposed that these domains were initially acquired from eukaryotes and subsequently spread by horizontal transfers between distantly related bacteria (Holmgren & Branden, 1989; Bork & Doolittle, 1992). Contrary to this, information from bacterial genomes suggested that the FnIII domain might have had its origins in bacteria (Aravind *et al.*, 2003). Along with this hypothesis, the bacterial FnIII domains of ApaG proteins, which are found in a wide variety of bacterial genomes and do not show significant sequence similarity with the FnIII domains of carbohydratases or PapD, may in fact have a bacterial origin (Cicero *et al.*, 2007).

Although the origin and the phylogenetic relationship among the different Ig-like types are controversial, various hypotheses have been proposed to explain the evolution of a primordial domain into its variants by gaining or losing β-strands. Depending on the authors, the original domain might be either the V domain (Williams & Barclay, 1988) or the C2 and I domains (Smith & Xue, 1997; Teichmann & Chothia, 2000). Smith and colleagues proposed that the V and the C1 types appear to be derived from the C2 type, either directly or through the type-I. The authors speculated that the C2 set might be the primordial domain because a more variable and simpler domain is likely to be more ancient (Smith & Xue, 1997). Teichmann and Chothia suggested that the IgSF was originated from the I set because this molecule has structural characteristics that combine structural features of V and C sets (Teichmann & Chothia, 2000). To our knowledge, the phylogeny of the FnIII, C3 and C4 types, which are generally agreed to be in an outgroup position relative to the V, C1, C2 and I sets (Klein & Nikolaidis, 2005), has not been addressed to date. Finally, it has been proposed that both mechanisms of convergent and divergent evolution might explain the IgSF: convergence of unrelated domains toward a simple and stable fold and divergence within each subtype (Halaby *et al.*, 1999). In case of divergent evolution of the domain, it has been suggested that the role of the primordial Ig-like domain was to mediate cell–cell homotypic interactions and, with gene duplication and divergence, developed into more complex heterotypic interactions. It also remains to be elucidated whether the primordial IgSF domain originated in cytosolic proteins or membrane proteins (Tilson & Rzhetsky, 2000; Barclay, 2003).

Interestingly, Ig-like domains have also been identified in certain bacteriophage proteins and bioinformatic analysis of all sequenced phage genomes have revealed that Ig-like domains are widely distributed in phage genomes, as insertions into the coding sequences of surface-exposed structural proteins of tailed double-stranded (ds) DNA bacteriophage particles (*Caudovirales*). Strikingly, no Ig-like domains were found in either ssDNA, RNA phages, or in other classes of dsDNA phages (Fraser *et al.*, 2006). The 68 phage Ig-like domains discovered in 41 genomes belonging to the three families of *Caudovirales* (*Siphoviridae*, *Myoviridae* and *Podoviridae*), were classified into three distinct sequence families of Ig-like domains (I set, FNIII and Big2) defined in the PFAM database (Finn *et al.*, 2008). The phage Ig-like domains appear to be present only in structural proteins such as tail fiber, baseplate wedge initiator, major tail, major head or outer capsid proteins. Although the precise function of these domains is unknown, several lines of evidence suggest that they may play a role in phage infection mediating processes of attachment to host cell (Fraser *et al.*, 2006, 2007; Sathaliyawala *et al.*, 2010). As the members of these bacteriophage families posses very limited sequence similarities to one another, it has been proposed that the presence of similar Ig-like domains among different virus families is attributable to horizontal gene transfer. Fraser and collaborators also hypothesized that *Caudoviridae*, which infect both Gram-negative and Gram-positive bacteria, might have been an important vector in the spread of Ig-like domains through diverse species of bacteria (Fraser *et al.*, 2006).

## Ig-like domains in *E. coli* and enterobacteria

Ig-like domains have been reported in a good number of *E. coli* and enterobacterial proteins, as will be reviewed in the following sections. To find known *E. coli* and enterobacterial proteins bearing an Ig-like fold we performed bibliographic and bioinformatic searches to screen current protein databases. We searched in the structural classification of proteins (SCOP, version 1.75; Andreeva *et al.*, 2008) and in the conserved domain protein family (Pfam; Finn *et al.*, 2008) databases. Next, an estimate of the number of *E. coli* and enterobacteria protein entries containing Ig-like domains was obtained performing a taxonomic search of the accession numbers of the Pfam domains in the Uniprot knowledgebase (UniProtKB) database (UniProt Consortium, 2010).

In Table 1 is represented the Pfam code and accession number of the Ig-like domains identified along with representative examples of polypeptide domains with Ig-like fold. For each Pfam domain, the current number of Uniprot entries found in *E. coli* and enterobacteria is indicated. In addition, a protein example of each category, with its PDB code, the presence or absence of disulfide bonds, function and localization is indicated. As seen in Table 1, the Ig-like domain is amply distributed among *E. coli* and enterobacteria, being present in proteins with different functions (such as fimbrial and afimbrial adhesins, chaperones, transporters, and enzymes among others) and not circumscribed to a particular subcellular location. Ig-like domains are found in proteins in different cellular compartments such as cytoplasm, inner and outer membrane (IM and OM), periplasm and attached to the cell surface or secreted to the extracellular medium. The presence of a disulfide bond is not a requisite for the folding of all Ig domains and is not a conserved feature among enterobacterial Ig-like domains (Table 1). As could be expected given the reducing environment of the *E. coli* cytoplasm (Ritz & Beckwith, 2001), the Ig-like domains from cytoplasmic enzymes

**Table 1.** Immunoglobulin domains in proteins from *Escherichia coli* and other Enterobacteria

| Organelles and proteins | Ig-like domain | Uniprot entries (*E. coli*/ Enterobacteria*) | Protein example (Uniprot accession) | PDB | Disulfide bond in Ig-like domain | Function | Localization |
|---|---|---|---|---|---|---|---|
| Fimbrial adhesins | | | | | | | |
| Pilin domain | *Fimbrial* (PF00419) FimH-C | 4366/8102 | FimA (P04128) | 2JTY | Yes | Adhesion | Cell surface |
| | *PapG_C* (PF03628) PapG-C | 55/59 | | | | | |
| | *Fimbrial_K88* (PF02432) FaeG | 88/259 | | | | | |
| | *Fimbrial_CS1* (PF04449) Cfae-C | 79/190 | | | | | |
| | *Adhesin_Dr* (PF04619) DraE | 60/60 | | | | | |
| | *AfaD* (PF05775) AfaE/D | 36/112 | | | | | |
| | *Antig_Caf1* (PF09255) Caf1 | 2/29 | | | | | |
| | *Saf-Nte_pilin* (PF09460) SafA | 0/59 | | | | | |
| Lectin domain (two-domain adhesins) | FimH_man-bind (PF09160) | 513/589 | FimH (P08191) | 1UWF | Yes | Adhesion | Cell surface |
| | Fim-adh_lectin (PF09222) | 12/21 | | | | | |
| | PapG_N (PF03627) PapG-N | 49/51 | | | | | |
| | Cbld (PF07434) Cfae-N | 72/150 | | | | | |
| Non-fimbrial adhesins | *Big_1* (PF02369) | 656/926 | Intimin (P19809) | 1F00 | No | Adhesion | Cell surface |
| | *Big_2* (PF02368)† | 640/884 | | | | | |
| | *Invasin_D3* (PF09134) | 5/47 | Invasin (P11922) | 1CWV | No | Invasion | Cell surface |
| Fimbrial usher | | | | | | | |
| Plug domains | *Usher* (PF00577) | 1938/4108 | FimD (P30130) | 3RFZ 3HN 1ZDX | No | Translocation of fimbrial subunits | OM |
| C-domains | *PapC_C (*PF13953*)* | 964/1669‡ | FimD (P30130) | 3RFZ 2KT6 | Yes | Translocation of fimbrial subunits | OM |
| Fimbrial chaperones | *Pili assembly_C* (PF02753) | 1947/4071 | PapD (P15319) | 3DPA | Yes | Folding of fimbrial subunits | Periplasm |
| | *Pili assembly_N* (PF00345) | 2206/4630 | PapD (P15319) | 3DPA | No | | |
| Enzymes | | | | | | | |
| Superoxide dismutases and oxidorreductases | Copper/zinc superoxide dismutase *Sod_Cu* (PF00080) | 209/427 | SodC (P0AGD1) | 1ESO | Yes | Physiologic response to oxygen toxicity | Periplasm |
| | Thiol–disulfide reductase DsbD N-terminal $_{D4\text{-}A124}$ domain | ND | DsbD (P33655) | 1L6P | Yes | Electron transfer | IM (periplasmic side) |

**Table 1.** Continued

| Organelles and proteins | Ig-like domain | Uniprot entries (*E. coli*/ Enterobacteria*) | Protein example (Uniprot accession) | PDB | Disulfide bond in Ig-like domain | Function | Localization |
|---|---|---|---|---|---|---|---|
| Sugar-binding enzymes | Glycogen branching enzyme *CBM_48* (PF02922) | 431/969 | GlgB (P07762) | 1M7X | No | Glycogen biosynthetic process | Cytoplasm |
| | *O-Glycosyl hydrolase family 2 Glyco_Hydro_2* (PF00703) | *633/854* | *LacZ (P00722)* | 1DP0 | No | Carbohydrate metabolic process | Cytoplasm |
| | O-Glycosyl hydrolase family 18 *Chitinase A_N* (PF08329) | 4/40 | ChiA (P07254) | 1EDQ | Yes | Chitin degradation | Extracellular |
| | Glucans biosynthesis protein G OpgG C-terminal $_{P397\text{-}V512}$ domain | ND | OpgG (P33136) | 1TXK | No | Biosynthesis of glucans | Periplasm |
| ATP-binding ABC-transporters | Maltose transport system a MalF(P2) $_{N93\text{-}K275}$ domain | ND | MalF (P02916) | 2R6G | No | Maltose uptake | IM (periplasmic side) |
| | Wzm/Wzt transport system C-Wzt $_{M264\text{-}Q431}$ domain | ND | Wzt (Q47591) | 2R5O | No | O-polysaccharide export | IM (cytoplasmic side) |
| Other Ig-like containing proteins | Chitin binding protein | ND | Cbp21 (O83009) | 2BEM | Yes | Chitin degradation | Extracellular |
| | Copper resistance protein *CopC* (PF04234) | 169/408 | PcoC (Q47454) | 1IX2 | No | Copper iron binding | Periplasm |

*Uniprot entries updated at December 2011.

†Big_2 (PF02368) also includes DsDNA viruses (*Caudovirales*, 63 entries) and two enterobacterial phages.

‡Number of sequences obtained from PFAM database and not from Uniprot database. No Uniprot entries yet available.

(C-Wzt, β-galactosidase and the Glycogen branching enzyme) lack disulfide bonds.

## Structure and function of bacterial Ig-like domains

### Fimbrial adhesins, chaperones and ushers

The Ig-like domain is frequently found among fimbrial adhesins, and interestingly, it is also found in the components that constitute their specific secretory and assembly pathway, the chaperone-usher (CU) pathway (Dodson *et al.*, 1993; Waksman & Hultgren, 2009). For recent review see (Thanassi *et al.*, 2012). Fimbrial adhesins are proteinaceus filamenteous appendages assembled on the bacterial surface that mediate attachment to host cells and tissues (Proft & Baker, 2009). The word fimbriae (fimbria in singular) is generally used to refer to protein threads or fibers located on the bacterial surface. These adhesive organelles are made of noncovalently linked pilin subunits that are constituted by the Ig-like structure. Interestingly, the polymerization of these Ig-like subunits on the surface of the bacteria enables the adhesin to reach the target located on eukaryotic cells or abiotic surface. The morphology of these organelles can vary from thin flexible polyadhesive fibrils (2 nm of diameter) to thick rigid monoadhesive organelles (7–10 nm of diameter). Consequently, fimbriae have been divided into two main structural and functional families (Zavialov *et al.*, 2007; Proft & Baker, 2009): the 'adhesive pilus' and the 'nonpilus adhesins' (Supporting information, Table S1).

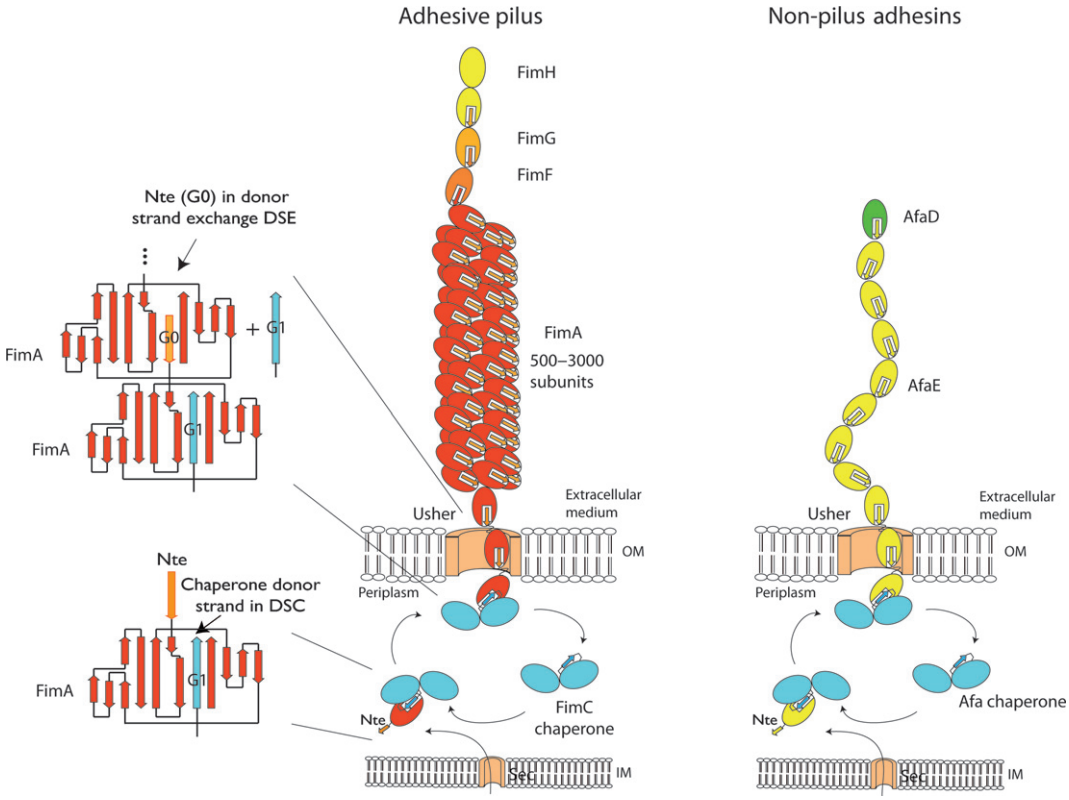


**Fig. 2.** Fimbrial adhesins carrying Ig-like domains in *Enterobacteriaceae*. Schematic representation of adhesive pilus (left) and nonpilus adhesins (right) and their assembly by the CU pathway. Type 1 pili (left) and Afa adhesins of *Escherichia coli* (right) are examples of these two families. The subunits that constituted these fimbrial adhesins are first translocated to the periplasm through the Sec translocon located in the inner membrane (IM). Afterward, the subunits are recognized by the periplasmic chaperone and are delivered to the usher in the OM. In the figure, FimC (left) and Afa chaperone (right) are examples of the periplasmic chaperones. Because each pili subunit is folded in an incomplete Ig-like domain (down and left, exemplified by FimA) the specific chaperone donates a β-strand (G1) that stabilizes the Ig fold. This interaction is known as DSC. Once in the usher, the pili subunits suffer an exchange of the chaperone G1 strand by an N-terminal extension (Nte or G0) of the next subunit that is going to be incorporated in the fimbrial adhesin (illustrated on the left with FimA-FimA polymerization). This reaction is known as donor strand exchange (DSE).

The 'adhesive pilus' (pili in plural) family comprises fimbriae that have a small number of different subunits at various stoichiometries and typically displaying only one specialized adhesive domain (adhesin) on the tip of the pilus fiber. The most extensively characterized examples of this family are the *E. coli* Type 1 and P-pili from uropathogenic *E. coli* (UPEC) strains. The Type 1 pilus (Fig. 2) is a 6.9 nm wide and 1–2 μm long helical rod formed by a right-handed helical array of 500–3000 copies of the main structural pilin subunit FimA connected via FimF to a 3 nm wide tip fibrillum containing FimG and the adhesin FimH (Krogfelt *et al.*, 1990; Hahn *et al.*, 2002). The P pilus is a 6.8 nm wide and several micrometers long right-handed helical cylinder composed of *c.* 1000 copies of the major structural protein, PapA, attached to the OM by a minor structural protein, PapH, and terminated by a 2–3 nm wide tip fibrillum containing the PapG adhesin and the three minor pilin proteins PapE, PapF and PapK (Kuehn *et al.*, 1992; Bullitt & Makowski, 1995).

The second fimbrial family comprises the 'nonpilus adhesins' that contain one or two types of subunits, being generally the main structural subunit the one implicated in adhesion. These polyadhesive fimbriae typically present amorphous or capsule-like morphology at low resolution (Zavialov *et al.*, 2007). Relevant examples of nonpilus adhesins are the Afa/Dr adhesins from *E. coli*, the polymeric F1 capsular antigen of *Yersinia pestis* and the atypical Saf fimbriae of *Salmonella enterica*. Dr adhesins from UPEC and diffusely adherent *E. coli* (DAEC) strains, are constituted by numerous copies of a single-domain adhesin (i.e. AfaE) that assemble in a thin, flexible, filamentous polymer capped with a different subunit (i.e. AfaD) at the tip that mediates invasion of the host cell

(Anderson *et al.*, 2004a, b; De Greve *et al.*, 2007; Knight & Bouckaert, 2009; Fig. 2). The F1 capsular antigen of *Y. pestis* consists of high-molecular weight polymers built from a single protein subunit (Caf1) that confers, as an additional role, protection to the bacteria against phagocytosis (Zavialov *et al.*, 2002). Saf fimbriae are composed by two subunit types, SafA and SafD. Whereas SafA forms the major pilus subunit and adhesin, SafD is classified as a putative invasin, based on sequence similarity with AfaD, and is predicted to be localized at the tip of the protein fiber (Strindelius *et al.*, 2004; Salih *et al.*, 2008).

Importantly, in addition to the adhesive functions that the Ig-like domain confers to fimbriae, their biogenesis also relies on the Ig-like domains found in each fimbrial subunit, in the dedicated fimbrial chaperones found in the periplasm, and in the fimbrial ushers involved in the translocation of these organelles across the OM (Sauer *et al.*, 2004; Remaut *et al.*, 2008; Waksman & Hultgren, 2009; Phan *et al.*, 2011). During the biogenesis of the fimbrial adhesins, the subunits are first secreted to the periplasm, via the general secretory pathway (Sec-system; Driessen & Nouwen, 2008), where a specific fimbrial chaperone assists their folding and prevents premature assembly of the subunits (Barnhart *et al.*, 2000; Vetsch *et al.*, 2004). Because each fimbrial subunit presents an incomplete and unstable Ig-like fold, this fold has to be completed and stabilized by an extra β-strand donated by the chaperone, a reaction known as donor strand complementation (DSC; Sauer *et al.*, 1999; Waksman & Hultgren, 2009). After DSC, the chaperone-subunit complexes are delivered to the OM usher, which recruits the complexes and catalyzes polymerization of subunits. Subunit polymerization is based on a concerted interaction between incomplete Ig-like domains, in this case of two consecutive fimbrial subunits, in the usher. In a reaction known as donor strand exchange (DSE) the Ig fold of the fimbrial subunit is completed by a N-terminal extension of the incoming subunit that replaces the β-strand donated by the chaperone (Fig. 2; Sauer *et al.*, 2002; Barnhart *et al.*, 2003; Remaut & Waksman, 2006; Vetsch *et al.*, 2006).

As mentioned previously, Ig-like domains are also found in both the periplasmic chaperone and the OM usher. Whereas the chaperone is constituted by two complete Ig-like domains (the N and C domains; Holmgren, 1989; Choudhury *et al.*, 1999; Waksman & Hultgren, 2009), the usher, in addition to the translocation channel, contains four complete Ig-like folds distributed among their periplasmic soluble domains (the N- terminus, C-terminus and the plug domain; Remaut *et al.*, 2008; Yu *et al.*, 2009; Ford *et al.*, 2010; Phan *et al.*, 2011). The structures and known functions of these Ig-like domains will be described later in this review.

## Pilin domains and the N-terminal domain of two domain adhesins

Multiple crystal structures of chaperone–subunit complexes have been reported, FimC–FimH (Choudhury *et al.*, 1999), PapD–PapK, PapD–PapE and PapD–PapA (Sauer *et al.*, 1999, 2002; Verger *et al.*, 2008), Caf1M–Caf1 (Zavialov *et al.*, 2003, 2005) and SafB–SafA (Remaut *et al.*, 2006). Structural data is also available for adhesive fimbrial subunit complexes FimF-FimG (Gossert *et al.*, 2008), FimA-FimA (Puorger *et al.*, 2011), the P pilus rod subunit PapA (Verger *et al.*, 2007), the tip of Type 1 fimbria (Le Trong *et al.*, 2010), and fimbrial polyadhesins subunits AfaE/DraE, DraD, DaaE, and SafA (Anderson *et al.*, 2004a, b; Pettigrew *et al.*, 2004; Cota *et al.*, 2006; Jedrzejczak *et al.*, 2006; Korotkova *et al.*, 2006; Remaut *et al.*, 2006). These studies have revealed that the structural components of fimbrial adhesins present two general types of Ig-like folding: an incomplete Ig-like domain or 'pilin domain' and a complete Ig-like fold in the N-terminal domain of two domain adhesins.

The incomplete Ig-like domain is present in all the structural subunits, this is, in the pilin subunits of adhesive pili (e.g. PapA, FimA) and in nonpilus adhesin fimbriae (e.g. AfaE, DraD, DaaE, Caf1, SafA), in addition it is also found in the C-terminal domain of the tip adhesins of pili (e.g. FimH, PapG). The incomplete Ig-like fold is composed of six anti-parallel β-strands (A–F) that form a β-sandwich with two β-sheets. Hence, the seventh and last strand (G) of Ig domains is missing in these incomplete Ig-like domains (Fig. 3a). The fimbrial structural subunits, with the exception of the tip subunits, contain an N-terminal extension (Nte) peptide. This disordered tail will act as a complementing G β-strand to complete the Ig fold of the previous subunit. Strands B, E and D form one β-sheet packed against the second one in such a way that the hydrophobic side chains of the two sheets face each other. This produces, in the absent of the seventh strand, a large hydrophobic groove on the side of the pilin subunit (Choudhury *et al.*, 1999; Gossert *et al.*, 2008). The second β-sheet consists, in the case of the subunits from the adhesive pili, of strand A, the N-terminal donor strand G and strands F and C. In the case of the subunits of nonpilus adhesin fimbriae, this second β-sheet does not include the A strand. Independently on the fimbriae family type, the Nte extension of pilus subunits (corresponding to G strand) contain a set of well-conserved alternating nonpolar residues (named as P1–P5 residues) that will occupy the groove of the subunit in positions known as P1–P5 pockets or sites (Remaut *et al.*, 2006; Waksman & Hultgren, 2009). The residue localized at P4 is a strictly conserved Gly, as it is the only one that can avoid steric constrains in the

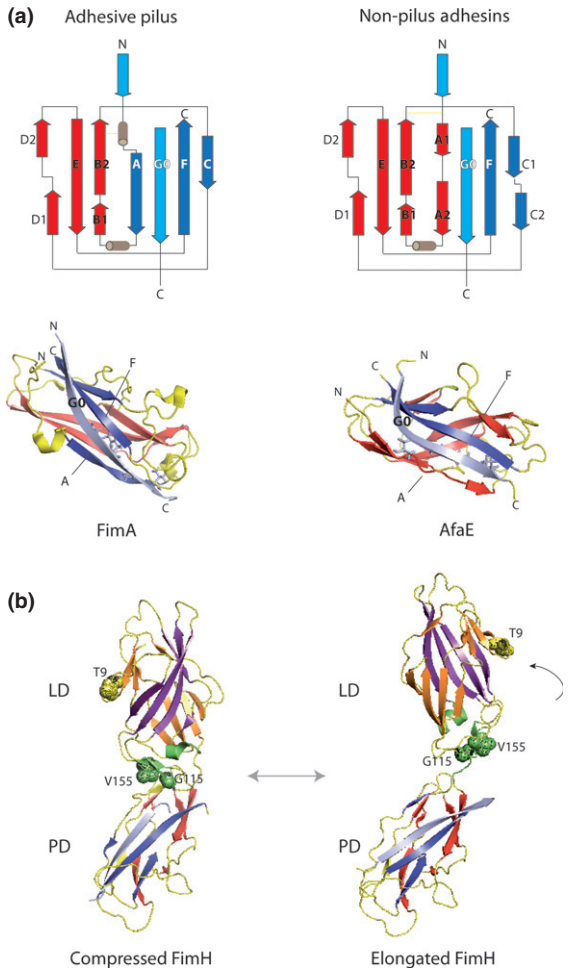


**Fig. 3.** Ig-like domains present in the subunits of fimbrial adhesins. (a) Incomplete Ig-like domain. FimA of Type 1 pili (left) and AfaE pili subunits of Dr adhesins (right) are examples of adhesive pilus and nonpilus adhesins, respectively. The incomplete Ig-like domain of these subunits is constituted by six β-strands (a–f) organized in two β-sheets (in red and in blue in the figure). A hydrophobic groove on the subunits is left because of the lack of the seventh β-strand of complete Ig domains. This hydrophobic groove is stabilized by the periplasmic chaperone or the next subunit incorporated in the fimbriae, which donate a β-strand (G1 of the chaperone and G0 of fimbriae subunits, represented in light blue). The 3D structures of FimA (pdb: 2JTY) and AfaE (pdb: 1RXL) show the incomplete Ig-like domain completed by the Nte of an incoming FimA (G0, light blue, on the left) or an incoming AfaE subunit (G0, light blue, on the right), respectively. (b) Complete Ig-like domains. Structure of FimH as an example of two domains adhesins carrying one complete Ig-like domain (lectin domain, LD) and one incomplete Ig-like domain (pilus domain, PD). The two β-sheets that constituted the Ig domains are shown in blue and red for the incomplete Ig-like domain and in orange and purple for the complete Ig-like domain. In the figure, it is showed the movement of residues Valine 155 and Glycine 115 after the interaction of the lectin domain with its receptor. Images generated with PyMOL program (Molecular Graphics Systems, LLC).

groove and properly adjusts at the corresponding position of P4 pocket (Gossert *et al.*, 2008).

The complete Ig-like fold is present in the N-terminal domain of the tip adhesins of fimbriae and pilus (e.g. F17-GII, PapG, FimH and CfaE). These proteins are two domain adhesins consisting of a N-terminal (carbohydrate) receptor-binding domain (lectin domain) and a C-terminal pilin domain joined by a short interdomain linker. Whereas the C domain presents an incomplete Ig fold of structural pilin domains, the N-terminal domain has a complete Ig fold (De Greve *et al.*, 2007) with variable number of β-sheets. Beside the difference in the number of β-sheets that constitutes the lectin domain, all of them encode a functional carbohydrates binding site, named sugar-binding pocket (Fig. 3b; De Greve *et al.*, 2007). For instance, in the case of FimH, the N-terminal region consists in an elongated Ig-like fold with 11 β-strands (Choudhury *et al.*, 1999) that mediate mannose-specific adhesion (Bouckaert *et al.*, 2005; Le Trong *et al.*, 2010). In the case of F17-G, the overall fold of the N-terminal domain consists in a β-sandwich with a back sheet of five anti-parallel β-strands and a front sheet of four anti-parallel strands that enable the bacteria to attach to *N*-acetylglucosamine (GlcNAc; Buts *et al.*, 2003). The PapGII N-terminal domain is larger than that of FimH and F17-G and is composed by a central antiparallel β-sheet of six strands flanked by two double-stranded β-sheets on one side, and by an α-helix on the other side, bearing in this sugar-binding structure the site for Galα1-4Gal (Dodson *et al.*, 2001).

An additional surprising feature of two domain adhesins has been recently described once the structure of the whole fimbrial tip of Type 1 fimbriae was solved (Le Trong *et al.*, 2010). This consists on an important conformational change of the N-terminal domain of FimH when it binds to its carbohydrate receptor at the tip of the fimbria. In the absence of mannose, the N-terminal domain of FimH is maintained in a compressed conformation against the C-terminal domain, being the sugar-binding pocket present in a loose (open) form and the short interdomain linker trapped by both domains (Fig. 3b, left). However, when FimH is bound to its ligand, the sugar-binding pocket is tighten around it, causing a switch from the compressed conformation to an elongated one making the N-terminal domain to be separated from the C-terminal domain (pilin domain) and the linker to be exposed (Le Trong *et al.*, 2010; Kisiela *et al.*, 2011; Tchesnokova *et al.*, 2011). This interesting phenomenon is proposed to be the basis for a mechanical force regulation of FimH adhesion (Aprikian *et al.*, 2011) and for the allosteric regulation of the pilin domain by which FimH is maintained in a low affinity state through internals contacts (Le Trong *et al.*, 2010). Upon interac-

tion with the ligand and/or tensile-force, both domains are separated from each other and the lectin domain untwists and holds tightly the ligand acquiring the elongated conformation of high affinity (Fig. 3b, right).

## Periplasmic fimbrial chaperones

After release from the Sec translocon, pilin subunits are recognized by a dedicated periplasmic chaperone. This interaction not only avoids their aggregation and degradation but also maintains the subunits in a polymerization-prone folding state until the complex is targeted to the OM usher (Jones *et al.*, 1997; Vetsch *et al.*, 2004). The structures of several periplasmic chaperones are known: PapD (P-pili; Sauer *et al.*, 1999), FimC (Type 1 fimbria; Pellecchia *et al.*, 1998; Choudhury *et al.*, 1999), SfaE (S pili; Knight *et al.*, 2002), Caf1M (F1 capsule; Zavialov *et al.*, 2003), FaeE (F4 Fimbria; Van Molle *et al.*, 2005), SafB (Saf Pilus; Remaut *et al.*, 2006). These structures reveal that they consist of two Ig-like domains joined at 90° angle, which are separated by a large cleft that accommodates the bound subunit (Holmgren & Branden, 1989; Holmgren *et al.*, 1992; Kuehn *et al.*, 1993). The Pfam database has classified the domains as Pili assembly_N and _C. As an example, the PapD domain 1 (Pili assembly_N) is located at the N-terminus of the protein and folds as a β-sandwich of seven anti-parallel β-strands that is very similar to that of the V domain of Igs (Fig. 4). The C-terminal domain of PapD (Pili assembly_C) presents an Ig-like β-sandwich of eight anti-parallel β-strands in which the H strand is disulfide bonded to strand G. However, as the H strand is not always present in PapD homologs it should not be considered as a conserved feature of periplasmic chaperones (Holmgren *et al.*, 1992).

Structural data of pilin-chaperone complexes show that the interactive surfaces between the pili subunits and the chaperone are formed two specific areas which include: (1) two conserved basic residues located in the large cleft left between the two Ig-like domains of the chaperone; and (2) by the first and the seventh β-strands of the N-terminal Ig-like domain (A1 and G1, respectively). As mentioned above, the incomplete Ig fold of pilin subunits left a hydrophobic groove in the barrel between strands A and F that makes these polypeptides unstable in the periplasm unless they interact with the chaperone. In DSC the chaperone donates a motif of four alternating hydrophobic residues (P1–P4 residues) of its G1 β-strand to the subunit, thus capping the hydrophobic groove (P1–P4 pockets) and stabilizing the Ig-like structure of the subunit. This reaction produces a complete, noncanonical, Ig-like fold by which the chaperone G1 β-strand aligns parallel to the F strand of the pilin subunits (Fig. 4a and b).

On the basis of conserved structural features found in the flexible loop that connects the F1 and G1 β-strands of the chaperone N domain, Hung and collaborators classified the periplasmic chaperones into two distinct classes: the FGL (F1G1 long) and FGS (F1G1 short) chaperones (Hung *et al.*, 1996). Interestingly, these two groups of

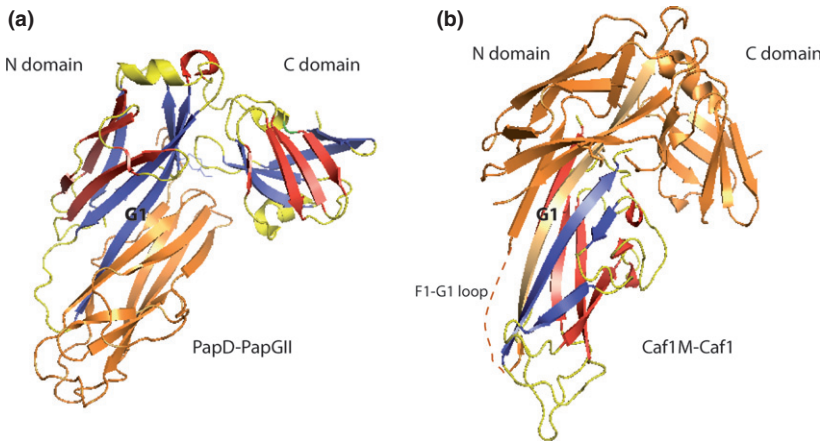


**Fig. 4.** The periplasmic fimbrial chaperones. (a) PapD chaperone complementing one of the pilus subunits of Pap-pili, PapGII (from PDB: 3M0). The N and C-domains of PapD chaperone consist of two Ig-like domain joined at 90° angle. In the figure, the two β-sheets that constitute the Ig domains of PapD chaperone are showed in red and in blue. The complementing G1 strand of the chaperone is labeled. (b) Structure of Caf1M chaperone as an example of FGL chaperone complementing a Caf1 pilus subunit (from PDB: 1Z9S). In this case, the two β-sheets that constitute the Ig-like domain of Caf1 pilus subunit are showed in red and in blue. The complementing G1 strand and the unstructured long F1–G1 loop (dashed line) of the chaperone are labeled. Images generated with PyMOL program (Molecular Graphics Systems, LLC).

chaperones assemble the two groups of adhesive pilus and nonpilus adhesive fimbriae, respectively (Soto & Hultgren, 1999; Waksman & Hultgren, 2009; Zav'yalov *et al.*, 2010). FGL chaperones contain a significantly longer G1 strand and longer F1-G1 loop stabilized by a disulfide bridge between two conserved Cys residues of F1 and G1. In addition, FGL chaperones also contain a longer binding motif at the N-terminus, which extends the A1 strand by at least three residues (Zavialov *et al.*, 2007). G1 strand of FGL chaperones contain an additional alternating hydrophobic residue (P5 residue) that inserts intermittently into a corresponding P5 pocket of the groove, which is never occupied in the case of FGS chaperones.

During subunit polymerization, the complementing G1 strand of the chaperone is replaced by the Nte extension on the incoming subunit. This reaction is known as DSE and takes place in the OM usher (Saulino *et al.*, 1998; Thanassi *et al.*, 1998, 2012; Sauer *et al.*, 2004; Vetsch *et al.*, 2006; Waksman & Hultgren, 2009). In contrast to the complementation achieved by the chaperone, the canonical Ig fold of the subunits is completed by the Nte strand in an anti-parallel orientation to the F β-strand (Fig. 2; Sauer *et al.*, 2002). DSE is proposed to occur through a concerted 'zip in-zip out' mechanism that is initiated by the insertion of the Nte P5 residue into the P5 pocket of the groove of the previous subunit bound to the usher (Zavialov *et al.*, 2003; Remaut *et al.*, 2006). Once the P5 pocket is occupied by Nte P5 residue of the next subunit, a transient ternary chaperone-subunit-subunit complex is formed allowing a gradual displacement of the chaperone G1 strand (Remaut *et al.*, 2006; Waksman & Hultgren, 2009). In addition to the initiating role in DSE of the P5 pocket, the accessibility of Nte to this pocket of specific subunit also marks DSE kinetics and the termination of the biogenesis of adhesive organelles (Rosen *et al.*, 2008; Verger *et al.*, 2008). As an example, in the P pilus system termination of the pilus depends on the last subunit PapH that lacks the P5 pocket. This subunit undergoes DSE with the last PapA subunit but is unable to undergo DSE with any other subunit, resulting in the truncation of pilus assembly (Verger *et al.*, 2006).

## Ig domains in OM fimbrial ushers

The polymerization of adhesive organelles takes place at the OM usher. This membrane protein not only recruits chaperone-subunit complexes but also coordinates the mentioned DSE reaction and translocates the growing adhesive organelle to the OM (Waksman & Hultgren, 2009). Ushers are integral OM proteins of *c.* 800 amino acid residues that comprise four functional domains: two periplasmic soluble N-terminal and C-terminal domains that interact with the pilus subunit-chaperone complexes, a β-barrel domain that constitutes the translocation channel and a plug domain, located in the middle of the channel, occluding the lumen in resting ushers, or underneath the translocation domain in active ushers (Fig. 5; Remaut *et al.*, 2008; Huang *et al.*, 2009; Phan *et al.*, 2011). The N-terminal domain of ushers has been shown to form the initial binding site for subunit-chaperone complexes (Nishiyama *et al.*, 2003, 2005; Ng *et al.*, 2004) with a very fast recruiting activity (Nishiyama & Glockshuber, 2010). Using crystallographic and nuclear magnetic resonance (NMR) to characterize the N-terminus domain of FimD usher from Type 1 fimbriae (Nishiyama *et al.*, 2005), it has been revealed that this domain is composed of a disordered N-terminal tail that becomes structured upon binding of chaperone-subunit complex, a folded core of six β-strands and a hinge segment to the β-barrel domain. Interestingly, in adhesive pilus such as Type 1 fimbriae, this binding features different affinities depending on the type of subunit-chaperone complex that parallel the relative position in the polymerized fimbriae (Nishiyama *et al.*, 2003; Ng *et al.*, 2004; Li *et al.*, 2010; Nishiyama & Glockshuber, 2010).

The three-dimensional structure of the C-terminal region of PapC usher (Ford *et al.*, 2010) and the structure of the full-length FimD usher (Phan *et al.*, 2011) have been solved, indicating that the C-terminal regions of these ushers are composed by two Ig-like domains of seven β-strands, named CTD1 and CTD2 (Fig. 5a; Phan *et al.*, 2011). These C-terminal domains also interact with subunit-chaperone complexes (Thanassi *et al.*, 2002; So & Thanassi, 2006; Phan *et al.*, 2011). *In vivo* and *in vitro* experiments carried out with deletion mutants indicate that this region is necessary for the organelle assembly (So & Thanassi, 2006; Huang *et al.*, 2009).

Structure of the embedded OM part of PapC and FimD showed that the β-barrel domain of ushers is composed by 24 β-strands (Remaut *et al.*, 2008; Huang *et al.*, 2009; Phan *et al.*, 2011). This translocation channel presents a kidney-shaped with a pore size large enough for the passage of individual folded subunit but not for any chaperone-subunit complexes (Fig. 5b; Huang *et al.*, 2009). The plug domain is inserted into the loop connecting β6 and β7 strands and is positioned laterally inside the β-barrel gating the lumen of the pore (in resting ushers) or underneath the translocation channel (in actively translocating ushers; Fig. 5b; Remaut *et al.*, 2008; Huang *et al.*, 2009; Phan *et al.*, 2011). The crystal structure of the plug domains in FimD and PapC ushers indicated that they are composed by six β-strands with a fold similar to an Ig fold (Remaut *et al.*, 2008; Huang *et al.*, 2009; Phan *et al.*, 2011). Crystallization of the isolated plug domain of Caf1A usher suggested that this domain

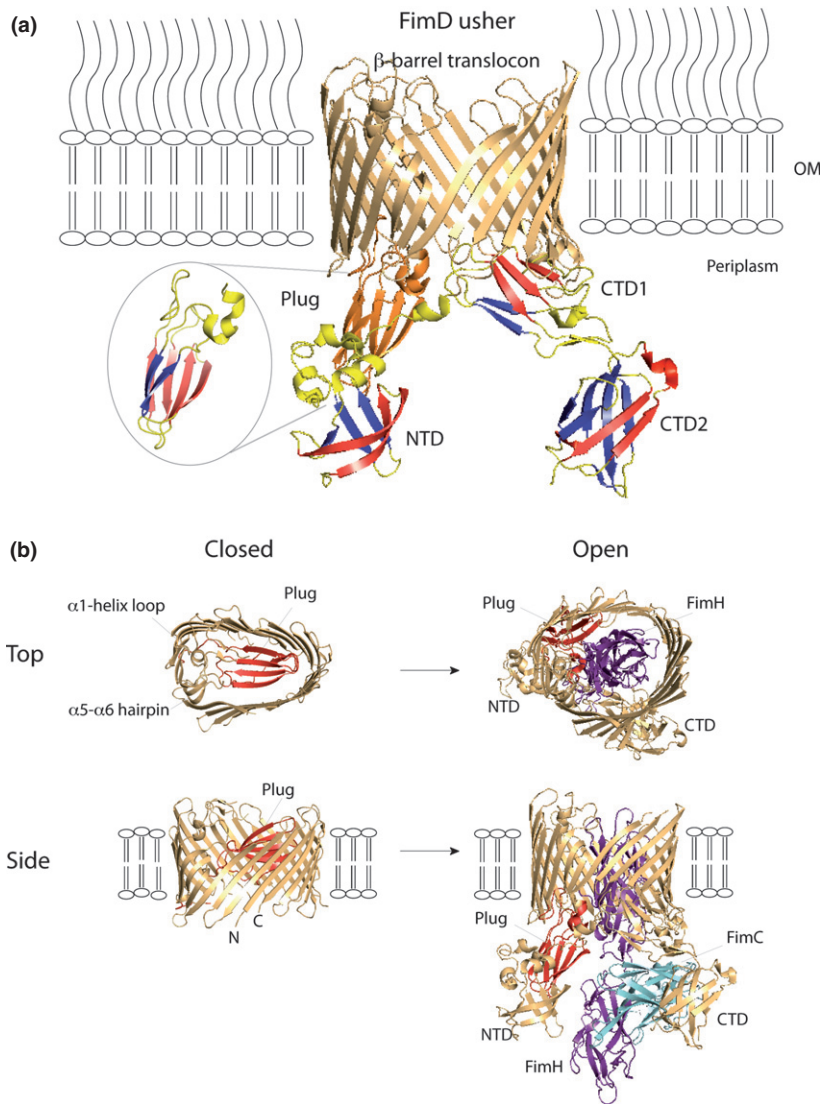


**Fig. 5.** The OM usher of fimbrial adhesins. (a) Structure of FimD protein as representative of fimbrial ushers. In the 3D structure of FimD (from PDB: 3RFZ) are indicated the Ig domains present in fimbrial ushers, the C-terminus domains 1 and 2 (CDT1 and CTD2) and the plug domain. The two β-sheets that constituted the Ig domains are shown in red and in blue. In addition, the β-barrel translocation domain and the N-terminus domain containing a β-strand structure are also shown. (b) Conformational change of fimbrial ushers shown with the structure of FimD-FimH-FimC complex (from PDB: 3RFZ) and FimD alone (from PDB: 3OHN). The plug domain (in red) containing an Ig-like domain moves from inside the lumen of the channel (left, in closed state) to a position beneath the translocation domain when the channel is open (right, in open activated state). Top and side views illustrate the rotation of the plug domain outside the channel and enable to show other structures inside the lumen of the channel, the β5-6 hairpin and the α-helix. Images generated with PyMOL program (Molecular Graphics Systems, LLC).

could be composed of seven β-strands as a result of the swapped dimerization of two monomers. Each monomer has the classical s-type Ig fold constituted by six β-strands and a seventh additional β-strand coming from the N-terminus of the second monomer (Yu *et al.*, 2009). Interestingly, mutations located in the corresponding N-region of the plug domain of PapC abolish the adhesive capacity of the bacteria to eukaryotic cells, suggesting the loss of P fimbriae (Henderson *et al.*, 2004). From *in vitro* and *in vivo* experiments it is deduced that the plug domain not only acts merely as a barrier occluding the translocation channel, but it appears to be involved in the polymerization process (Huang *et al.*, 2009; Mapingire *et al.*, 2009; Yu *et al.*, 2009).

The three-dimensional structure of FimD bound to its cognate FimH-FimC substrate (Phan *et al.*, 2011) revealed that the usher suffers an important conformational change from a closed nonactive state to an open active

one (Fig. 5b). Although this conformational change implies the movement of the plug domain from the lumen of the channel to a position beneath the translocon, the mechanism by which this conformational change occurs is not yet known. However, in some of the adhesive fimbriae system it has been identified a natural activator of the usher. This is the case of the type 1 fimbriae, in which it has been shown that FimH is involved in the activation of FimD usher (Nishiyama *et al.*, 2008). Importantly, it has been demonstrated that the lectin N-terminal domain of FimH, in addition to its adhesive capacity, is required for recognition of FimH by FimD and is essential for activation of the usher (Munera *et al.*, 2007, 2008; Nishiyama *et al.*, 2008). An interesting possibility is that the N-terminal domain of FimH, being the activator, could trigger the initial displacement of the plug domain to leave free space for its passage across the lumen of FimD.

## Nonfimbrial OM adhesins: the family of intimins and invasins

The nonfimbrial adhesins containing Ig-like domains belong to the family of intimins in strains of enteropathogenic *E. coli* (EPEC), enterohaemorrhagic *E. coli* (EHEC), *Citrobacter* spp., and *Hafnia alvei* and invasins found in strains of *Yersinia* spp. Intimins and invasins are large OM proteins (OMPs) of *c.* 900 amino acids related to each other in terms of sequence and structure, that protrude from the bacterial membrane to mediate bacterial invasion (invasin) or adhesion (intimin) to their host cell. Invasin is encoded by the chromosomal *inv* gene and this polypeptide mediates bacterial entry into M cells at the Peyer's patches of the gut (Isberg *et al.*, 1987; Marra & Isberg, 1997) by binding to β1-integrin receptors (Isberg & Tran Van Nhieu, 1994). This process triggers the internalization of the pathogen by a zipper-like mechanism (Isberg *et al.*, 2000). Intimin is encoded by the *eae* gene located in the locus of enterocyte effacement (LEE; McDaniel *et al.*, 1995), and is responsible for the intimate adhesion of EPEC, EHEC and *Citrobacter* bacteria to the enterocytes of the gastrointestinal tract where they produce a characteristic cytopathic effect known as an attaching and effacing (A/E) lesion (Moon *et al.*, 1983). This event is mediated, in part, by binding of intimin to its cognate receptor termed Tir (Tanslocated intimin receptor; Kenny *et al.*, 1997), which is a bacterial effector injected by the bacteria in the host cell plasma membrane via a type-III secretion system (Frankel & Phillips, 2008).

Intimins and invasins contain two functionally distinct regions, an amino (N) and a carboxy (C)-terminal region (Fig. 6a). The N-region comprises a signal peptide for Sec-dependent secretion across the IM, a hydrophilic periplasmic domain (Lys-M type in *Gammaproteobacteria*) that could tether the polypeptide to the peptidoglycan layer, and a large β-domain of about 500 amino acids that forms a β-barrel predicted to contain at least 12 amphipathic β-strands (Touze *et al.*, 2004; Bodelón *et al.*, 2009; Tsai *et al.*, 2010). The β-domain participates in the secretion of the C-region to the extracellular milieu and directs both the dimerization of intimin and Tir-clustering in the cell host membrane. The reported structures of the polypeptides comprising the extracellular C- region of 497 amino acids of invasin (Inv497) from *Yersinia pseudotuberculosis* (Hamburger *et al.*, 1999), the C-terminal 280 amino acids of intimin (Int280) from EPEC (Kelly *et al.*, 1999; Luo *et al.*, 2000) and 188 amino acids (Int188) from EHEC (Yi *et al.*, 2010), along with structural predictions of the rest of the C-region, show a rod-like structure consisting of four (invasin D1–D4) or three (intimin D0–D3) Ig-like domains (SCOP 49373) followed by a C-type lectin-like domain (invasin D5 and intimin D4) responsible for receptor-binding and located at the C-terminal tip of the molecule (Fig. 6b). The lectin-like domain and the most distal Ig-like domain of intimins and invasins form a rigid superdomain that participates in the binding to Tir (Tir-binding domain; Batchelor *et al.*, 2000; Yi *et al.*, 2010) or β1-Integrins (Leong *et al.*, 1990), respectively (Niemann *et al.*, 2004). The Tir–binding domain of EHEC and EPEC, despite its low sequence identity (48%), adopt a very similar structure, confirming previous data in which both intimins can cross-complement *in vitro* (Yi *et al.*, 2010). The cell-binding activity of intimins and invasins, depends on a conserved intradomain disulfide bond that is present in similar places in the lectin-like domain of these proteins (Leong *et al.*, 1993; Frankel *et al.*, 1995; Batchelor *et al.*, 2000). The overall topology of Inv497 D1-D3, EPEC Int280 D1-D2 and EHEC Int188 D2 Ig-like domains (there is no structural data available from intimin D0) resembles that of the IgSF type-I set. Inv497 D4 is similar to the IgSF C1 set (Hamburger *et al.*, 1999; Kelly *et al.*, 1999; Yi *et al.*, 2010). The Int/Inv Ig-like domains are classified into the Pfam database as Bacterial Ig-like domains 1 and 2 (Big_1 and Big_2) and invasin_D3: The D1–D4 domains of invasin are classified as Big_1 (D1), invasin_D3 (D3) and Big_2 (D2 and D4). The D0 to D2 domains of intimin are classified as Big_1 (D0 and D1) and Big_2 (D2; Fig. 6b).

Interestingly, a recent study has identified a new member of the intimin/invasin family in extraintestinal pathogenic *E. coli* strains causing urinary tract infections and sepsis, but which is also conserved in the genomes of pathogenic *E. coli* strains causing enteric diseases (Nesta *et al.*, 2012). This novel OM adhesin, called factor adherence *E. coli* (FdeC), is predicted to have an OM

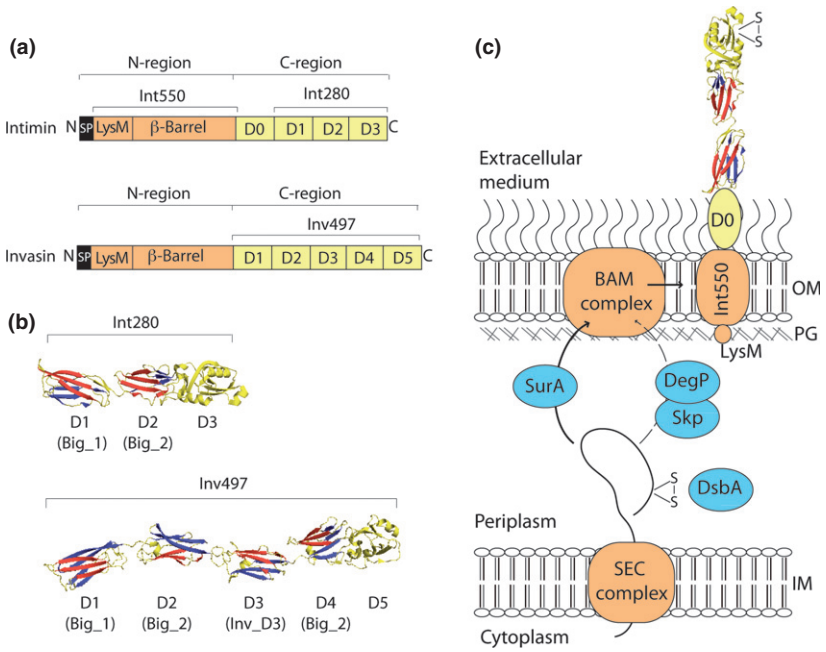


**Fig. 6.** Intimin and Invasin nonfimbrial adhesins. (a) Scheme of the N and C-regions of intimin and invasin showing the signal peptide (SP) for Sec-dependent secretion across the inner membrane (IM), the LysM domain, the β-domain and the secreted D domains. (b) Ribbon diagrams of the Int280 and Inv497 crystals showing the string of Ig-like domains and the C-terminal lectin-like domain. The PFAM terms for the Ig-like domains is indicated below each domain. (c) Intimin secretion to the extracellular medium starts with its translocation across the IM via the Sec translocon. In the periplasm DsbA catalyzes the formation of the disulfide bond (S-S) present in the D3 domain. The SurA pathway delivers the β-domain (Int550) to the BAM complex, where BamA is responsible for its folding and OM insertion. The alternate DegP/Skp pathway is depicted with a dashed line. The LysM domain is predicted to bind the peptidoglycan (PG) layer. Translocation of the secreted domains D0 (green oval), D1, D2 and D3 is believed to take place concomitantly with insertion of the β-domain in the OM via the BAM complex.

β-barrel and a large C-terminal region of about 900 amino acids containing nine surface-exposed Ig-like domains, but no evidence of a lectin-like domain. The 3D structure of a soluble FdeC fragment comprising three internal domains of the surface-exposed region confirmed their Ig-like structure (Nesta *et al.*, 2012). It has been also demonstrated that the exposed Ig-like domains of FdeC are able to bind epithelial cells and different collagens types *in vitro*. FdeC is expressed *in vivo* during colonization of the uroepithelium and, remarkably, immunization with the exposed Ig-like domains protects from infection (Nesta *et al.*, 2012).

A phylogenetic study of the Int/Inv family identified 69 sequence-divergent proteins present in *Alpha-*, *Beta-* and *Gammaproteobacteria* as well as *Chlamydia*, demonstrating that these adhesins comprising a β-barrel and Big motifs are not limited to enterobacteria, but they are widely distributed among other Gram negatives (Tsai *et al.*, 2010). This tandemly arrayed Ig-like domains, typical of many eukaryotic cell adhesion molecules or cell surface receptors, is also present in adhesins of Gram-positive pathogenic bacteria that belong to the microbial surface components recognizing adhesive matrix molecules (MSCRAMMs) such as the staphylococcal adhesin binding to fibrinogen (Ponnuraj *et al.*, 2003), or in Gram-positive pilin proteins like SpaA from *Corynebacterium diphteriae* (Kang *et al.*, 2009). Also, cell surface proteins of pathogenic Leptospiral species termed Leptospiral Ig-like (Lig) proteins present a structural organization of its extracellular domain very similar to that of intimins and invasins with tandem Ig-like domains identified as Big_2 in Pfam (Matsunaga *et al.*, 2003).

The mechanism of secretion of intimin and invasin (Int/Inv) family of proteins has not been elucidated. However, secretion of their members follows a multistep process described for most OMPs, which involves their translocation through the IM via the Sec-system, transport across the periplasm interacting with chaperones (e.g. SurA, DegP, Skp) and disulfide bond enzymes (e.g. DsbA), and insertion in the OM mediated by the β-barrel assembly machinery (BAM) complex (Fig. 6c; Ruiz *et al.*, 2006; Knowles *et al.*, 2009; Hagan *et al.*, 2011; Ricci & Silhavy, 2011; Dalbey & Kuhn, 2012). It has been demonstrated that intimin requires the BAM complex for OM insertion and that SurA is the major periplasmic chaperone involved in folding of intimin

β-barrel (Bodelón *et al.*, 2009). The protease activity of DegP participates in the degradation of unassembled intimin β-barrel in the periplasm. In addition, periplasmic DsbA catalyzes the formation of the disulfide bond of the lectin-like domain of intimin (Bodelón *et al.*, 2009), therefore indicating that this domain is at least partially folded in the periplasm prior to its translocation across the OM.

The mechanism of secretion of Int/Inv proteins may be similar to that used by autotransporters (ATs), a large superfamily of secreted virulence factors that contain an OM-anchored β-domain at the C-teminus, and a secreted 'passenger' domain at the N-terminus (Dautin & Bernstein, 2007; Leo *et al.*, 2012; Leyton *et al.*, 2012). The secreted domains of Int/Inv and ATs present different structures (a rod of Ig-like domains vs. a β-helix rod, respectively), and they are located in opposite sites of the polypeptides, but their translocation across the OM might follow similar mechanisms (Touze *et al.*, 2004; Adams *et al.*, 2005; Bodelón *et al.*, 2009; Leo *et al.*, 2012). Originally, it was proposed that these proteins could use a 'self-translocation' mechanism in which the β-barrel inserts in the OM forming a hydrophilic protein-conducting channel that is used for secretion of the passenger domain (Pohlner *et al.*, 1987; Henderson *et al.*, 1998). The crystallographic structures of ATs β-barrels show the existence of an α-helix that fills the hydrophilic lumen of the β-barrel connecting it to the 'passenger' domain (Oomen *et al.*, 2004; Barnard *et al.*, 2007; van den Berg, 2010). This α-helix is essential for the translocation process in ATs and, along with the β-barrel, constitutes a functional transport unit (Marín *et al.*, 2010). Interestingly, a similar region is predicted in the members of the Int/Inv family after the putative twelfth β-strand (Tsai *et al.*, 2010). However, data indicating that the Int/Inv and AT proteins can translocate Ig-like and other protein domains in, at least, a partially folded conformation, are difficult to explain in the context of the narrow hydrophilic channel provided by the β-barrels (Brandon & Goldberg, 2001; Veiga *et al.*, 2004; Skillman *et al.*, 2005; Purdy *et al.*, 2007; Bodelón *et al.*, 2009; Marín *et al.*, 2010). Therefore, an 'assisted-translocation' mechanism has been proposed for AT and Int/Inv proteins, in which insertion of the β-barrel and translocation of the surface-exposed domains are both assisted by the BAM complex (Bernstein, 2007; Bodelón *et al.*, 2009; Ieva *et al.*, 2011; Rossiter *et al.*, 2011; Leyton *et al.*, 2012). Interestingly, a recent study has discovered a novel conserved, but nonessential, OM protein member of the BamA (Omp85)-family, called TamA, which interacts with an integral IM protein (TamB) to assemble a transmembrane complex that promotes efficient secretion of ATs (Selkrig *et al.*, 2012). Whether this newly identified TAM-complex, or the essential BAM complex, forms the actual OM translocon for the Ig-like domains of Int/Inv proteins is unknown.

## Enzymes with Ig-like domains in *E. coli* and enterobacteria

Interestingly, the Ig-like domain is not only limited to structures that mediate bacterial adhesion, and thus, enterobacteria have adopted this type of folding in proteins bearing enzymatic functions. Ig-like domains are present in enzymes involved in an ample variety of cellular processes such as anti-oxidative damage, protein folding and, biosynthesis, degradation or transport of sugars (Table 1). The majority of these enzymes are complex structures formed by different functional domains, in which the Ig-like domain usually displays binding functions. For example, some of the Ig-like domains identified in sugar-binding enzymes are carbohydrate-binding modules (CBMs) that bring catalytic domains into close proximity with their cognate substrates (Boraston *et al.*, 2004; Guillen *et al.*, 2010). However, the striking example of the Copper, Zinc Superoxide Dismutase, an enzyme formed by a single Ig-like domain that has catalytic activity, confirms the versatility and evolutionary success of this topology.

## Superoxide dismutases and oxidoreductases

### Periplasmic Cu,Zn-superoxide dismutases (SODs)

The Cu,ZnSODs are metalloenzymes present in nearly all eukaryotic cells and in a large number of bacteria (Fang *et al.*, 1999; Battistoni, 2003) that catalyze the dismutation of superoxide to oxygen and hydrogen peroxide (McCord & Fridovich, 1969; Fridovich, 1995) protecting these organisms from oxygen-mediated free-radical damage (Bannister *et al.*, 1987). Both eukaryotic and prokaryotic Cu,ZnSODs share a conserved structure based on an Ig-like fold (Richardson *et al.*, 1976; Halaby *et al.*, 1999; Khare *et al.*, 2003; Culotta *et al.*, 2006; SCOP 49330). Whereas eukaryotic Cu,ZnSODs are homodimeric enzymes (Tainer *et al.*, 1982) characterized by high structural conservation (Bordo *et al.*, 1994; Perry *et al.*, 2010), the prokaryotic Cu,ZnSODs present larger structural variability than the eukaryotic counterparts (Pesce *et al.*, 2000) and, interestingly, some bacterial variants were isolated as active enzymes either as monomers [e.g. *E. coli* (Pesce *et al.*, 1997), *S. enterica* (Mori *et al.*, 2008), *Brucella abortus* (Chen *et al.*, 1995)] or as homodimers [e.g. *Salmonella typhimurium* (Pesce *et al.*, 2000), *Photobacterium leiognathi* (Bourne *et al.*, 1996)].

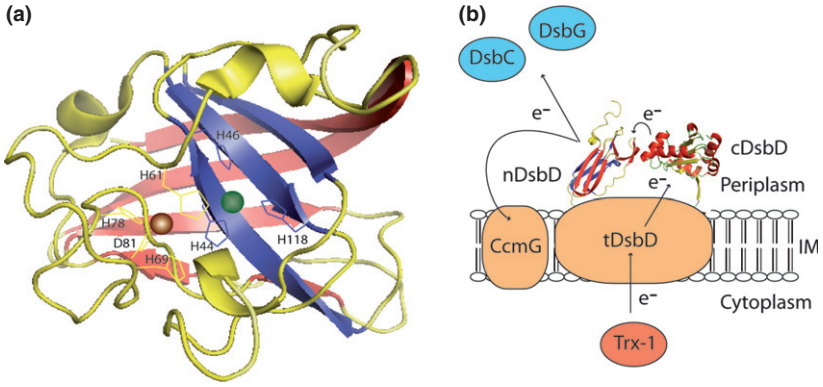


**Fig. 7.** Ig-like domains in *Escherichia coli* oxidoreductases. (a) Copper/zinc superoxide dismutase SodC. Ribbon diagram of the Ig-like periplasmic SOD SodC indicating the amino acid residues that contribute to the Copper (green sphere) and Zinc (brown sphere) binding sites. (b) DsbD thiol disulfide oxidoreductase. The thioredoxin (Trx-1) transports two electrons from cytoplasmic NADPH to the transmembrane domain of DsbD (tDsbD), and across the membrane via periplasmic cDsbD to the Ig-like domain nDsbD. This polypeptide functions as a hub and delivers the electrons to periplasmic substrates (e.g. DsbC, DsbG) or to IM proteins (e.g. CcmG). The black arrows indicate the electron flow.

The Cu,ZnSOD of *E. coli*, encoded by the *sodC* gene, is a periplasmic polypeptide of about 15 kDa that is specifically induced in stationary phase (Benov & Fridovich, 1994; Imlay & Imlay, 1996). The crystal structure of this enzyme (Pesce *et al.*, 1997) revealed a monomer composed by eight antiparallel β-strands with the two β-sheets of the sandwich each containing four strands (A, B, C and F and E, D, G and H) that are connected by seven loops (Fig. 7a). Two of them, termed the electrostatic and zinc loops (Tainer *et al.*, 1982; Bordo *et al.*, 1994), enclose the enzyme active center that is composed by one copper- and one zinc-binding sites linked via a histidine ligand (Pesce *et al.*, 1997). Whereas Copper is at the catalytic center of the enzyme and is cyclically oxidized and reduced during the successive encounters with the superoxide anion, it is believed that Zinc greatly contributes to the structural stability of the enzyme, modulation of the redox properties of the Copper ion, resistance to denaturing agents and proteolytic enzymes (Spagnolo *et al.*, 2004; Perry *et al.*, 2010). The Cu,ZnSODs of *E. coli* (SodC) and *S. enterica* (SodCII) have an intra-subunit disulfide bond between two cysteines located in the S-S subloop and in strand H (Cys 50 and Cys 144 in SodC) that is required for proper stability and formation of the Zinc-binding pocket of the active site of the enzyme (Pesce *et al.*, 1997; Mori *et al.*, 2008). The *in vivo* folding of the Cu,ZnSOD of *E. coli* depends on DsbA, which catalyzes the formation of its disulfide bond (Battistoni *et al.*, 1999).

As many Gram-negative bacteria contain a periplasmic Cu,ZnSOD (Gort *et al.*, 1999), it has been suggested that this enzyme plays an important role for scavenging superoxide species from the periplasm generated during aerobic growth, possibly attributable to electron leakage from the respiratory chain (Korshunov & Imlay, 2006). However, the fact that several pathogenic bacteria including *B. abortus*, *Legionella* sp., *E. coli*, *Salmonella* sp., or *Neisseria meningitidis* contain this enzyme, has led to the hypothesis that the Cu,ZnSOD could serve to protect these microorganisms against host defense-derived free-radical-mediated damage, thus facilitating bacterial survival within the infected host (Lynch & Kuramitsu, 2000; Battistoni, 2003; Perry *et al.*, 2010). Interestingly, in contrast to the nonpathogenic *E. coli* K-12 strain that has a single *sodC* gene, the pathogenic EHEC O157:H7 strain contains three *sodC* genes. One of them is homologous to that of the K-12 strain, and the other two, which encode almost identical proteins, are embedded within the sequences of two lambdoid prophages (CP-933R and CP-933V in the EDL933 strain; D'Orazio *et al.*, 2008). This redundancy also occurs in virulent *Salmonella* strains, which, besides the chromosomal *sodCII* gene, have a bacteriophage-encoded copy (*sodCI*) that contributes to *Salmonella* virulence (Fang *et al.*, 1999). These and other studies have established a role of the prophage associated SodC in virulence (Ammendola *et al.*, 2008; D'Orazio *et al.*, 2008).

### Inner membrane disulfide oxidoreductase DsbD

The *E. coli* DsbD is an inner membrane (IM) protein of 546 amino acids composed of a periplasmic N-terminal (nDsbD) domain (residues 1–143) with an Ig-like fold, a central transmembrane (tDsbD) domain (residues 144–418) and a C-terminal (cDsbD) periplasmic domain (residues 419–546; Chung *et al.*, 2000; Rozhkova *et al.*, 2004). Each of the domains contains two redox-active cysteines that are involved in the transfer of electrons

from cytoplasmic NADPH, via thioredoxin-1 (Trx-1; Rietsch *et al.*, 1997; Cho & Beckwith, 2009), to various periplasmic targets (Stewart *et al.*, 1999; Katzen & Beckwith, 2000). In the first step, tDsbD transfers electrons from Trx-1 to cDsbD, which in turn reduces the Cys103-Cys109 disulfide of nDsbD (Chung *et al.*, 2000; Katzen & Beckwith, 2000; Collet *et al.*, 2002), which enables a transient disulfide bond between Cys 109 of nDsbD and Cys 461 of cDsbD leading to the formation of a mixed disulfide complex termed nDsbD-SS-cDsbD (Rozhkova *et al.*, 2004). Next, nDsbD subsequently transfers the reducing potential to various substrates including periplasmic DsbC (Missiakas *et al.*, 1994) and DsbG (Bessette *et al.*, 1999; Depuydt *et al.*, 2009), and the membrane-anchored protein disulfide oxidoreductase DsbE/CcmG involved in c-type cytochrome maturation (Fabianek *et al.*, 1998).

The structures of the oxidized free form of *E. coli* nDsbD (Goulding *et al.*, 2002) and of its mixed disulfide complex with either cDsbD (Rozhkova *et al.*, 2004) or the periplasmic partners DsbC (Haebel *et al.*, 2002) and DsbE/CcmG (Stirnimann *et al.*, 2005) reveal that, unlike other Dsb proteins such as DsbA, DsbC, DsbE and cDsbD which all possess a thioredoxin-like fold, nDsbD features an Ig-like fold (SCOP 74863; Fig. 7b). This Ig-like domain of nDsbD consists of a β-sandwich formed by two β-sheets, each of which is constituted by three antiparallel strands (β1, β2, β8 and β4, β10, β12, respectively; Goulding *et al.*, 2002). However, in the nDsbD-DsbC complex the Ig-like domain is composed by a four-stranded β-sheet (β1, β2, β8 and β5) that packs against a three-stranded β-sheet (β4, β9 and β12; Haebel *et al.*, 2002), indicating that the topology of the domain varies depending on whether it is bound or not to its periplasmic substrates. The structural data also shows that the active site containing Cys103 (strand β11) and Cys109 (strand β10) is located at the N-terminus of the Ig-like domain. Interestingly, the cystein-active site of the oxidized form of nDsbD is protected from illegitimate redox reactions by a cap-loop region located between strands β6 and β7 of the Ig-like domain. In contrast, the active site in complexed-nDsbD is in an open conformation exposing the catalytic sulfur of Cys109. Therefore, opening of the cap-loop could be a prerequisite for the formation of an interdisulfide bond between Cys109 of nDsbD and the catalytic Cys of its partners (Stirnimann *et al.*, 2006; Quinternet *et al.*, 2009).

## Sugar-binding enzymes

### Cytoplasmic glycogen branching enzymes (GBEs)

The biosynthesis of glycogen in bacteria requires the participation of GBEs (EC 2.4.1.18) that cleave the linear α-1-4 linked glucose chain and the resulting oligosaccharide is subsequently linked in α-1-6 position to the carbohydrate chain. This branching activity increases the number of nonreducing ends, thus making glycogen more reactive to synthesis and digestion. Bacterial GBEs are classified into the glycoside hydrolase family 13 (GH13; Stam *et al.*, 2006) that belongs to the α-amylase superfamily (Jespersen *et al.*, 1991; Abad *et al.*, 2002). Structurally, bacterial GBEs are multidomain enzymes containing a central catalytic A-domain with a 'TIM' barrel folding that is invariably present in all the members of the α-amylase superfamily, and two C- and N-terminal extensions termed C and N domains, respectively, domains (Jespersen *et al.*, 1991; Palomo *et al.*, 2009). The C domain is believed to protect the hydrophobic residues of the catalytic domain from contacts with the solvent, and it been suggested to be involved in substrate binding. The N domain contains a CBMs (CBM48) of about 150 amino acids (Palomo *et al.*, 2009), which is classified into the SCOP database as having an Ig-like fold (SCOP 81282).

Some GBEs like the one from *E. coli* possess a long N-terminal region that contains two N domains of about 100 aminoacids termed N1 and N2, respectively (Lo Leggio *et al.*, 2002; Palomo *et al.*, 2009). Abad and collaborators elucidated the 3D structure of the *E. coli* GBE lacking its first 113 aminoacids (N113BE) and the crystal revealed that the N2 domain indeed adopts an Ig-like fold (Abad *et al.*, 2002). Although the N1 domain is absent in the reported crystal, its Ig-like folding has been predicted elsewhere (Lo Leggio *et al.*, 2002). The truncated N113BE shows around 60% of activity compared to the full-length enzyme, but its substrate preference and $K_m$ value were similar to the full-length protein. Amino terminal truncations of the *E. coli* GBE resulted in almost half reduction of enzyme activity (Hilden *et al.*, 2000), altered branching pattern (Binderup *et al.*, 2002) and to a gradual increase in the length of the glucan chains, indicating that this domain is involved in determining the size of the chain transferred (Abad *et al.*, 2002; Devillers *et al.*, 2003). It has been also reported that the putative CBM48 region of the N domain of two GBEs from two species of *Deinococcus*, dictates both the substrate and chain length specificity (Palomo *et al.*, 2009).

### Cytoplasmic glycosyl hydrolase family 2 (GHF2)

The *O*-glycosyl hydrolase superfamily (EC 3.2.1) is a widespread group of enzymes responsible for the hydrolysis and/or transglycosylation of glycosidic bonds. A classification system for glycosyl hydrolases (GH) which is available on the CAZy database, has led to the definition of 113 different protein families. The *E. coli*

β-galactosidase, encoded by the *lacZ* gene, is a member of this family of enzymes and its structure revealed two Ig-like domains (Juers *et al.*, 2000). The enzyme is a 464 kDa tetramer that is comprised of four polypeptide chains, named A–D, each of 1023 amino acids. Each monomer is made up of five domains, 1–5, being most of the enzyme active site residues within Domain 3 (residues 335–624), which has an α/β or 'TIM' barrel structure. The domains 2 (residues 219–334) and 4 (residues 625–725) of β-galactosidase have an identical Ig-like fold topology (SCOP 49305) composed by seven anti-parallel β-strands arranged in two β-sheets similar to that of FnIII (Jacobson *et al.*, 1994; Juers *et al.*, 2000). Importantly, the Ig-like Domain 2 contributes via a loop with amino acid residues to the completion of the active site within Domain 3 of the neighboring subunit (residues 272–288). Thus, monomer A donates its Domain 2 loop to complete the active site of monomer D and vice versa. The reciprocal situation also occurs between monomer B and C to give a total of four functional active sites (Juers *et al.*, 2000; Matthews, 2005). The Ig-like Domain 4 presents low homology among different β-galactosidases and it may just function as a linker between Domains 3 and 5. The Ig-like fold of Domains 2 and 4 might not only be limited to the β-galactosidase of *E. coli* and, based on structural comparisons, similar domains are found in other glycosidases (Juers *et al.*, 1999).

### Extracellular chitinase A

Chitin is the second most abundant polysaccharide in nature. It is a linear insoluble polymer of β(1–4)-linked *N*-acetylglucosamine (GlcNAc), present in fungal cell walls, shells of crustaceans, and exoskeletons of insects (Khoushab & Yamabhai, 2010). Genes encoding chitin-binding proteins or proteins containing chitin-binding domains have been found in many organisms including viruses, bacteria, fungi, insects, higher plants, and mammals. Based on their amino acid sequences, three-dimensional structures, and molecular mechanisms of catalytic reactions the chitinases can be grouped into GH families 18 and 19. Originally, it was indicated that bacterial chitinases belong to family 18 of glycoside hydrolases (Henrissat & Davies, 1997), although some members are now classified under family 19 of GH (Kawase *et al.*, 2004; Khoushab & Yamabhai, 2010). *Serratia marcescens* secretes to the extracellular medium four distinct types of chitinases (A, B and C1 and C2) and a chitin-binding protein (CBP21) lacking chitinase activity, which are encoded by the ChiA, B, C and Cbp21 genes, respectively (Horn *et al.*, 2006). ChiA and CPB21 contain Sec-dependent N-terminal signal peptides that are cleaved off by a periplasmic peptidase during their secretion to the extracellular medium. Nevertheless, its mechanism of OM secretion is still unknown.

The chitinase A (ChiA) from *S. marcescens* is a 58 kDa protein that comprises three domains: (1) an N-terminal module (amino acids 24–137) with an Ig-like fold (SCOP 49233) similar to a FnIII domain, termed ChiN domain and identified in the Pfam database as Chitinase A_N, (2) a catalytic domain with (α/β) 8-barrel fold structure, and (3) a C-terminal domain with (α + β) structure (Perrakis *et al.*, 1994). The ChiN domain is classified into the Cazy database as a CBM5. Although the overall topology and folding of ChiN and FnIII Ig-like domains is similar, they present a number of differences: ChiN strand A is shorter, there is long insertion between strands A and B, and strands F and G are linked by a cysteine bridge, which is absent in FnIII domains (Perrakis *et al.*, 1994). The catalytic mechanism of this type of chitinases involves substrate-assisted catalysis, and although the function of the ChiN domain is not well defined, the crystallographic data suggests that it may help the enzyme to remain bound to the chitin chain and direct the terminal sugar residues toward the catalytic groove. The surface of the ChiN domain has four conserved tryptophan residues important for substrate affinity that could be involved in interacting with the loose ends of chitin chains (Perrakis *et al.*, 1997). Two adjacently arranged tryptophans (Trp-33 and Trp-69), exposed on a continuous surface with the conserved aromatic residues of catalytic domain (Trp-245 and Phe-232), play important roles in guiding a chitin chain into the catalytic cleft to be hydrolyzed at the catalytic site (Uchiyama *et al.*, 2001).

### Periplasmic glucans biosynthesis protein G

*Escherichia coli* protein named OpgG is a 56 kDa periplasmic enzyme necessary for the production of glucans, which are linear-branched polysaccharides composed of 8–10 glucose units per molecule linked by β-1,2 linkages and branched by β-1,6 linkages, that constitute the bacterial envelope of many Gram-negative bacteria (Lacroix & Bohin, 2010). Although the precise function of OpgG remains unknown, under low nutrients and osmolarity growth conditions, *E. coli* and other Gram-negative bacteria synthesize glucans, which contribute to maintain the osmolarity of the periplasm preventing the swelling and rupturing of cytoplasmic membrane (Kennedy, 1982). It has been reported that OpgG is secreted via the Sec-system to the periplasm (Lequette *et al.*, 2004) where it may interact with the glycosyltransferase OpgH, the other member of the bicistronic OpgGH operon, to catalyze the formation of β-1,6 glucose branches (Debarbieux *et al.*, 1997; Hanoulle *et al.*, 2004). The structure of OpgG showed that it is composed of two β-sandwich domains

connected by a helix (Hanoulle *et al.*, 2004). Whereas the N-terminal domain (residues 22–388), that bears the putative catalytic activity, displays a 25-stranded β-sandwich fold, the C-terminal domain (residues 401–512) has a seven-stranded Ig-like fold (SCOP 110054) formed by two β-sheets constituted by A, B, D, E and C, F and G anti-parallel strands. Folding of the OpgG Ig-like domain resembles that of the *E. coli* α-1,4-glucan branching enzyme (GlgB), suggesting that, as GlgB, it may act to modulate the enzymatic activity of the OpgG catalytic N-terminal domain.

## Ig-like domains in ATP-binding cassette (ABC) membrane transporters

ABC transporters are ubiquitous across all cells and utilize the free energy of ATP hydrolysis to import or export a wide variety of substrates across biological membranes, ranging from small molecules such as ions, sugars or amino acids to larger compounds such as antibiotics, drugs, lipids and oligopeptides (Moussatova *et al.*, 2008; Cuthbertson *et al.*, 2010). Bacterial ABC transporters are involved in the uptake of nutrients (e.g. sugars and vitamins) and they also participate in the export of molecules (e. g. proteins, lipids, and oligo- and polysaccharides). Two examples of Ig-like domains have been reported ABC transporters in *E. coli*. One is present in the maltose transporter of *E. coli* and *Salmonella* (importer) and the other forms part of the LPS O-antigen transporter (exporter) of certain *E. coli* strains.

Despite their ample functional diversity ABC transporters share a general domain architecture: Two transmembrane domains (TMDs) that dimerize to form the substrate translocation pathway, coupled to two nucleotide-binding domains (NBDs), also known as ABC, that control the conformation of the TMDs through ATP-induced dimerization and hydrolysis-induced separation (Davidson *et al.*, 2008; Procko *et al.*, 2009). In addition to these four domains that form the core transporter, many ABC transporters have accessory domains for regulation or protein–protein interactions (Biemans-Oldehinkel *et al.*, 2006; Kos & Ford, 2009). This seems to be the case for the Ig-like domains present in the periplasmic loop region 2 of the MalF subunit (MalF-P2) of the maltose transporter and in the C-terminal region of the Wzt protein (cWzt) of the LPS O-antigen transporter.

### The periplasmic MalF-P2 Ig-like domain

The maltose transporter of *E. coli* and *Salmonella* is composed of the periplasmic soluble maltose-binding protein (MBP or MalE), encoded by *malE*, two integral IM proteins MalF and MalG, and two copies of the ATPase subunit MalK protein, which mediate the transport of maltose across the IM (Fig. 8a). The crystal structure of the reconstituted maltose transporter (MalFGK2) in complex with MalE (MalFGK2-E) in open conformation revealed that MalF contains a long periplasmic loop,

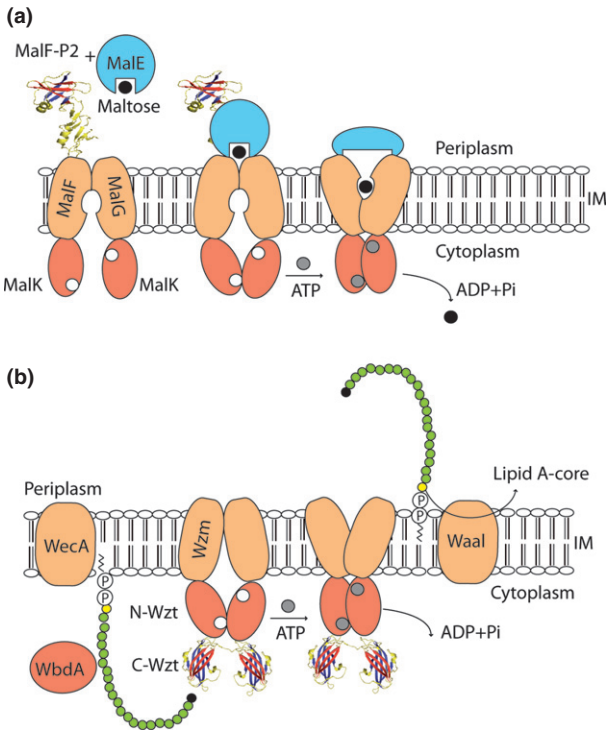


**Fig. 8.** Ig domains in ABC transporters. (a) The maltose importer MalFGK2. Maltose binds in the periplasm to MalE and stabilizes it in a closed conformation that interacts with the periplasmic Ig-like domain of MalF (MalF-P2). Binding of MalE to MalFGK2 (1) brings the NBDs of MalK closer such that ATP would promote a concerted motion of MalK closure (2), reorientation of the TM subunits, and opening of MalE (3). The change in conformation of MalE allows the transfer of maltose to the TM subunits binding site and positions ATP at the catalytic site for hydrolysis. Once ATP is hydrolyzed, the TMDs of MalF and MalG will reorient toward the cytoplasm, the transporter returns to the resting state and maltose completes its translocation across the membrane. (b) Export of *O*-polysaccharides (O-PS) by Wzt/Wzm. The synthesis of O-PS takes place at the cytoplasmic face of the IM starting with the WecA-mediated transfer of an *N*-acetylglucosamine-1-phosphate (yellow sphere) to form the undecaprenol pyrophosphate-linked intermediate. Next the mannosyltransferases WbdA, WbdB, and WbdC synthesize the O-PS by addition of mannose residues (green spheres) to the nonreducing end of the growing chain. WbdD finishes the extension of the nascent polymannan incorporating a methyl (O8 serotype) or a phospho-methyl (O9a serotype) residue (black sphere) to block the nonreducing terminus. The Ig-like domain of C-Wzt binds to the polymannose O-antigen substrate, possibly recognizing structural features in the nonreducing terminal modification. The export of the O-PS to the outer leaflet of the IM is initiated by ATP binding and hydrolysis in N-Wzt. Next, the O-antigen ligase WaaL, joins lipid A-core and O-antigen and the mature LPS is transported to the cell surface.

termed (MalF-P2), connecting its transmembrane helices 3 and 4, that adopts an Ig-like fold (residues N93-K275) and contacts periplasmic MalE in a cap-like manner (Oldham *et al.*, 2007). Although the MalF-P2 region loop is characteristic of enterobacterial MalF proteins (Tapia *et al.*, 1999), is almost unique among bacterial maltose ABC transporters and is missing in homologous systems of Archaea (Daus *et al.*, 2009). The MalF-P2 domain is able to interact with MalE independently of the transmembrane region of MalF (Jacso *et al.*, 2009). Binding of the maltose-loaded MalE to MalFGK2 brings together the MalK NBDs in the cytoplasm such that ATP would promote reorientation of the MalF and MalG subunits, opening MalE and the concomitant release of maltose toward the transmembrane binding site (Oldham & Chen, 2011).

### The cytoplasmic C-Wzt Ig-like domain

ABC transporters are also responsible for the export of an ample variety of glycans in cell surface glycoconjugates from bacteria. Examples include glycans from glycoproteins, teichoic acids, capsular polysaccharides, and the O-antigenic polysaccharide (O-PS) of lipopolysaccharide (LPS). LPS typically consists of three structural and functional regions: (1) the lipid A, (2) a nonrepeat core oligosaccharide, and (3) and a linear polymer termed O-polysaccharide (O-PS or O-antigen), which gives rise to about 180 O serotypes in *E. coli* (Raetz & Whitfield, 2002; Cuthbertson *et al.*, 2007). During LPS biogenesis, the lipid A-core oligosaccharide and O-PS are synthesized independently at the cytoplasmic side of the IM (Raetz & Whitfield, 2002) and the two pathways converge at a ligation reaction, which transfers the O-PS from undecaprenol-PP to lipid A-core oligosaccharides at the periplasmic face of the IM. Once assembled, LPS molecules are shuttled to the OM through a process involving the LptABCDE complex (Sperandeo *et al.*, 2009).

The synthesis of the O-PS follows two different pathways: The Wzm/Wzt ABC transporter-dependent pathway or the Wzy-dependent pathway (Raetz & Whitfield, 2002; Davidson *et al.*, 2008; Ruiz *et al.*, 2009). The Wzm/Wzt ABC transporter is comprised of two transmembrane (Wzm) and two nucleotide-binding polypeptides (Wzt) and translocates to the periplasm the O-PSs of *E. coli* serotypes O8, O9, and O9a (Fig. 8b; Clarke *et al.*, 2011). Interestingly, in contrast to other glycan-ABC transporters that export different types of polysaccharides (Davidson *et al.*, 2008), the *E. coli* Wzm/Wzt (ABC)-transporter is specific for its cognate substrate and this specificity relies in the C-terminal portion of Wzt (Cuthbertson *et al.*, 2005).

The Wzt polypeptides from the O8 and O9a serotypes are homologous proteins of 404 and 431 amino acids, respectively, that contain two functional regions: The transmembrane N-terminal region of Wzt (N-Wzt) is involved in ATP binding and hydrolysis and is highly conserved between the two proteins. The cytosolic C-terminal region (C-Wzt) is relatively less conserved and binds specifically the O-PS chain-terminating residue, thus dictating the serotype specificity for either the O8 or O9a (Fig. 8b; Cuthbertson *et al.*, 2005).

The structure of C-Wzt from serotype O9a reveals a dimer in which each monomer adopts an Ig-like fold that is composed by two anti-parallel β-sheets formed by the β-strands 2, 3 and 6, and 1, 4, 5, 7 and 8, respectively. The later β-sheet forms a concave groove that bears the O-PS pocket essential for binding. This carbohydrate-binding groove, conserved among O9a-like Wzt homologs, contains six critical aromatic residues, which allow substrate binding through a ring-stacking mechanism (Cuthbertson *et al.*, 2007). It has been proposed that this binding domain introduces the polymer into the transport channel. In an alternative model, it may be required to disengage the nascent O-PS from the assembly enzymes to allow it to enter the export pathway (Cuthbertson *et al.*, 2010).

## Other Ig-like containing proteins

### The extracellular sugar-binding Cbp21 protein

The Chitin-binding protein 21 (CBP21) secreted by *S. marcescens* (Suzuki *et al.*, 1998), is a 21 kDa noncatalytic protein, classified as a CBM family 33 CBM in the CaZy database. Although the precise biological function of CBP21 is not clear, it has been suggested that it may play a role to enhance efficient chitin degradation. CBP21 binds to the insoluble crystalline substrate, leading to structural changes in the substrate and increased accessibility promoting hydrolysis (Vaaje-Kolstad *et al.*, 2005a).

The structure of Cbp21 consists in an Ig-like fold (SCOP 117045) of two β-sheets (β-strands A, B, and E and β-strands C, D, F, and G, respectively) with a 65-residue 'bud' of three short helices located between β-strands A and B. The protein contains four cysteine residues forming two disulfide bridges, one in the loop/helical region (C41 and C49) and one joining β-strands D and E (C145 and C162; Vaaje-Kolstad *et al.*, 2005a, b). The structural data showed that Cbp21 and the ChiN domain of ChiA from *S. marcescens* (Perrakis *et al.*, 1994) present a similar FnIII-like fold, although the ChiN domain lacks the Cbp21 bud-like extension. In contrast to other CBMs, and ChiN domain in particular, CBP21 does not have a cluster of aromatic amino acids on the surface that is used for carbohydrate binding. Remark-

ably, it contains a conserved surface patch of hydrophilic residues that was shown to be essential for the disruption of the polysaccharide, as variants with single mutations on the largely polar binding surface lost their ability to promote chitin degradation while retaining considerable affinity for the polymer. Thus, the interaction of CBP21 with chitin could be primarily governed by polar interactions that may disrupt the hydrogen-binding network between individual polysaccharide chains (Vaaje-Kolstad *et al.*, 2005a, b).

### The periplasmic copper resistance protein (CopC)

Copper is as a catalytic and structural cofactor for enzymes that is essential for most prokaryotic and eukaryotic organisms because it is involved in numerous biological processes (e.g. respiration, iron transport, oxidative stress protection; Puig & Thiele, 2002; Turski & Thiele, 2009). However, this metal ion can become toxic when is present free in the cell because it participates in redox reactions that result in the transfer of electrons to hydrogen peroxide with the concomitant generation of damaging hydroxyl radicals. Therefore, the levels of intracellular Copper are tightly regulated by homeostatic systems, which involve Copper acquisition, sequestration, and efflux (Banci *et al.*, 2009). Certain bacterial strains contain extra-chromosomal operons that confer Copper resistance and allow survival of the cell under extremely high Copper levels (Zhang *et al.*, 2006; Djoko *et al.*, 2008). The best-characterized Copper resistance loci have been isolated from Gram-negative bacteria colonizing areas contaminated by the use of Copper salts (Munson *et al.*, 2000). In *E. coli* and *Pseudomonas syringae* resistance to high levels of Copper is conferred by plasmid-borne clusters termed *pcoABCDRSE* (Brown *et al.*, 1995) and *copABCDRS* (Cha & Cooksey, 1991), respectively.

The Copper resistance protein PcoC from *E. coli* (and its homologous CopC in *Pseudomonas syringae*) are periplasmic soluble monomeric proteins of about 100 amino acids (Djoko *et al.*, 2007), constituted by seven anti-parallel β-strands having an Ig-like fold (SCOP 81969). Both PcoC and CopC are proposed to function as Copper carriers in the periplasm exchanging copper with the multicopper oxidase CopA, that converts CuI to the less toxic CuII (Djoko *et al.*, 2008), and with the membrane-bound copper pumps CopB and CopD ((Koay *et al.*, 2005). PcoC and CopC have two distinct (CuI) and (CuII) binding-regions (Arnesano *et al.*, 2003; Wernimont *et al.*, 2003). Crystallographic and spectroscopy studies of the *E. coli* PcoC have shown that the CuI site is constituted by a methionine rich loop located at the C-region of the protein, and the CuII site is formed both by β-strand-interconnecting loops containing histidine residues and by the N-terminal amino group located at the N-region of PcoC (Fig. 9; Drew *et al.*, 2008). Interestingly, the CuII site in PcoC is classified as type-II2, like that of Cu,Zn-SOD, nevertheless Cu-PcoC does not exhibit superoxide dismutase activity *in vitro* (Huffman *et al.*, 2002).

## Chaperones involved in folding of Ig-like proteins in *E. coli*

### General periplasmic chaperones

With a few exceptions (e.g. LacZ, GlgB, c-Wzt), most Ig-like domains of *E. coli* proteins are located in the periplasm or exposed to the periplasm at some point during their secretion and/or biogenesis (Table 1). These periplasmic-exposed *E. coli* proteins with Ig-like domains are transported across the IM by SecYEG translocon (Driessen & Nouwen, 2008), reaching the periplasm in an unfolded conformation. Folding of Ig-like proteins in the periplasm is assisted by different folding factors. There are three different types of periplasmic folding catalysts that in some cases have overlapping functions: chaperones, which bind to proteins and help to prevent undesirable off-pathway interactions or aggregation of their substrates (e.g. Skp, SurA, DegP), peptidyl–prolyl *cis/trans* isomerases (PPIases) that catalyze *cis/trans* isomerization of proline peptide bonds in proteins (e.g. SurA, FkpA, PpiA, PpiD), and protein disulfide forming enzymes and isomerases that catalyze the formation and exchange of disulfide bonds (e.g. DsbA, DsbC; Allen *et al.*, 2009). These folding factors assist folding of mutiple proteins in the periplasm and also have a role in folding of OMPs, including intimin and fimbrial ushers, during their periplasmic transit to the OM (Justice *et al.*, 2005; Bos *et al.*, 2007; Bodelón *et al.*, 2009; Knowles *et al.*, 2009; Palomino *et al.*, 2011). The dedicated periplasmic chaperones

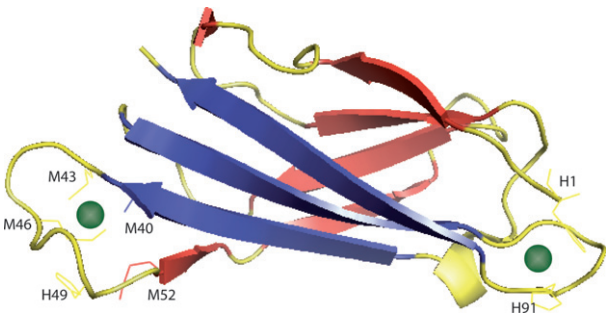


**Fig. 9.** Copper resistance protein PcoC. Ribbon diagram of the Ig-like periplasmic protein PcoC indicating the Copper (green spheres) CuI (N-region) and CuII (C-region) binding sites and the amino acid residues involved.

(e.g. PapD, FimC) of fimbrial subunits were discussed in a different section.

Some of the above mentioned periplasmic chaperones, which play an essential role for the homeostasis of the cell, are also key players in the cellular response under periplasmic and membrane stress conditions. The accumulation of aggregated protein in the periplasm leads to cell damage, and therefore, bacteria have evolved stress response pathways to sense and combat this effect. In *E. coli* the σE and Cpx pathways, which are induced upon extracytoplasmic stress, upregulate periplasmic chaperones such as SurA, Skp or DegP that bind unfolded or misfolded proteins to alleviate folding defects (SurA and Skp) or to proteolyze misfolded polypeptides (DegP; Duguay & Silhavy, 2004; Mogensen & Otzen, 2005).

The SurA polypeptide presents four domains: an N-terminal domain, two central parvulin-like domains (PPIase 1 and 2) observed in other PPIases and a short C-terminal domain (Bitto & McKay, 2002). The N- and the C-terminal regions bear the chaperone activity, whereas the PPIase activity resides in the PPIase domain 2 (Behrens *et al.*, 2001). Xu and collaborators reported the crystal structure of SurA in complex with short peptides known to specifically interact with SurA and suggested that the PPIase domain 1 is responsible for substrate selection (Xu *et al.*, 2007). SurA binds preferentially to Ar-X-Ar tripeptide motifs common in OMPs, being Ar and X any aromatic and polar residues, respectively (Bitto & McKay, 2003), and it was shown that SurA binds OMP assembly intermediates immediately after they leave the Sec translocon before signal sequence cleavage (Ureta *et al.*, 2007).

Skp is a highly basic protein of about 17 kDa that forms a homotrimer in solution, which binds its substrate in a 1 : 1 stoichiometry. The crystal structure of the homotrimer resembles a jellyfish with tentacles defining a substrate-binding cavity that is likely to hold substrates of different sizes protecting them from aggregation (Korndorfer *et al.*, 2004; Walton & Sousa, 2004; Walton *et al.*, 2009). Functionally, the arms of the Skp oligomer have a net positive charge on their external side and hydrophobic patches on their internal side that are thought to bind to LPS and hydrophobic substrates, respectively (Allen *et al.*, 2009). This periplasmic chaperone was identified in a screen for periplasmic folding factors because it was retained on an affinity column with Sepharose-bound OmpF (Chen & Henning, 1996). Interestingly, Skp was also identified as a folding factor for heterologous antibody Ig domains expressed in *E. coli* (Bothmann & Plückthun, 1998). Skp it is known to bind OMPs to prevent their aggregation and to escort them from the IM to the OM (Schafer *et al.*, 1999; Harms *et al.*, 2001). Skp was co-purified in complex with 31 envelope proteins being 19 of them β-barrel proteins, indicating that Skp exhibits a broad substrate spectrum (Jarchow *et al.*, 2008).

DegP belongs to the high temperature requirement (HtrA) family of serine proteases and is able to perform the antagonistic functions of protein repair and degradation. It rescues and refolds aggregated or slightly misfolded proteins or proteolyzes irretrievably misfolded proteins with the aim of reducing extracytoplasmic stress and cellular damage. DegP mediates the encapsulation and/or refolding of misfolded proteins via its chaperone-like activity. Recent biochemical and structural analysis have shed light into the molecular architecture and function of DegP (Allen *et al.*, 2009; Ortega *et al.*, 2009; Subrini & Betton, 2009; Sawa *et al.*, 2010). Each DegP monomer is composed of three domains: a N-terminal trypsin-like protease domain and two C-terminal PDZ domains that contain the structural features allowing DegP to recognize and tether its substrates for proteolytic cleavage. The first crystal structure reported for DegP consists in a hexamer formed by two homotrimers stabilized by interactions between the PDZ domains (Krojer *et al.*, 2002). Subsequently, two independent studies reported that, in the presence of the substrate, the homotrimers further oligomerize via PDZ1–PDZ2 interactions into cage-like spherical structures of 12- and 24-mers that exhibit both protease and chaperone activity. Inside of the cage, the substrate protein is then subjected to either protease degradation or refolding (Jiang *et al.*, 2008; Krojer *et al.*, 2008).

FkpA is a member of the FKBP family of periplasmic PPIases with general chaperone activity (Horne & Young, 1995; Arie *et al.*, 2001). The PPIase activity catalyzes the *cis/trans* isomerization of peptide bonds before proline (Xaa-Pro or prolyl bonds). Although this is a sequence-specific reaction, the ample specificity of the chaperone domain for its substrate allows these enzymes to act over a broad range of protein substrates. Also, catalysis of isomerization is several orders of magnitude faster than chaperone-mediated substrate delivery (Jakob *et al.*, 2009). The crystal structure of dimeric FkpA shows a V-shaped molecule comprised of two monomers, each containing a globular C-terminal domain and a N-terminal dimerization domain linked together by a long α-helix, that functions as a flexible arm. The cleft formed between the two arms of the homodimer might accommodate the substrates (Saul *et al.*, 2004; Hu *et al.*, 2006). The chaperone activity of FkpA is independent of its PPIase activity. A structural and functional study of FkpA indicated that the PPIase and the chaperone activities reside in the N- and C-terminal domains, respectively, and act independently of one another (Saul *et al.*,

2004; Hu *et al.*, 2006). There is limited data available regarding the natural substrates of this folding factor. FkpA does not seem to have a specific affinity for OMPs, and to date it still remains unclear whether the protein is involved in OMP biogenesis (Allen *et al.*, 2009). An *in vivo* target of FkpA is the toxin colicin M, which is a polypeptide produced by *E. coli* that kills sensitive *E. coli* cells. Colicin M appears to be refolded by FkpA upon its entry into the periplasm of sensitive *E. coli* cells (Hullmann *et al.*, 2008). Interestingly, it was also demonstrated that FkpA improves the functional expression of antibody Ig domains in *E. coli* (Bothmann & Plückthun, 2000; Ramm & Plückthun, 2000).

## Disulfide bond catalysts

The presence of disulfide bonds is not a conserved feature in all *E. coli* and enterobacterial Ig-like domains (Table 1). Nevertheless, many of them contain structural disulfide bonds that may be essential for the stability and functionality of these polypeptides. The available crystal structures of Ig-like domains from *E. coli* proteins reveal disulfide bonds in structural components of secreted adhesive organelles such as fimbria, pili, fibrills or capsule-like structures, some periplasmic chaperones and ushers of the CU pathway, the Cu,ZnSOD and DsbD enzymes, Chitin-hydrolase ChiA and Chitin-binding protein Cbp21. Most disulfide bonds of *E. coli* Ig-like domains are present in the structural components of fimbriae such as the pilin domain of major (e.g. PapA and FimA) and minor fimbrial subunits (e.g. PapE, PapH, PapK, FimF, FimG, and FimH) or the pilin subunits of polyadhesive fimbria (e.g. AfaE-III). Sequence analysis of fimbrial proteins assembled by classical CU systems showed that they contain two conserved cysteine residues that may form a noncanonical intradomain disulfide bond between the beginning of β-strand A1 and the end of β-strand B (Fig. 3a; Piatek *et al.*, 2010). This cysteine bond possesses a unique localization (joining two adjacent β-strands) not found in any other protein family belonging to the IgSF.

Most organisms have specialized machinery to catalyze formation and isomerization of disulfide bonds (Heras *et al.*, 2007), and the best-characterized disulfide bond (DSB) system is that of *E. coli* K-12 (Heras *et al.*, 2009). A bioinformatic analysis of the cysteine content of predicted cell envelope proteins from 375 bacterial genomes showed that, although the bacterial species posses an ample diversity in the mechanisms for disulfide bond formation, *Gamma-* and *Betaproteobacteria* have DSB systems similar to that of *E. coli* K-12 (Dutton *et al.*, 2008; Heras *et al.*, 2009). In *E. coli* formation of disulfide bonds takes place in the periplasm and is mainly catalyzed by the DsbA oxidoreductase (Bardwell *et al.*, 1991), which is a member of the thioredoxin (Trx) superfamily (Kadokura *et al.*, 2003; Sevier & Kaiser, 2006) that contains a Trx domain with a redox-active cysteine pair (Cys30–Cys33; Martin *et al.*, 1993). DsbA introduces disulfide bonds nonspecifically by donating its disulfide. It has been suggested that DsbA may have more than 300 *in vivo* substrates of the 700 periplasmic and membrane proteins listed in the Swiss Protein Database having two or more cysteine residues (Hiniker & Bardwell, 2004). The DsbB IM protein (Bardwell *et al.*, 1993), which is able to generate disulfides de novo, restores the oxidizing activity of DsbA by reoxidizing its cysteine pair.

In addition to the DsbA–DsbB oxidation pathway the periplasmic space contains a DsbC-DsbD disulfide-isomerizing pathway, in which DsbC functions as the major chaperone/protein disulfide isomerase repairing incorrectly paired cysteines (Rietsch *et al.*, 1996). a DsbC must be maintained in its reduced state for proper isomerase activity and the bacterial periplasm lacks a source of reducing equivalents, *E. coli* has evolved a mechanism for importing electrons from the cytoplasm that are used for reduction of DsbC in a process mediated by the IM thiol disulfide reductase DsbD (Missiakas *et al.*, 1995). The *E. coli* DSB system also includes DsbG, a homolog of DsbC that acts to protect single cysteines from sulfenylation (Depuydt *et al.*, 2009), and DsbE, which is restricted to the pathway of cytochrome c maturation (Fabianek *et al.*, 1999). In-depth reviews regarding the *E. coli* and other DSB systems have been published elsewhere (Ito & Inaba, 2008; Heras *et al.*, 2009; Inaba, 2009).

It is also becoming apparent the importance of the DSBs for bacterial pathogenesis and numerous Gram-negative pathogens either encode additional copies of their DSB genes or posses functional DSB paralogs (e.g. DsbL, DsbI) that, despite of some levels of redundancy, showed differences in substrate specificity (Bouwman *et al.*, 2003; Totsika *et al.*, 2009). The implications of DSBs in the assembly of functional fimbria was demonstrated in bacterial pathogens assembling different types of adhesive organelles such as P fimbriae of UPEC (Jacob-Dubuisson *et al.*, 1994), bundle-forming pili of EPEC (Zhang & Donnenberg, 1996), plasmid-encoded fimbriae (Pef) of *S. enterica* (Bouwman *et al.*, 2003), the type-IV pilus of *N. meningitidis* (Tinsley *et al.*, 2004). As an example, UPEC bacteria lacking DsbA are defective in the assembly of P fimbriae owing to the inability of the PapD chaperone to fold into a conformation able to bind subunits *in vivo* (Jacob-Dubuisson *et al.*, 1994). Zav'yalov and collaborators reported that capsule production was significantly decreased in an *E. coli dsbA* mutant carrying the *Y. pestis* F1 operon. As Caf1 subunit contains no cysteine residues, the most likely explanation for the

observed decrease in subunit assembly is owing to an effect on the Caf1M chaperone. Two cysteine residues of Caf1M chaperone form a disulfide bond and reduction of this bond increases the dissociation constant for the Caf1M-Caf1 complex (Zavialov *et al.*, 2007). Interestingly, the *dsbA* mutant strain accumulates considerably lower amount of Caf1M than the wild type, suggesting that a misfolded intermediate would be more susceptible to degradation by periplasmic proteases. Totsika and collaborators also showed that DsbL catalyzes disulfide bond formation of PapD as overexpression of DsbL in the absence of DsbA restored the production of functional P fimbriae in *E. coli* CFT073 *dsbAB dsbLI* mutant (Totsika *et al.*, 2009). Finally, DsbA has been shown to catalyze disulfide bond formation in the C-terminal lectin-like domain of the afimbrial adhesin intimin (Bodelón *et al.*, 2009). Intimin expressed in an *E. coli dsbA* mutant lacks this disulfide bond, shows lower display levels on the cell surface and has higher sensitivity to protease digestion (Bodelón *et al.*, 2009).

## Biotechnological applications of Ig proteins in *E. coli*

Natural *E. coli* proteins with Ig-like domains, such as Type 1 fimbrial subunits and intimin, have been exploited biotechnologically for the display of peptides and small protein domains on the surface of *E. coli* cells (Pallesen *et al.*, 1995; Klemm & Schembri, 2000; Wentzel *et al.*, 2001). Nonetheless, the major biotechnological breakthrough related to the expression of Ig proteins in *E. coli* is the development of technologies for the expression and selection of full-length antibodies and small antibody fragments in this bacterium.

Well before the discovery of a natural *E. coli* protein containing Ig-like domains (Holmgren & Brändén, 1989), biotechnologists had already focused their attention in the possibility of producing antibodies in *E. coli*. This objective was spurred by the success of an emerging recombinant DNA technology for the production of small human polypeptide hormones of biomedical interest in *E. coli*, such as insulin (Goeddel *et al.*, 1979), and by the therapeutic expectations raised by the technology of monoclonal antibodies (mAbs) from murine hybridomas (Kohler & Milstein, 1975). However, expression of IgGs in *E. coli* was a formidable task owing to their secreted nature and large mass (*c.* 150 kDa), with heterotetrameric structure composed of two identical light (L) chains (*c.* 25 kDa) and two identical heavy (H) chains (*c.* 50 kDa), the presence of multiple intra- and interchain disulfide bonds and post-translational modifications (i.e. glycosylation; Elgert, 1998; Maynard & Georgiou, 2000; Schroeder & Cavacini, 2010). The difficulties found for producing full-length H and L chains in stoichiometric amounts and good yields to assemble functional IgGs in *E. coli* led scientist to focus on the expression of small antibody fragments with antigen-binding activity. Currently, it is also possible to produce and select in *E. coli* full-length IgGs and other antibody formats. These developments will be described in the following sections.

### Structure of classical antibodies and recombinant Fab and scFv fragments

The classical IgGs of humans, mice, and most mammal species, are Y-shaped bivalent molecules with identical antigen-binding sites at the tip of the two arms (Fig. 10a). Each arm is referred to as the antigen-binding fragment (Fab), which is composed by the association of the VH and CH1 domains of the H chain with the VL and CL domains of the L chain. The actual antigen-bind-

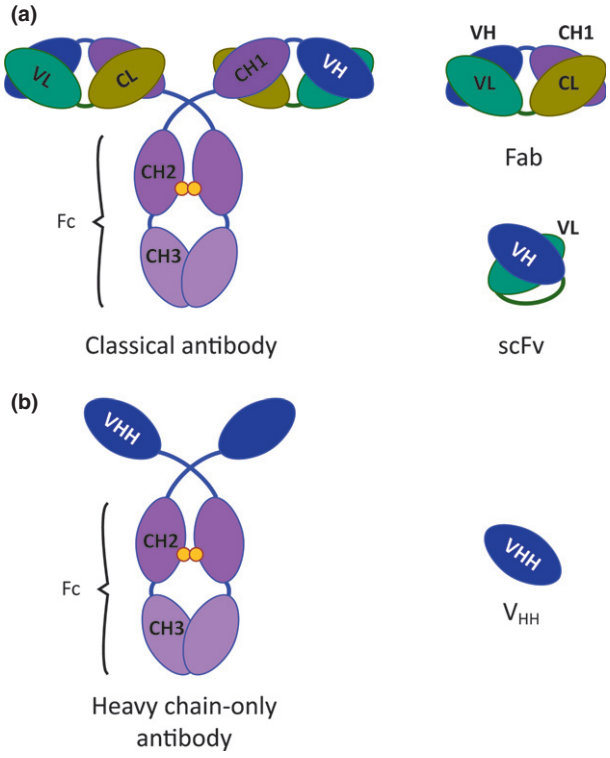


**Fig. 10.** Schematic diagram of antibodies and small antigen-binding fragments. (a) Structure of classical IgGs with heavy (H) and light (L) chains indicating the crystallizable fragment (Fc) region with conserved glycosylation site (orange circle) of CH2 domain. The small antigen-binding fragments Fab and scFv derived from classical IgGs are shown on the right. (b) Structure of camelid HCAbs lacking the light (L) chains and CH1 domains. The glycosylated Fc region of HCAbs is similar to that of classical IgGs. The single-domain antibody (sdAb) fragment derived from the VH of camelid HCAbs, referred to as VHH (VH of HCAb) or nanobody, is shown on the right.

ing site is formed of three hypervariable loops found in the VH and VL domains, also known as complementarity-determining regions (CDRs). The sequence and structural diversity of these six CDRs, juxtaposed in the folded antibody structure forming flat or concave antigen-binding surfaces, provide the capacity to recognize multiple antigenic structures (Padlan, 1996; Schroeder & Cavacini, 2010). The stalk of the Y, or crystallizable fragment (Fc), is assembled by the constant CH2-CH3 domains of the two H chains in the IgG molecule, and contains a conserved N-linked glycan at residue N297 in the CH2 domain. The Fc region mediates the effector functions of antibodies by recruiting the complement, in complement dependent cytotoxicity (CDC), and immune cells, in antibody dependent cell-mediated cytotoxicity (ADCC) and phagocytosis (ADCP; Raju, 2008; Schroeder & Cavacini, 2010; Desjarlais & Lazar, 2011). To exert these effector functions the glycosylated Fc region interacts with blood serum proteins (e.g. C1q) and specific Fcγ receptors (FcγRs) found in immune cells (e.g. Natural Killer, Macrophages; Carroll, 2008; Nimmerjahn & Ravetch, 2008). In addition, the Fc region plays an important role for the long-circulating half-life of full-length IgGs in the body owing to its recognition by the neonatal Fc receptor (FcRn) expressed in endothelial cells of the vasculature, epithelial cells of intestine, kidneys and lungs, among other organs (Roopenian & Akilesh, 2007). The FcRn is involved in the recycling of endocyted IgG molecules by antigen presenting cells (APCs) and endothelial cells, and in the transcytosis of IgGs through several epithelial and endothelial layers, including kindney podocytes (Roopenian & Akilesh, 2007).

The first recombinant antibodies (rAbs) expressed in *E. coli* were Fab fragments (Better *et al.*, 1988) and single-chain variable fragments (scFv), which are assembled by linking the VH and VL domains into a single polypeptide chain by means of a short flexible peptide (Fig. 10a; Bird *et al.*, 1988; Huston *et al.*, 1988). These formats have been the more common antibody fragments expressed in *E. coli* because of their simple structure and the ability to select specific binders from large combinatorial libraries of Fab and scFvs genes (Marks *et al.*, 1991; Plückthun *et al.*, 1996; Hoogenboom, 2005). Libraries of antibody fragments can be amplified from B-cells of diverse origin (e.g. mouse, human) after immunization or disease ('immune libraries') or from nonimmunized organisms ('naïve libraries'). Alternatively, the sequence of V genes can be mutagenized *in vitro* to diversify the CDRs generating the so-called 'synthetic libraries' (Griffiths *et al.*, 1994; Nissim *et al.*, 1994; Aujame *et al.*, 1997; Benhar, 2007; Zhai *et al.*, 2011). These repertoires of antibody fragments can be displayed on the capsid of filamenteous *E. coli* bacteriophages (e.g. M13) for selection of high-affinity antigen-binding clones, which allows the isolation and affinity maturation of antibodies for therapeutic and *in vivo* diagnostic applications (Clackson *et al.*, 1991; Hoogenboom *et al.*, 1998; Hoogenboom, 2002; Kretzschmar & von Ruden, 2002; Bradbury *et al.*, 2011). As discussed later in the text, full-length IgGs can also be expressed in *E. coli* cells and display on filamentous bacteriophages for selection of high-affinity clones (Mazor *et al.*, 2007, 2008, 2009, 2010).

## Heavy-chain antibodies and single-domain antibodies (sdAbs)

Besides Fab and scFvs other recombinant antibody fragments are currently expressed in *E. coli* and selected by phage display from antibody libraries. Among them, the most remarkable are the single-domain antibodies (sdAbs; Holliger & Hudson, 2005; Saerens *et al.*, 2008; Wesolowski *et al.*, 2009). These small fragments (*c.* 12–14 kDa) are based on a single V domain with full antigen-binding capacity isolated from a special class of natural homodimeric antibodies devoid of L chains called heavy-chain antibodies (HCAbs; Fig. 10b) found in camelids (e.g. dromedaries, llamas) and cartilaginous fish (e.g. sharks), where they are called IgNARs (Ig of shark new antigen receptor). The sdAb from a camelid HCAb is referred to as VHH (VH of HCAb; Hamers-Casterman *et al.*, 1993; Muyldermans *et al.*, 1994; Arbabi Ghahroudi *et al.*, 1997; van der Linden *et al.*, 2000), whereas that from an IgNAR is called VNAR (Diaz *et al.*, 2002; Dooley *et al.*, 2003). VHHs and VNARs are able to bind with high affinity and specificity their cognate antigens in the absence of VL domains. Although their overall Ig domain folding is similar to that of VH domains from classical antibodies (with H and L chains), the lack of a VL partner has driven the evolution of special features for antigen recognition and domain stability and solubility (Desmyter *et al.*, 1996; Vu *et al.*, 1997; Decanniere *et al.*, 1999; van der Linden *et al.*, 1999; Muyldermans *et al.*, 2001; Stanfield *et al.*, 2004; Flajnik *et al.*, 2011). The CDR3 loops of VHHs and VNARs, and to a lesser extent the CDR1 loops, are longer than those of classical VHs providing novel conformations and an extra interaction surface for antigen recognition. The long CDR3 loop provides the major antigen-interaction surface in VHHs and VNARs and often protrudes from the Ig domain as a finger-like structure. This long CDR allows the recognition of epitopes in narrow clefts and concave cavities in protein structures (e.g. active sites of enzymes), which are frequently not accessible to conventional antibodies with larger size and flat or concave antigen-binding surfaces (Fig. 11a; Desmyter *et al.*, 1996). Potent enzyme inhibitors based on VHHs binding to active sites of enzymes have been

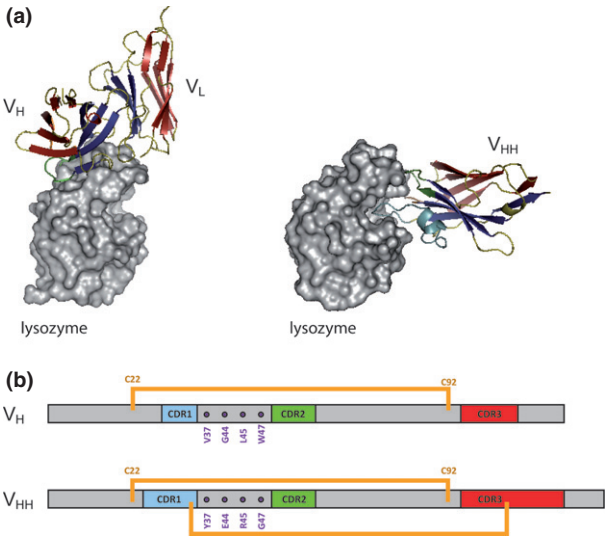


**Fig. 11.** Comparison between conventional VH/VL and VHH antibody fragments. (a) Three-dimensional structure of chicken lysozyme recognized by an scFv (VH/VL domains) (left; PDB code 1mlc) or by a VHH domain of camelid HCAb (right; PDB code 1mel). The CDRs of VH, VL and VHHs are shown in different colors (CDR1 in orange, CDR2 in green, CDR3 in cyan). The long CDR3 loop of the VHH protrudes as a finger-like structure toward an epitope located in the cavity of the active site of lysozyme (right), whereas the flat VH/VL paratope binds to an epitope located at the exposed surface of lysozyme (left). Images generated with PyMOL program (Molecular Graphics Systems, LLC). (b) Cartoon showing the polypeptide chains of conventional VH and camelid VHHs, indicating the amino acids of framework region 2 of classical VHs (V37, G44, L45, W47) that are substituted by more hydrophilic amino acids (e.g. Y37, E44, R45, G47) in VHHs. Other distinct features shown are the different length of CDR1 and CDR3 in VHs and VHHs, the presence of the conserved canonical disulfide bond between residues C22 and C92, and the extra-disulfide bond between Cys residues of CDR1 and CDR3 in VHHs.

isolated (Lauwereys *et al.*, 1998; Transue *et al.*, 1998; Conrath *et al.*, 2001; Desmyter *et al.*, 2002). Interestingly, small VHHs and VNARs can also reach to the less-accessible inner regions of proteins found on the cell surface of pathogens, and not only to the outermost epitopes of these proteins, which are frequently variable among strains or during infection to evade the immune system (Stijlemans *et al.*, 2004; Henderson *et al.*, 2007).

In addition, the protein solubility and stability of VHHs are higher than isolated VH domains from classical antibodies (van der Linden *et al.*, 1999; Dumoulin *et al.*, 2002). This is owing to a combination of subtle amino acid changes along their sequence (Fig. 11b), including substitutions of conserved hydrophobic residues in classical VHs (V37, G44, L45, and F47 or W47) for more hydrophilic amino acids (Y37/F37; E44/Q44; R45/C45; G47/R47/L47/S47; Muyldermans *et al.*, 2001, 2009). In some cases, the long CDR3 loops of VHHs participate in the higher solubility of these domains by protecting exposed hydrophobic residues of the core-Ig domain from aqueous environment. Furthermore, VHHs and VNARs often contains, besides the canonical disulfide bond of Ig domains, an extra-disulfide bond connecting CDR3 and CDR1 (in camels and sharks) or CDR3 with CDR2 (in llamas) that assist in stabilizing the conformation of these CDRs and the overall stability of the domain (Govaert *et al.*, 2011). Interestingly, the protease resistance and thermal stability of VHHs can be enhanced by the introduction of an additional disulfide bond in the core-Ig domain (A/G54C and I78C; Hussack *et al.*, 2011).

The distinct properties of VHHs and VNARs explain their high level of expression in *E. coli*, increased solubility, monomeric behavior, resistance to degradation, and to heat and chemical denaturation. Given the unique biophysical and antigen-binding properties of sdAbs, these antibody fragments have become attractive molecules for diverse biotechnological applications (Harmsen & De Haard, 2007; De Marco, 2011). Remarkably, camelid VHHs also show high sequence identity with the human VH family 3 (the most commonly expressed human VH family) opening the possibility of their therapeutic application (Holliger & Hudson, 2005; Saerens *et al.*, 2008; Vincke *et al.*, 2009). This potential and their small nanometer size (*c.* 2.5 nm width and *c.* 4 nm height) prompted a 'marketing' rename of VHHs as Nanobodies (Nbs), which is now a widely accepted term for these sdAbs also in scientific literature.

There has been considerable interest in developing sdAbs based on human VH (or VL) sequences that could mimic the binding and stability properties of natural VHHs. Isolated heavy and light chains of antibodies are overproduced as Bence-Jones proteins in multiple myeloma and other human pathologies (Hilschmann & Craig, 1965; Seligmann *et al.*, 1979; Hendershot *et al.*, 1987; Prelli & Frangione, 1992). In fact, murine VH sdAbs against lysozyme were isolated from phage display in *E. coli* before the discovery of natural HCAbs in camels (Ward *et al.*, 1989). However, most VH sequences from conventional antibodies tend to aggregate in solution and show low affinity for antigens. Introduction of 'camel' mutations G44E, L45R, and W47G, in a human VH3 sequence significantly increased its solubility (Davies & Riechmann, 1994). Human VH clones with low tendency to aggregate were isolated by repeated cycles of heating and cooling of phages with displayed VH sequences (Jespers *et al.*, 2004). Based on these findings, synthetic libraries of human VHs have been developed by randomizing CDRs in 'camelized' and/or selected human VH sequences and used for isolation of functional sdAbs with nanomolar affinity (Davies & Riechmann, 1994, 1995;

Martin *et al.*, 1997; Reiter *et al.*, 1999; Dumoulin *et al.*, 2002; Holt *et al.*, 2008; Arbabi-Ghahroudi *et al.*, 2009). SdAbs based on human VL domains have also been reported and expressed in *E. coli*. These include VL domains from conventional IgGs and naive scFv libraries (Colby *et al.*, 2004; Martsev *et al.*, 2004; Cossins *et al.*, 2007; Schiefner *et al.*, 2011) and sdAbs isolated from synthetic libraries (van den Beucken *et al.*, 2001; Holt *et al.*, 2008).

## Periplasmic expression of antibody fragments and their selection by phage display

Ab fragments (Fab, scFv and sdAbs) are commonly exported into the periplasm of *E. coli* cells (Fig. 12a) where disulfide bond forming and isomerization enzymes (i.e. DsbA, DsbC) and other protein chaperones (e.g. Skp, FkpA, SurA) enable their correct folding (Skerra & Pluckthun, 1991; Glockshuber *et al.*, 1992; Missiakas & Raina, 1997; Jurado *et al.*, 2002; Merdanovic *et al.*, 2011; Schlapschy & Skerra, 2011). Standard expression vectors are derived from phagemids with a replication origin of high-copy number pBR/pUC-plasmids and a phage origin of replication and packaging from M13/fd filamentous phages (Qi *et al.*, 2012). These vectors carry an inducible *E. coli* promoter (e.g. P*tac*, P*lac*, $P_{BAD}$, P*phoA*) and an N-terminal secretion signal (e.g. PelBss, OmpAss, PhoAss, STIIss) to which the gene segment encoding the Ab fragment is fused in frame (Choi & Lee, 2004). In addition, these vectors commonly incorporate epitope tags (e.g. HA, myc, 6xhis) in the C-terminus for detection and purification of the overproduced Ab fragment. Vectors for expression of scFvs and sdAbs (e.g. pHEN) have a single monocistronic transcriptional unit. Fab expression vectors export two polypeptide chains (the L and VH-CH1 chains) from either a single bicistronic transcriptional unit (e.g. pComb3) or from two independent monocistronic transcriptional units with separate promoters (e.g. pComb3X; Qi *et al.*, 2012).

Prolonged high-level expression of the Ab fragment at 37 °C frequently leads to protein misfolding and aggregation and induces permeabilization of the bacterial

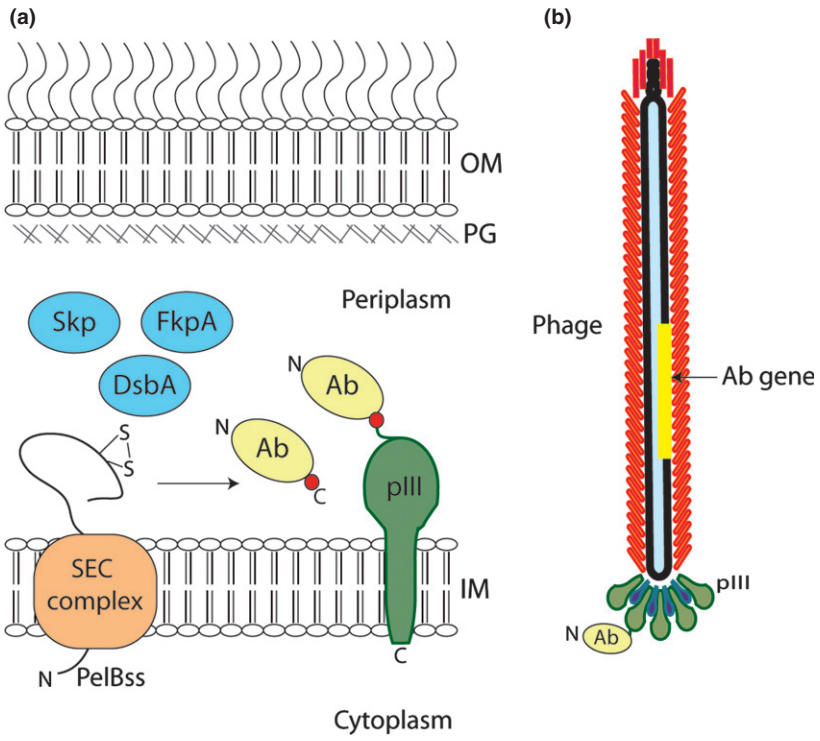


**Fig. 12.** Expression of antibody fragments in the periplasm of *Escherichia coli* cells. (a) Soluble Ab fragments, depicted as a single-domain Ab in this scheme, are exported to the periplasm with an N-terminal signal peptide (e.g. PelBss) fused to epitope tags (e.g. myc, 6xhis) at the C-terminus for immune detection and protein purification (red circle). The Ab fragment can also be fused at its C-terminus to the minor coat protein III (pIII) of filamentous phages for phage display. In this case, the Ab fragment is exposed to the periplasm but tethers to the inner membrane (IM) of *E. coli* by pIII. (b) Phage antibody particles assembled upon infection of *E. coli* cells expressing Ab-pIII fusions with a helper phage. The Ab-pIII fusion is displayed on the phage capsid (usually one copy per phage), which contains the phagemid DNA with the Ab gene.

OM with leakage of periplasmic proteins, releasing a significant fraction of the overproduced Ab fragment soluble in the culture media (Skerra, 1993). Misfolding and aggregation of the Ab fragment, along with toxicity and lysis of the overproducing *E. coli* cells, can be minimized growing the cells at lower temperatures (16–30 °C; Somerville *et al.*, 1994). The coexpression of protein chaperones (Skp, FkpA, SurA) and disulfide bond enzymes (DsbA, DsbC, DsbG, human PDI) in the periplasmassists the correct folding of Ab fragments and improve their final yield in *E. coli* (Humphreys *et al.*, 1996; Plückthun *et al.*, 1996; Bothmann & Plückthun, 1998, 2000; Hayhurst & Harris, 1999; Zhang *et al.*, 2002; Friedrich *et al.*, 2010). In addition, *E. coli* strains deficient in various periplasmic proteases (DegP, Prc, Spr) increase the accumulation of certain Ab fragments (Chen *et al.*, 2004).

Phagemid vectors for Ab expression also enable the fusion of the Ab fragment (the VH-CH1 chain in the case of Fabs, the VH-VL chain of scFvs or the single V domain of sdAbs) with the minor coat protein III (pIII) of filamentous phages (e.g. M13, fd) for phage display (Fig. 12a). Frequently, between the Ab gene (with an N-terminal secretion signal and C-terminal epitope tags) and the gene III segment (encoding pIII lacking its own secretion signal) is placed an amber stop codon (TAG) that is suppressed to Glutamine in *E. coli* strains carrying the *supE* mutation. As a consequence, these vectors encode the soluble periplasmic Ab fragment in 'wild-type' *E. coli* strains (e.g. WK6, HB2151) and the hybrid protein between the Ab fragment and pIII in *E. coli supE* strains (e.g. TG1, XL-1Blue). The Ab moiety of the hybrid protein is exposed to the periplasm but is tethered to the IM of *E. coli* by the pIII. As mentioned above, the N-terminal signal peptides commonly used for Ab fragments are post-translational Sec-signals, such as PelBss, OmpAss, and PhoAss. Interestingly, highly stable protein domains need the use of co-translational SRP-signals (e.g. DsbAss) for efficient display on filamentous phages (Steiner *et al.*, 2006).

Bacteriophages displaying the Ab fragments, also called Phabs (Fig. 12b), are produced by infecting *E. coli* F+ *supE* cells (e.g. TG1, XL1-Blue) carrying the phagemid vector, with a helper phage encoding the bacteriophage proteins (e.g. M13-K07, VCS-M13). These helper phages have a defective origin of replication and packaging that reduces incorporation of its own genome into phage particles. Thus, released phage particles contain phagemid DNA instead of helper phage DNA, linking the genotype (antibody gene) and phenotype (displayed antibody fragment) in an infective particle. Ab gene libraries packaged as Phabs are incubated with a chosen antigen and/or cell, and clones with the desired antigen-binding specificity and affinity can be recovered and amplified by infecting fresh *E. coli* F+ *supE* cells, in a process called biopanning (Clackson *et al.*, 1991; Marks *et al.*, 1991; Winter *et al.*, 1994; Chames *et al.*, 2002; Mutuberria *et al.*, 2004; Hoogenboom, 2005). Assembled Phabs usually contain a single copy of the hybrid Ab-pIII fusion and up to 3–5 copies of the wild-type pIII encoded by the helper phage. Monovalency of Phabs results in the selection of high-affinity Ab binders (O'Connell *et al.*, 2002). In some situations, such us panning against rare antigens on cell surfaces, phage multivalency may be desirable. Multivalent Phabs can be produced by infection of *E. coli* with mutant helper phages having a deletion in gene III (Rakonjac *et al.*, 1997; Rondot *et al.*, 2001; Baek *et al.*, 2002; Soltes *et al.*, 2003; Oh *et al.*, 2007).

## Expression and selection of full-length IgGs in the periplasm of *E. coli*

Access to libraries of human V sequences and their engineering is a major application of the phage display of Ab fragments in *E. coli* (Lonberg, 2008). However, given their small size and lack of the Fc region, Ab fragments miss effector functions and show a significantly reduced *in vivo* half-live (Roopenian & Akilesh, 2007; Carroll, 2008; Nimmerjahn & Ravetch, 2008). Hence, Ab fragments have been traditionally formatted as full-length IgGs for therapeutic applications and produced in bioreactors of mammalian cell culture systems (e.g. CHO cells) being released into the culture media as fully assembled and glycosylated IgG molecules that have followed the mammalian secretory route (Chadd & Chamow, 2001; Brekke & Sandlie, 2003, Birch & Racher, 2006; Beck *et al.*, 2010).

Given the structural complexity of IgGs, early works that attempted their expression in *E. coli* only reported the production of insoluble aggregates that need to be refolded *in vitro* (Boss *et al.*, 1984; Cabilly *et al.*, 1984). Although refolding of IgGs produced in *E. coli* as inclusion bodies has been greatly optimized (Hakim & Benhar, 2009), it is not widely used for IgG expression. Production of fully assembled, functional, full-length IgGs in the periplasm of *E. coli* was not reported until 2002 (Simmons *et al.*, 2002). The geness encoding L and H chains of a human IgG1 molecule were cloned in a plasmid vector with two independent transcriptional units controlled by independent *phoA* promoters. To balance the expression of L and H chains, *E. coli* translation initiation regions (TIR) of various strengths were placed upstream of the fusions between the N-terminal signal peptide of STII and the L and H chains. Optimal TIRs for the L and H chains were found empirically by analyzing the amount of assembled IgG in the periplasm of *E. coli* cells. These *E. coli*-produced aglycosylated IgGs were purified and

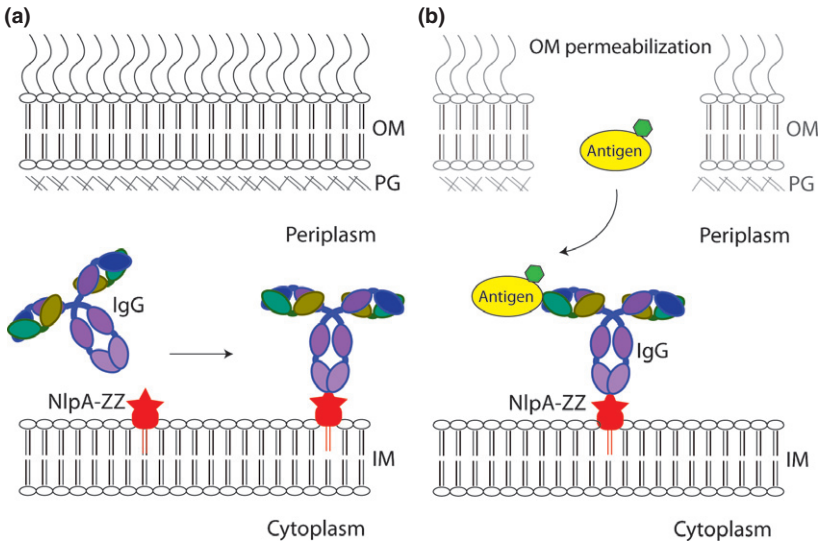


**Fig. 13.** Expression of full-length IgGs in the periplasm of *Escherichia coli*. (a) Scheme showing the periplasm of *E. coli* cells containing a full-length IgG tethered to the bacterial inner membrane (IM) with chimeric lipoprotein NlpA-ZZ (shown in red). (b) IgGs expressed in the periplasm can be selected for specific binders by permeabilization of the OM and incubation of the spheroplasts with a labeled antigen for fluorescence-activated cell sorting.

compared to the glycosylated IgGs from mammalian CHO cells, demonstrating identical antigen-binding activity. In addition, they exhibited equal affinity for human FcRn receptor, whose binding does not depend on the glycosylation of the Fc region, having similar half-lives in chimpanzees. However, aglycosylated *E. coli* IgGs do not bind human C1q and FcYRs (Simmons *et al.*, 2002; Reilly & Yansura, 2010) and, therefore, are not able to elicit Ab effector functions such as CDC, ADCC and ADCP. Although the absence of Ab effector functions may be desirable in some cases, such as those in which binding or neutralization of the antigen (e.g. hormone/toxin/virus) is the only therapeutic effect required for the Ab, aglycosylation could represent a major limitation for the IgGs produced in *E. coli*. Despite this, aglycosylated IgG antibodies obtained through protein engineering have entered clinical trials with excellent therapeutic results, tolerance, and pharmacokinetics (Jung *et al.*, 2011). In addition, variants of therapeutic IgGs with mutations in the Fc region can, at least partially, bypass the need of glycosylation and selectively activate certain FcγRs (Sazinsky *et al.*, 2008; Jung *et al.*, 2011).

The therapeutic potential of aglycosylated IgGs and simplified bioprocessing triggered an interest in producing and selecting full-length human IgGs directly in *E. coli*. Research conducted by Dr. George Georgiou and collaborators (University of Texas) has achieved this major goal. This group reported an expression vector with a single P*lac* promoter and a bicistronic operon for L and H chains, both fused to the PelBss, which efficiently assembled several IgGs in *E. coli* (Mazor *et al.*, 2007). In addition, based on the same vector, they developed a library screening system in which IgGs against any given antigen can be selected and directly expressed in *E. coli* (Mazor *et al.*, 2007). Pools of VH and VL gene segments were cloned in this vector, replacing the VH and VL domains of a well-expressed human IgG, and selected for specific antigen binders using fluorescent sorting of spheroplasted *E. coli* cells, with a modification of the anchored-periplasmic expression (APEx) technology (Chen *et al.*, 2001; Harvey *et al.*, 2004). To this end, the library of IgGs is co-expressed with a chimeric lipoprotein in the periplasm, named NlpA-ZZ, which tethers IgGs to the bacterial IM Fig. 13a). NlpA-ZZ comprises the N-terminal signal peptide and the first six amino acids of the *E. coli* lipoprotein NlpA fused to a synthetic analog of the IgG-binding B domain of protein A from *Staphylococcus aureus*, called ZZ (Jendeberg *et al.*, 1995). Spheroplasts obtained by OM permeabilization of *E. coli* cells are incubated with the antigen of interest, labeled with a fluorophore (Fig. 13b), and sorted in a cytometer. After a few cycles of selection, full-length IgGs of high-affinity for the antigen are obtained and expressed in *E. coli* at high levels (Mazor *et al.*, 2007, 2008; Makino *et al.*, 2011).

Screening of large IgG libraries ($> 10^8$ clones) is complex using only fluorescent cell sorting. Selection of binders can be optimized by combining phage panning with subsequent rounds of fluorescent cell sorting of

*E. coli* spheroplast. To this end, full-length IgGs are displayed on phage particles having the ZZ domain fused to pIII (Mazor *et al.*, 2010). Collectively, these developments enable the fast and versatile isolation and production of functional aglycosylated full-length IgGs in *E. coli*, or E-clonal antibodies, against an antigen of interest from combinatorial libraries of VH and VL genes. Finally, these technologies have also been used for selection of aglycosylated Fc variants that selectively bind certain FcγRs (Jung *et al.*, 2010).

## Systems for multimerization, surface display, extracellular secretion and intracellular expression of antibody fragments in *E. coli*

Albeit the expression of Ab fragments and full-length IgGs in the periplasm of *E. coli* is the most common strategy, other Ab formats and expression systems have been developed, which have potential for specialized applications.

Protein engineering provided dimerization and multimerization of Fab, scFv, and sdAbs to produce bi-, tri- and multi-valency for increased antigen avidity (apparent affinity; Cuesta *et al.*, 2010). This has been accomplished by joining monovalent fragments by means of short peptide linkers (e.g. Diabodies; Perisic *et al.*, 1994; Conrath *et al.*, 2001) and stable dimerization and oligomerization domains, such as CH3 domains and Fc regions derived from natural Igs (e.g. Minibodies and Fc fusions; Hu *et al.*, 1996; Li *et al.*, 2000), Leucine zippers (e.g. ZIP-miniantibodies; Pack & Pluckthun, 1992), dimeric Alkalyne Phosphatase from *E. coli* (Furuta *et al.*, 1998), Barnase-Barstar complex (Deyev *et al.*, 2003), trimeric coiled coils (Fan *et al.*, 2008; Cuesta *et al.*, 2009), tetrameric Streptavidin (Dubel *et al.*, 1995; Cloutier *et al.*, 2000), and pentameric B-subunits of toxins (Zhang *et al.*, 2004), among others. In addition these multimerization strategies enable the generation of bi- and multi-specific Ab molecules for the simultaneous recognition of several antigens. For instance, diabodies that simultaneously bind the target antigen and recruit TNFα (Schmidt & Wels, 1996), serum Igs (Holliger *et al.*, 1997), the complement (Kontermann *et al.*, 1997), T cells (Helfrich *et al.*, 1998; Kipriyanov *et al.*, 1998), or NK cells (Arndt *et al.*, 1999). Bi-specific diabodies based on two sdAbs have also been constructed for increasing circulating half-life *in vivo* through binding to serum albumin (Harmsen *et al.*, 2005; Coppieters *et al.*, 2006; Roovers *et al.*, 2007; Tijink *et al.*, 2008).

Alternative expression strategies have been developed for the production of Abs in *E. coli*. For instance, functional scFvs have been displayed on the surface of *E. coli* cells anchored to the OM with the chimeric Lpp-OmpA protein (Francisco *et al.*, 1993; Daugherty *et al.*, 1998) and the autotransporter (AT) β-domain of the *Neisseria gonorrhoeae* IgA protease (Veiga *et al.*, 1999), but the frequent aggregation associated to scFv fragments limits their translocation across the OM (Wörn & Plückthun, 1999; Veiga *et al.*, 2004). Contrary to scFvs, VHHs have been efficiently displayed on the surface of *E. coli* cells with the AT β-domain of IgA protease from *N. gonorrhoeae* (Veiga *et al.*, 2004), and the β-domain of EhaA from enterohemorrhagic *E. coli* O157:H7 (Marín *et al.*, 2010).

Functional scFv and VHHs can be secreted to the extracellular media of *E. coli* cultures, at concentrations *c.* 0.5–5 mg $L^{-1}$ in shake flasks, with the *E. coli* α-hemolysin (HlyA) transport system (Fernández *et al.*, 2000; Fraile *et al.*, 2004). In this expression system, a fusion between the Ab fragment and the C-terminal secretion signal of HlyA (Koronakis *et al.*, 1989) is transported from the cytoplasm to the extracellular media, without a periplasmic step, through the HlyB/HlyD/TolC complex that connects the bacterial IM and OM (Koronakis *et al.*, 2004). The Ab fragments secreted with the HlyA system do not accumulate in the periplasm of *E. coli* and elicit very low toxicity for the overproducing bacterium. This situation contrasts to the release of Ab fragments to the culture media owing to OM leakage and cell lysis after their Sec-dependent periplasmic overexpression in *E. coli*. The HlyA transport system could be particularly suited for continuous Ab fragment production in bioreactors and the delivery of Ab fragments by commensal *E. coli* strains *in vivo* (Fernández *et al.*, 2000; Rao *et al.*, 2005).

Another secretion strategy of interest for Ab fragments in *E. coli* is the use of the type-III secretion system (T3SS) form EPEC strains. T3SSs of EPEC and other Gram-negative pathogens form supramolecular protein complexes, or injectisomes, expanding the bacterial cell envelope and protruding from the cell surface with a needle-like structure and a tip translocon complex able to form a pore in the plasma membrane of eukaryotic cells (Knutton *et al.*, 1998; Garmendia *et al.*, 2005; Cornelis, 2006). Injectisomes allow bacteria to translocate a number of proteins (i.e. effectors) to the cytoplasm of the eukaryotic host cells, which subvert different cell functions to benefit infection (Galan & Wolf-Watz, 2006; Galan, 2009). Functional sdAbs (VHHs) have been translocated to the cytoplasm of human cells using the T3SS of EPEC (Blanco-Toribio *et al.*, 2010). The VHHs were fused to the 20 amino acid N-terminal T3S signal from a natural EPEC effector (EspF). The fusions were injected by wild-type and attenuated EPEC strains into the cytoplasm of HeLa cells cultured *in vitro* at levels *c.* $10^5$–$10^6$ molecules per cell (Blanco-Toribio *et al.*, 2010). *Escherichia coli*

bacteria remain extracellular during translocation of sdAbs to HeLa cells. Additional improvements are needed in the infecting *E. coli* strain for selective delivery of therapeutic sdAbs in the cytoplasm of human cells (intrabodies). The high-stability properties of VHHs make them excellent candidates to function as intrabodies to abrogate or modulate the activity of intracellular proteins (Kirchhofer *et al.*, 2010; Peréz-Martínez *et al.*, 2010; Vercruysse *et al.*, 2010).

Functional Ab fragments have also been expressed in the cytoplasm of *E. coli* cells. As most Ig V domains contain a conserved disulfide bond that is required for the stability of the folded polypeptide (Wörn & Plückthun, 2001), their expression in the reducing conditions found in the cytoplasm of *E. coli* mostly produce misfolded polypeptides unable to bind the antigen. Lack of disulfide bonds in the cytoplasm of *E. coli* is owed to the thioredoxin and glutaredoxin pathways that keep free thiol groups in a reduced state (Ritz & Beckwith, 2001; Kadokura *et al.*, 2003). Production of functional Fab and scFv fragments in the cytoplasm of *E. coli* was achieved in double mutant strains in *trxB* (thioredoxin reductase) and *gor* (glutathione oxidoreductase) that simultaneously coexpress in the cytoplasm the chaperone Skp or the disulfide bond isomerase and chaperone DsbC, both devoid of the N-terminal signal peptide (ΔssSkp and ΔssDsbC, respectively; Levy *et al.*, 2001; Jurado *et al.*, 2002). In *E. coli trxB gor* cells coexpressing ΔssSkp or ΔssDsbC, disulfide bonds are formed correctly in cytoplasmic Fab and scFv fragments. It has also been demonstrated that N-terminal fusions between thioredoxin (Trx1) and scFv fragments allow their correct folding in the cytoplasm of *E. coli trxB gor* cells (Jurado *et al.*, 2006b). In this case, coexpression ΔssDsbC is not needed because Trx1 acts itself as a chaperone assisting the proper folding of the scFv (Jurado *et al.*, 2006b). Using these strategies, immune libraries of scFvs have been screened *in vivo* for selection of intrabodies blocking a prokaryotic transcriptional activator (Jurado *et al.*, 2006a) and a relaxase involved in plasmid conjugation in *E. coli* (Garcillan-Barcia *et al.*, 2007). Intrabodies selected *in vivo* with *E. coli trxB gor* cells may be stabilized later for correct folding in the reducing environment of the cytoplasm of wild-type cells. Interestingly, APEx has been employed for selection of stable scFvs able to fold in reducing environments taking advantage of their expression in the periplasm of *E. coli dsbA* mutants (Seo *et al.*, 2009), in which oxidation of disulfide bonds in periplasmic Ab fragments is abolished (Jurado *et al.*, 2002). Using this methodology highly stable scFvs that fold in the periplasm of *E. coli dsbA* mutants were selected and shown to be active when expressed in the cytoplasm of wild-type *E. coli* cells.

## Concluding remarks

The Ig-like domain is frequently found in *E. coli* and enterobacterial extracellular and cell surface proteins that function in adhesion and cell host invasion, especially structural and components of multiple fimbrial systems and members of the intimin and invasin family of OM proteins. Periplasmic fimbrial chaperones and OM ushers involved in the assembly of fimbriae also contain Ig-like domains that are essential for their biological function. But the presence of the Ig-like domain is not limited to adhesive organelles and their assembly machineries. Ig-like domains are found in enzymes with oxidoreductase and hydrolytic activities, ABC transporters, sugar-binding and metal-binding proteins. Most of these domains are exposed to the periplasm where general periplasmic chaperones, PPIases and DSB forming and isomerization enzymes participate in their folding. In the light of the 'skills learnt' by *E. coli* during evolution for expression, secretion, and folding and their endogenous Ig-like-containing proteins, it is not surprising the biotechnological succes obtained in the heterologous expression of full-length antibodies and small antibody fragments in the periplasm of this bacterium and its filamentous bacteriophages. A better understanding of the structure and the mechanisms of secretion, folding and assembly of the natural Ig-like-containing proteins of pathogenic *E. coli* strains will not only provide a rationale basis for the design of new anti-microbial compounds to combat infection but will help to improve current expression systems of heterologous Igs in nonpathogenic laboratory *E. coli* strains and to expand their potential biotechnological applications.

## Note added in proof

While this manuscript was at the proof stage the crystal structures of the β-barrels of Intimin and Invasin have been reported (Fairman *et al.*, 2012). These structures reveal β-barrels of 12 antiparallel β-strands with a peptide linker that runs from the periplasm toward the extracellular space through the internal barrel pore.

## Acknowledgements

This work has been supported by Grants of the Spanish Ministry of Science and Innovation (BIO2008-05201; BIO2011-26689), the Autonomous Community of Madrid (S-BIO-236-2006; S2010-BMD-2312), CSIC (PIE 2011 20E049), '*la Caixa*' Foundation, and the VI Framework Program from the European Union (FP6-LSHB-CT-2005-512061 NoE 'EuroPathogenomics'). The authors declare that there is no conflict of interest.

## References


Abad MC, Binderup K, Rios-Steiner J, Arni RK, Preiss J & Geiger JH (2002) The X-ray crystallographic structure of *Escherichia coli* branching enzyme. *J Biol Chem* **277**: 42164–42170.

Adams TM, Wentzel A & Kolmar H (2005) Intimin-mediated export of passenger proteins requires maintenance of a translocation-competent conformation. *J Bacteriol* **187**: 522–533.

Allen WJ, Phan G & Waksman G (2009) Structural biology of periplasmic chaperones. *Advances in Protein Chemistry and Structural Biology*, Vol. 78 (Alexander M, ed), pp. 51–97. Elsevier Academic Press Inc, San Diego, CA.

Amit AG, Mariuzza RA, Phillips SE & Poljak RJ (1986) Three-dimensional structure of an antigen-antibody complex at 2.8 A resolution. *Science* **233**: 747–753.

Ammendola S, Pasquali P, Pacello F, Rotilio G, Castor M, Libby SJ, Figueroa-Bossi N, Bossi L, Fang FC & Battistoni A (2008) Regulatory and structural differences in the Cu, Zn-superoxide dismutases of *Salmonella enterica* and their significance for virulence. *J Biol Chem* **283**: 13688–13699.

Anderson KL, Cota E, Simpson P, Chen HA, Du Merle L, Bouguenec CL & Matthews S (2004a) Complete resonance assignments of a 'donor-strand complemented' AfaE: the afimbrial adhesin from diffusely adherent *E. coli*. *J Biomol NMR* **29**: 409–410.

Anderson KL, Billington J, Pettigrew D *et al.* (2004b) An atomic resolution model for assembly, architecture, and function of the Dr adhesins. *Mol Cell* **15**: 647–657.

Andreeva A, Howorth D, Chandonia JM, Brenner SE, Hubbard TJ, Chothia C & Murzin AG (2008) Data growth and its impact on the SCOP database: new developments. *Nucleic Acids Res* **36**: D419–D425.

Aprikian P, Interlandi G, Kidd BA *et al.* (2011) The bacterial fimbrial tip acts as a mechanical force sensor. *PLoS Biol* **9**: e1000617.

Aravind L, Anantharaman V & Iyer LM (2003) Evolutionary connections between bacterial and eukaryotic signaling systems: a genomic perspective. *Curr Opin Microbiol* **6**: 490–497.

Arbabi Ghahroudi M, Desmyter A, Wyns L, Hamers R & Muyldermans S (1997) Selection and identification of single domain antibody fragments from camel heavy-chain antibodies. *FEBS Lett* **414**: 521–526.

Arbabi-Ghahroudi M, MacKenzie R & Tanha J (2009) Selection of non-aggregating VH binders from synthetic VH phage-display libraries. *Methods Mol Biol* **525**: 187–216, xiii.

Arie JP, Sassoon N & Betton JM (2001) Chaperone function of FkpA, a heat shock prolyl isomerase, in the periplasm of *Escherichia coli*. *Mol Microbiol* **39**: 199–210.

Arndt MA, Krauss J, Kipriyanov SM, Pfreundschuh M & Little M (1999) A bispecific diabody that mediates natural killer cell cytotoxicity against xenotransplantated human Hodgkin's tumors. *Blood* **94**: 2562–2568.

Arnesano F, Banci L, Bertini I, Mangani S & Thompsett AR (2003) A redox switch in CopC: an intriguing copper trafficking protein that binds copper(I) and copper(II) at different sites. *P Natl Acad Sci USA* **100**: 3814–3819.

Aujame L, Geoffroy F & Sodoyer R (1997) High affinity human antibodies by phage display. *Hum Antibodies* **8**: 155–168.

Baek H, Suk KH, Kim YH & Cha S (2002) An improved helper phage system for efficient isolation of specific antibody molecules in phage display. *Nucleic Acids Res* **30**: e18.

Banci L, Bertini I & Ciofi-Baffoni S (2009) Copper trafficking in biology: an NMR approach. *HFSP J* **3**: 165–175.

Bannister JV, Bannister WH & Rotilio G (1987) Aspects of the structure, function, and applications of superoxide dismutase. *CRC Crit Rev Biochem* **22**: 111–180.

Barclay AN (2003) Membrane proteins with immunoglobulin-like domains–a master superfamily of interaction molecules. *Semin Immunol* **15**: 215–223.

Bardwell JC, McGovern K & Beckwith J (1991) Identification of a protein required for disulfide bond formation *in vivo*. *Cell* **67**: 581–589.

Bardwell JC, Lee JO, Jander G, Martin N, Belin D & Beckwith J (1993) A pathway for disulfide bond formation *in vivo*. *P Natl Acad Sci USA* **90**: 1038–1042.

Barnard TJ, Dautin N, Lukacik P, Bernstein HD & Buchanan SK (2007) Autotransporter structure reveals intra-barrel cleavage followed by conformational changes. *Nat Struct Mol Biol* **14**: 1214–1220.

Barnhart MM, Pinkner JS, Soto GE *et al.* (2000) PapD-like chaperones provide the missing information for folding of pilin proteins. *P Natl Acad Sci USA* **97**: 7709–7714.

Barnhart MM, Sauer FG, Pinkner JS & Hultgren SJ (2003) Chaperone-subunit-usher interactions required for donor strand exchange during bacterial pilus assembly. *J Bacteriol* **185**: 2723–2730.

Batchelor M, Prasannan S, Daniell S *et al.* (2000) Structural basis for recognition of the translocated intimin receptor (Tir) by intimin from enteropathogenic *Escherichia coli*. *EMBO J* **19**: 2452–2464.

Bateman A, Eddy SR & Chothia C (1996) Members of the immunoglobulin superfamily in bacteria. *Protein Sci* **5**: 1939–1941.

Battistoni A (2003) Role of prokaryotic Cu,Zn superoxide dismutase in pathogenesis. *Biochem Soc Trans* **31**: 1326–1329.

Battistoni A, Mazzetti AP & Rotilio G (1999) *In vivo* formation of Cu,Zn superoxide dismutase disulfide bond in *Escherichia coli*. *FEBS Lett* **443**: 313–316.

Beck A, Wurch T, Bailly C & Corvaia N (2010) Strategies and challenges for the next generation of therapeutic antibodies. *Nat Rev Immunol* **10**: 345–352.

Behrens S, Maier R, de Cock H, Schmid FX & Gross CA (2001) The SurA periplasmic PPIase lacking its parvulin domains functions *in vivo* and has chaperone activity. *EMBO J* **20**: 285–294.

Benhar I (2007) Design of synthetic antibody libraries. *Expert Opin Biol Ther* **7**: 763–779.

Benov LT & Fridovich I (1994) *Escherichia coli* expresses a copper- and zinc-containing superoxide dismutase. *J Biol Chem* **269**: 25310–25314.

Bernstein HD (2007) Are bacterial 'autotransporters' really transporters? *Trends Microbiol* **15**: 441–447.

Bessette PH, Cotto JJ, Gilbert HF & Georgiou G (1999) *In vivo* and *in vitro* function of the *Escherichia coli* periplasmic cysteine oxidoreductase DsbG. *J Biol Chem* **274**: 7784–7792.

Better M, Chang CP, Robinson RR & Horwitz AH (1988) *Escherichia coli* secretion of an active chimeric antibody fragment. *Science* **240**: 1041–1043.

Biemans-Oldehinkel E, Doeven MK & Poolman B (2006) ABC transporter architecture and regulatory roles of accessory domains. *FEBS Lett* **580**: 1023–1035.

Binderup K, Mikkelsen R & Preiss J (2002) Truncation of the amino terminus of branching enzyme changes its chain transfer pattern. *Arch Biochem Biophys* **397**: 279–285.

Birch JR & Racher AJ (2006) Antibody production. *Adv Drug Deliv Rev* **58**: 671–685.

Bird RE, Hardman KD, Jacobson JW *et al.* (1988) Single-chain antigen-binding proteins. *Science* **242**: 423–426.

Bitto E & McKay DB (2002) Crystallographic structure of SurA, a molecular chaperone that facilitates folding of outer membrane porins. *Structure* **10**: 1489–1498.

Bitto E & McKay DB (2003) The periplasmic molecular chaperone protein SurA binds a peptide motif that is characteristic of integral outer membrane proteins. *J Biol Chem* **278**: 49316–49322.

Blanco-Toribio A, Muyldermans S, Frankel G & Fernández LA (2010) Direct injection of functional single-domain antibodies from *E. coli* into human cells. *PLoS One* **5**: e15227.

Bodelón G, Marín E & Fernández LA (2009) Role of periplasmic chaperones and BamA (YaeT/Omp85) in folding and secretion of intimin from enteropathogenic *Escherichia coli* strains. *J Bacteriol* **191**: 5169–5179.

Boraston AB, Bolam DN, Gilbert HJ & Davies GJ (2004) Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *Biochem J* **382**: 769–781.

Bordo D, Djinovic K & Bolognesi M (1994) Conserved patterns in the Cu,Zn superoxide dismutase family. *J Mol Biol* **238**: 366–386.

Bork P & Doolittle RF (1992) Proposed acquisition of an animal protein domain by bacteria. *P Natl Acad Sci USA* **89**: 8990–8994.

Bork P, Holm L & Sander C (1994) The immunoglobulin fold. Structural classification, sequence patterns and common core. *J Mol Biol* **242**: 309–320.

Bos MP, Robert V & Tommassen J (2007) Biogenesis of the gram-negative bacterial outer membrane. *Annu Rev Microbiol* **61**: 191–214.

Boss MA, Kenten JH, Wood CR & Emtage JS (1984) Assembly of functional antibodies from immunoglobulin heavy and light chains synthesised in *E. coli*. *Nucleic Acids Res* **12**: 3791–3806.

Bothmann H & Plückthun A (1998) Selection for a periplasmic factor improving phage display and functional periplasmic expression. *Nat Biotechnol* **16**: 376–380.

Bothmann H & Plückthun A (2000) The periplasmic *Escherichia coli* peptidylprolyl cis,trans-isomerase FkpA. I. Increased functional expression of antibody fragments with and without cis-prolines. *J Biol Chem* **275**: 17100–17105.

Bouckaert J, Berglund J, Schembri M *et al.* (2005) Receptor binding studies disclose a novel class of high-affinity inhibitors of the *Escherichia coli* FimH adhesin. *Mol Microbiol* **55**: 441–455.

Bourne Y, Redford SM, Steinman HM, Lepock JR, Tainer JA & Getzoff ED (1996) Novel dimeric interface and electrostatic recognition in bacterial Cu,Zn superoxide dismutase. *P Natl Acad Sci USA* **93**: 12774–12779.

Bouwman CW, Kohli M, Killoran A, Touchie GA, Kadner RJ & Martin NL (2003) Characterization of SrgA, a *Salmonella enterica* serovar Typhimurium virulence plasmid-encoded paralogue of the disulfide oxidoreductase DsbA, essential for biogenesis of plasmid-encoded fimbriae. *J Bacteriol* **185**: 991–1000.

Bradbury ARM, Sidhu S, Dubel S & McCafferty J (2011) Beyond natural antibodies: the power of *in vitro* display technologies. *Nat Biotechnol* **29**: 245–254.

Brandon LD & Goldberg MB (2001) Periplasmic transit and disulfide bond formation of the autotransported Shigella protein IcsA. *J Bacteriol* **183**: 951–958.

Brekke OH & Sandlie I (2003) Therapeutic antibodies for human diseases at the dawn of the twenty-first century. *Nat Rev Drug Discov* **2**: 52–62.

Brown NL, Barrett SR, Camakaris J, Lee BT & Rouch DA (1995) Molecular genetics and transport analysis of the copper-resistance determinant (pco) from *Escherichia coli* plasmid pRJ1004. *Mol Microbiol* **17**: 1153–1166.

Bullitt E & Makowski L (1995) Structural polymorphism of bacterial adhesion pili. *Nature* **373**: 164–167.

Buts L, Bouckaert J, De Genst E *et al.* (2003) The fimbrial adhesin F17-G of enterotoxigenic *Escherichia coli* has an immunoglobulin-like lectin domain that binds N-acetylglucosamine. *Mol Microbiol* **49**: 705–715.

Cabilly S, Riggs AD, Pande H *et al.* (1984) Generation of antibody activity from immunoglobulin polypeptide chains produced in *Escherichia coli*. *P Natl Acad Sci USA* **81**: 3273–3277.

Carroll MC (2008) Complement and humoral immunity. *Vaccine* **26**(Suppl 8): I28–I33.

Cha JS & Cooksey DA (1991) Copper resistance in *Pseudomonas syringae* mediated by periplasmic and outer membrane proteins. *P Natl Acad Sci USA* **88**: 8915–8919.

Chadd HE & Chamow SM (2001) Therapeutic antibody expression technology. *Curr Opin Biotechnol* **12**: 188–194.

Chames P, Hoogenboom HR & Henderikx P (2002) Selection of antibodies against biotinylated antigens. *Methods Mol Biol* **178**: 147–157.

Chattopadhyay K, Lazar-Molnar E, Yan Q *et al.* (2009) Sequence, structure, function, immunity: structural genomics of costimulation. *Immunol Rev* **229**: 356–386.

Chen R & Henning U (1996) A periplasmic protein (Skp) of *Escherichia coli* selectively binds a class of outer membrane proteins. *Mol Microbiol* **19**: 1287–1294.

Chen YL, Park S, Thornburg RW, Tabatabai LB & Kintanar A (1995) Structural characterization of the active site of *Brucella abortus* Cu-Zn superoxide dismutase: a 15N and 1H NMR investigation. *Biochemistry* **34**: 12265–12275.

Chen G, Hayhurst A, Thomas JG, Harvey BR, Iverson BL & Georgiou G (2001) Isolation of high-affinity ligand-binding proteins by periplasmic expression with cytometric screening (PECS). *Nat Biotechnol* **19**: 537–542.

Chen C, Snedecor B, Nishihara JC *et al.* (2004) High-level accumulation of a recombinant antibody fragment in the periplasm of *Escherichia coli* requires a triple-mutant (degP prc spr) host strain. *Biotechnol Bioeng* **85**: 463–474.

Cho SH & Beckwith J (2009) Two snapshots of electron transport across the membrane: insights into the structure and function of DsbD. *J Biol Chem* **284**: 11416–11424.

Choi JH & Lee SY (2004) Secretory and extracellular production of recombinant proteins using *Escherichia coli*. *Appl Microbiol Biotechnol* **64**: 625–635.

Choudhury D, Thompson A, Stojanoff V, Langermann S, Pinkner J, Hultgren SJ & Knight SD (1999) X-ray structure of the FimC-FimH chaperone-adhesin complex from uropathogenic *Escherichia coli*. *Science* **285**: 1061–1066.

Chung J, Chen T & Missiakas D (2000) Transfer of electrons across the cytoplasmic membrane by DsbD, a membrane protein involved in thiol-disulphide exchange and protein folding in the bacterial periplasm. *Mol Microbiol* **35**: 1099–1109.

Cicero DO, Contessa GM, Pertinhez TA *et al.* (2007) Solution structure of ApaG from *Xanthomonas axonopodis* pv. citri reveals a fibronectin-3 fold. *Proteins* **67**: 490–500.

Clackson T, Hoogenboom HR, Griffiths AD & Winter G (1991) Making antibody fragments using phage display libraries. *Nature* **352**: 624–628.

Clarke BR, Richards MR, Greenfield LK, Hou D, Lowary TL & Whitfield C (2011) *In vitro* reconstruction of the chain termination reaction in biosynthesis of the *Escherichia coli* O9a O-polysaccharide; the chain-length regulator, WbdD, catalyzes the addition of methyl phosphate to the non-reducing terminus of the growing glycan. *J Biol Chem* **11**: 11.

Cloutier SM, Couty S, Terskikh A *et al.* (2000) Streptabody, a high avidity molecule made by tetramerization of *in vivo* biotinylated, phage display-selected scFv fragments on streptavidin. *Mol Immunol* **37**: 1067–1077.

Colby DW, Garg P, Holden T *et al.* (2004) Development of a human light chain variable domain (V(L)) intracellular antibody specific for the amino terminus of huntingtin via yeast surface display. *J Mol Biol* **342**: 901–912.

Collet JF, Riemer J, Bader MW & Bardwell JC (2002) Reconstitution of a disulfide isomerization system. *J Biol Chem* **277**: 26886–26892.

Conrath KE, Lauwereys M, Galleni M *et al.* (2001) Beta-lactamase inhibitors derived from single-domain antibody fragments elicited in the camelidae. *Antimicrob Agents Chemother* **45**: 2807–2812.

Coppieters K, Dreier T, Silence K *et al.* (2006) Formatted anti-tumor necrosis factor alpha VHH proteins derived from camelids show superior potency and targeting to inflamed joints in a murine model of collagen-induced arthritis. *Arthritis Rheum* **54**: 1856–1866.

Cornelis GR (2006) The type III secretion injectisome. *Nat Rev Microbiol* **4**: 811–825.

Cossins AJ, Harrison S, Popplewell AG & Gore MG (2007) Recombinant production of a VL single domain antibody in *Escherichia coli* and analysis of its interaction with peptostreptococcal protein L. *Protein Expr Purif* **51**: 253–259.

Cota E, Jones C, Simpson P *et al.* (2006) The solution structure of the invasive tip complex from Afa/Dr fibrils. *Mol Microbiol* **62**: 356–366.

Cuesta AM, Sanchez-Martin D, Sanz L *et al.* (2009) *In vivo* tumor targeting and imaging with engineered trivalent antibody fragments containing collagen-derived sequences. *PLoS One* **4**: e5381.

Cuesta ÁM, Sainz-Pastor N, Bonet J, Oliva B & Álvarez-Vallina L (2010) Multivalent antibodies: when design surpasses evolution. *Trends Biotechnol* **28**: 355–362.

Culotta VC, Yang M & O'Halloran TV (2006) Activation of superoxide dismutases: putting the metal to the pedal. *Biochim Biophys Acta* **1763**: 747–758.

Cuthbertson L, Powers J & Whitfield C (2005) The C-terminal domain of the nucleotide-binding domain protein Wzt determines substrate specificity in the ATP-binding cassette transporter for the lipopolysaccharide O-antigens in *Escherichia coli* serotypes O8 and O9a. *J Biol Chem* **280**: 30310–30319.

Cuthbertson L, Kimber MS & Whitfield C (2007) Substrate binding by a bacterial ABC transporter involved in polysaccharide export. *P Natl Acad Sci USA* **104**: 19529–19534.

Cuthbertson L, Kos V & Whitfield C (2010) ABC transporters involved in export of cell surface glycoconjugates. *Microbiol Mol Biol Rev* **74**: 341–362.

Dalbey RE & Kuhn A (2012) Protein traffic in Gram-negative bacteria – how exported and secreted proteins find their way. *FEMS Microbiol Rev* in press. DOI: 10.1111/j.1574-6976.2012.00327.x

Daugherty PS, Chen G, Olsen MJ, Iverson BL & Georgiou G (1998) Antibody affinity maturation using bacterial surface display. *Protein Eng* **11**: 825–832.

Daus ML, Grote M & Schneider E (2009) The MalF P2 loop of the ATP-binding cassette transporter MalFGK2 from *Escherichia coli* and *Salmonella enterica* serovar Typhimurium interacts with maltose binding protein (MalE) throughout the catalytic cycle. *J Bacteriol* **191**: 754–761.

Dautin N & Bernstein HD (2007) Protein secretion in gram-negative bacteria via the autotransporter pathway. *Annu Rev Microbiol* **61**: 89–112.

Davidson AL, Dassa E, Orelle C & Chen J (2008) Structure, function, and evolution of bacterial ATP-binding cassette systems. *Microbiol Mol Biol Rev* **72**: 317–364, table of contents.

Davies J & Riechmann L (1994) 'Camelising' human antibody fragments: NMR studies on VH domains. *FEBS Lett* **339**: 285–290.

Davies J & Riechmann L (1995) Antibody VH domains as small recognition units. *Biotechnology (N Y)* **13**: 475–479.

De Greve H, Wyns L & Bouckaert J (2007) Combining sites of bacterial fimbriae. *Curr Opin Struct Biol* **17**: 506–512.

De Marco A (2011) Biotechnological applications of recombinant single-domain antibody fragments. *Microb Cell Fact* **10**: 44.

Debarbieux L, Bohin A & Bohin JP (1997) Topological analysis of the membrane-bound glucosyltransferase, MdoH, required for osmoregulated periplasmic glucan synthesis in *Escherichia coli*. *J Bacteriol* **179**: 6692–6698.

Decanniere K, Desmyter A, Lauwereys M, Ghahroudi MA, Muyldermans S & Wyns L (1999) A single-domain antibody fragment in complex with RNase A: non-canonical loop structures and nanomolar affinity using two CDR loops. *Structure* **7**: 361–370.

Depuydt M, Leonard SE, Vertommen D *et al.* (2009) A periplasmic reducing system protects single cysteine residues from oxidation. *Science* **326**: 1109–1111.

Desjarlais JR & Lazar GA (2011) Modulation of antibody effector function. *Exp Cell Res* **317**: 1278–1285.

Desmyter A, Transue TR, Ghahroudi MA *et al.* (1996) Crystal structure of a camel single-domain VH antibody fragment in complex with lysozyme. *Nat Struct Biol* **3**: 803–811.

Desmyter A, Spinelli S, Payan F, Lauwereys M, Wyns L, Muyldermans S & Cambillau C (2002) Three camelid VHH domains in complex with porcine pancreatic alpha-amylase. Inhibition and versatility of binding topology. *J Biol Chem* **277**: 23645–23650.

Devillers CH, Piper ME, Ballicora MA & Preiss J (2003) Characterization of the branching patterns of glycogen branching enzyme truncated on the N-terminus. *Arch Biochem Biophys* **418**: 34–38.

Deyev SM, Waibel R, Lebedenko EN, Schubiger AP & Pluckthun A (2003) Design of multivalent complexes using the barnase*barstar module. *Nat Biotechnol* **21**: 1486–1492.

Diaz M, Stanfield RL, Greenberg AS & Flajnik MF (2002) Structural analysis, selection, and ontogeny of the shark new antigen receptor (IgNAR): identification of a new locus preferentially expressed in early development. *Immunogenetics* **54**: 501–512.

Djoko KY, Xiao Z, Huffman DL & Wedd AG (2007) Conserved mechanism of copper binding and transfer. A comparison of the copper-resistance proteins PcoC from *Escherichia coli* and CopC from *Pseudomonas syringae*. *Inorg Chem* **46**: 4560–4568.

Djoko KY, Xiao Z & Wedd AG (2008) Copper resistance in *E. coli:* the multicopper oxidase PcoA catalyzes oxidation of copper(I) in Cu(I)Cu(II)-PcoC. *ChemBioChem* **9**: 1579–1582.

Dodson KW, Jacob-Dubuisson F, Striker RT & Hultgren SJ (1993) Outer-membrane PapC molecular usher discriminately recognizes periplasmic chaperone-pilus subunit complexes. *P Natl Acad Sci USA* **90**: 3670–3674.

Dodson KW, Pinkner JS, Rose T, Magnusson G, Hultgren SJ & Waksman G (2001) Structural basis of the interaction of the pyelonephritic *E. coli* adhesin to its human kidney receptor. *Cell* **105**: 733–743.

Dooley H, Flajnik MF & Porter AJ (2003) Selection and characterization of naturally occurring single-domain (IgNAR) antibody fragments from immunized sharks by phage display. *Mol Immunol* **40**: 25–33.

D'Orazio M, Scotti R, Nicolini L, Cervoni L, Rotilio G, Battistoni A & Gabbianelli R (2008) Regulatory and structural properties differentiating the chromosomal and the bacteriophage-associated *Escherichia coli* O157:H7 Cu, Zn superoxide dismutases. *BMC Microbiol* **8**: 166.

Drew SC, Djoko KY, Zhang L *et al.* (2008) Electron paramagnetic resonance characterization of the copper-resistance protein PcoC from *Escherichia coli*. *J Biol Inorg Chem* **13**: 899–907.

Driessen AJ & Nouwen N (2008) Protein translocation across the bacterial cytoplasmic membrane. *Annu Rev Biochem* **77**: 643–667.

Dubel S, Breitling F, Kontermann R, Schmidt T, Skerra A & Little M (1995) Bifunctional and multimeric complexes of streptavidin fused to single chain antibodies (scFv). *J Immunol Methods* **178**: 201–209.

Duguay AR & Silhavy TJ (2004) Quality control in the bacterial periplasm. *Biochimica Biophys Acta* **1694**: 121–134.

Dumoulin M, Conrath K, Van Meirhaeghe A *et al.* (2002) Single-domain antibody fragments with high conformational stability. *Protein Sci* **11**: 500–515.

Dutton RJ, Boyd D, Berkmen M & Beckwith J (2008) Bacterial species exhibit diversity in their mechanisms and capacity for protein disulfide bond formation. *P Natl Acad Sci USA* **105**: 11933–11938.

Elgert KD (1998) Antibody structure and function. *Immunology: Understanding the Immune System*. pp. 58–78. John Wiley & Sons, Inc., Hoboken, NJ.

Fabianek RA, Hennecke H & Thony-Meyer L (1998) The active-site cysteines of the periplasmic thioredoxin-like protein CcmG of *Escherichia coli* are important but not essential for cytochrome c maturation *in vivo*. *J Bacteriol* **180**: 1947–1950.

Fabianek RA, Hofer T & Thony-Meyer L (1999) Characterization of the *Escherichia coli* CcmH protein reveals new insights into the redox pathway required for cytochrome c maturation. *Arch Microbiol* **171**: 92–100.

Fairman JW, Dautin N, Wojtowicz D *et al.* (2012) Crystal Structures of the Outer Membrane Domain of Intimin and Invasin from Enterohemorrhagic *E. coli* and Enteropathogenic *Y. pseudotuberculosis*. *Structure*, doi:10.1016/j.str.2012.04.011.

Fan CY, Huang CC, Chiu WC, Lai CC, Liou GG, Li HC & Chou MY (2008) Production of multivalent protein binders using a self-trimerizing collagen-like peptide scaffold. *FASEB J* **22**: 3795–3804.

Fang FC, DeGroote MA, Foster JW *et al.* (1999) Virulent *Salmonella* typhimurium has two periplasmic Cu, Zn-superoxide dismutases. *P Natl Acad Sci USA* **96**: 7502–7507.

Fernández LA, Sola I, Enjuanes L & de Lorenzo V (2000) Specific secretion of active single-chain Fv antibodies into the supernantants of *Escherichia coli* cultures by use of the hemolysin system. *Appl Environ Microbiol* **66**: 5024–5029.

Finn RD, Tate J, Mistry J *et al.* (2008) The Pfam protein families database. *Nucleic Acids Res* **36**: D281–D288.

Flajnik MF, Deschacht N & Muyldermans S (2011) A case of convergence: why did a simple alternative to canonical antibodies arise in sharks and camels? *PLoS Biol* **9**: e1001120.

Ford B, Rego AT, Ragan TJ *et al.* (2010) Structural homology between the C-terminal domain of the PapC usher and its plug. *J Bacteriol* **192**: 1824–1831.

Fowler SB & Clarke J (2001) Mapping the folding pathway of an immunoglobulin domain: structural detail from Phi value analysis and movement of the transition state. *Structure* **9**: 355–366.

Fraile S, Munoz A, de Lorenzo V & Fernandez LA (2004) Secretion of proteins with dimerization capacity by the haemolysin type I transport system of *Escherichia coli*. *Mol Microbiol* **53**: 1109–1121.

Francisco JA, Campbell R, Iverson BL & Georgiou G (1993) Production and fluorescence-activated cell sorting of *Escherichia coli* expressing a functional antibody fragment on the external surface. *Proc Natl Acad Sci USA* **90**: 10444–10448.

Frankel G & Phillips AD (2008) Attaching effacing *Escherichia coli* and paradigms of Tir-triggered actin polymerization: getting off the pedestal. *Cell Microbiol* **10**: 549–556.

Frankel G, Candy DC, Fabiani E *et al.* (1995) Molecular characterization of a carboxy-terminal eukaryotic-cell-binding domain of intimin from enteropathogenic *Escherichia coli*. *Infect Immun* **63**: 4323–4328.

Fraser JS, Yu Z, Maxwell KL & Davidson AR (2006) Ig-like domains on bacteriophages: a tale of promiscuity and deceit. *J Mol Biol* **359**: 496–507.

Fraser JS, Maxwell KL & Davidson AR (2007) Immunoglobulin-like domains on bacteriophage: weapons of modest damage? *Curr Opin Microbiol* **10**: 382–387.

Fridovich I (1995) Superoxide radical and superoxide dismutases. *Annu Rev Biochem* **64**: 97–112.

Friedrich L, Stangl S, Hahne H, Kuster B, Kohler P, Multhoff G & Skerra A (2010) Bacterial production and functional characterization of the Fab fragment of the murine IgG1/lambda monoclonal antibody cmHsp70.1, a reagent for tumour diagnostics. *Protein Eng Des Sel* **23**: 161–168.

Furuta M, Uchikawa M, Ueda Y *et al.* (1998) Construction of mono- and bivalent human single-chain Fv fragments against the D antigen in the Rh blood group: multimerization effect on cell agglutination and application to blood typing. *Protein Eng* **11**: 233–241.

Galan JE (2009) Common themes in the design and function of bacterial effectors. *Cell Host Microbe* **5**: 571–579.

Galan JE & Wolf-Watz H (2006) Protein delivery into eukaryotic cells by type III secretion machines. *Nature* **444**: 567–573.

Garcillan-Barcia MP, Jurado P, Gonzalez-Perez B, Moncalian G, Fernández LA & de la Cruz F (2007) Conjugative transfer can be inhibited by blocking relaxase activity within recipient cells with intrabodies. *Mol Microbiol* **63**: 404–416.

Garmendia J, Frankel G & Crepin VF (2005) Enteropathogenic and enterohemorrhagic *Escherichia coli* infections: translocation, translocation, translocation. *Infect Immun* **73**: 2573–2585.

Glockshuber R, Schimidt T & Plückthun A (1992) The disulfide bonds in antibody variable domains: effects on stability, folding *in vitro*, and functional expression in *Escherichia coli*. *Biochemistry* **31**: 1270–1279.

Goeddel DV, Kleid DG, Bolivar F *et al.* (1979) Expression in *Escherichia coli* of chemically synthesized genes for human insulin. *P Natl Acad Sci USA* **76**: 106–110.

Gort AS, Ferber DM & Imlay JA (1999) The regulation and role of the periplasmic copper, zinc superoxide dismutase of *Escherichia coli*. *Mol Microbiol* **32**: 179–191.

Gossert AD, Bettendorff P, Puorger C, Vetsch M, Herrmann T, Glockshuber R & Wuthrich K (2008) NMR structure of the *Escherichia coli* type 1 pilus subunit FimF and its interactions with other pilus subunits. *J Mol Biol* **375**: 752–763.

Goulding CW, Sawaya MR, Parseghian A, Lim V, Eisenberg D & Missiakas D (2002) Thiol-disulfide exchange in an immunoglobulin-like fold: structure of the N-terminal domain of DsbD. *Biochemistry* **41**: 6920–6927.

Govaert J, Pellis M, Deschacht N, Vincke C, Conrath K, Muyldermans S & Saerens D (2011) Dual beneficial effect of an interloop disulfide bond for single domain antibody fragments. *J Biol Chem* **287**: 1970–1979.

Griffiths AD, Williams SC, Hartley O *et al.* (1994) Isolation of high affinity human antibodies directly from large synthetic repertoires. *EMBO J* **13**: 3245–3260.

Guillen D, Sanchez S & Rodriguez-Sanoja R (2010) Carbohydrate-binding domains: multiplicity of biological roles. *Appl Microbiol Biotechnol* **85**: 1241–1249.

Haebel PW, Goldstone D, Katzen F, Beckwith J & Metcalf P (2002) The disulfide bond isomerase DsbC is activated by an immunoglobulin-fold thiol oxidoreductase: crystal structure of the DsbC-DsbDalpha complex. *EMBO J* **21**: 4774–4784.

Hagan CL, Silhavy TJ & Kahne D (2011) beta-Barrel membrane protein assembly by the Bam Complex. *Annu Rev Biochem* **80**: 189–210.

Hagihara Y, Mine S & Uegaki K (2007) Stabilization of an immunoglobulin fold domain by an engineered disulfide

bond at the buried hydrophobic region. *J Biol Chem* **282**: 36489–36495.
Hahn E, Wild P, Hermanns U *et al.* (2002) Exploring the 3D molecular architecture of *Escherichia coli* type 1 pili. *J Mol Biol* **323**: 845–857.
Hakim R & Benhar I (2009) "Inclonals": IgGs and IgG-enzyme fusion proteins produced in an *E. coli* expression-refolding system. *MAbs* **1**: 281–287.
Halaby DM & Mornon JP (1998) The immunoglobulin superfamily: an insight on its tissular, species, and functional diversity. *J Mol Evol* **46**: 389–400.
Halaby DM, Poupon A & Mornon J (1999) The immunoglobulin fold family: sequence analysis and 3D structure comparisons. *Protein Eng* **12**: 563–571.
Hamburger ZA, Brown MS, Isberg RR & Bjorkman PJ (1999) Crystal structure of invasin: a bacterial integrin-binding protein. *Science* **286**: 291–295.
Hamers-Casterman C, Atarhouch T, Muyldermans S *et al.* (1993) Naturally occurring antibodies devoid of light chains. *Nature* **363**: 446–448.
Hanoulle X, Rollet E, Clantin B *et al.* (2004) Structural analysis of *Escherichia coli* OpgG, a protein required for the biosynthesis of osmoregulated periplasmic glucans. *J Mol Biol* **342**: 195–205.
Harms N, Koningstein G, Dontje W, Muller M, Oudega B, Luirink J & de Cock H (2001) The early interaction of the outer membrane protein phoe with the periplasmic chaperone Skp occurs at the cytoplasmic membrane. *J Biol Chem* **276**: 18804–18811.
Harmsen MM & De Haard HJ (2007) Properties, production, and applications of camelid single-domain antibody fragments. *Appl Microbiol Biotechnol* **77**: 13–22.
Harmsen MM, Van Solt CB, Fijten HP & Van Setten MC (2005) Prolonged *in vivo* residence times of llama single-domain antibody fragments in pigs by binding to porcine immunoglobulins. *Vaccine* **23**: 4926–4934.
Harpaz Y & Chothia C (1994) Many of the immunoglobulin superfamily domains in cell adhesion molecules and surface receptors belong to a new structural set which is close to that containing variable domains. *J Mol Biol* **238**: 528–539.
Harvey BR, Georgiou G, Hayhurst A, Jeong KJ, Iverson BL & Rogers GK (2004) Anchored periplasmic expression, a versatile technology for the isolation of high-affinity antibodies from *Escherichia coli*-expressed libraries. *P Natl Acad Sci USA* **101**: 9193–9198.
Hayhurst A & Harris WJ (1999) *Escherichia coli* Skp chaperone coexpression improves solubility and phage display of single chain antibody fragments. *Protein Expr Purif* **15**: 336–343.
Helfrich W, Kroesen BJ, Roovers RC, Westers L, Molema G, Hoogenboom HR & de Leij L (1998) Construction and characterization of a bispecific diabody for retargeting T cells to human carcinomas. *Int J Cancer* **76**: 232–239.
Hendershot L, Bole D, Kohler G & Kearney JF (1987) Assembly and secretion of heavy chains that do not associate posttranslationally with immunoglobulin heavy chain-binding protein. *J Cell Biol* **104**: 761–767.
Henderson IR, Navarro-Garcia F & Nataro JP (1998) The great escape: structure and function of the autotransporter proteins. *Trends Microbiol* **6**: 370–378.
Henderson NS, So SS, Martin C, Kulkarni R & Thanassi DG (2004) Topology of the outer membrane usher PapC determined by site-directed fluorescence labeling. *J Biol Chem* **279**: 53747–53754.
Henderson KA, Streltsov VA, Coley AM *et al.* (2007) Structure of an IgNAR-AMA1 complex: targeting a conserved hydrophobic cleft broadens malarial strain recognition. *Structure* **15**: 1452–1466.
Henrissat B & Davies G (1997) Structural and sequence-based classification of glycoside hydrolases. *Curr Opin Struct Biol* **7**: 637–644.
Heras B, Kurz M, Shouldice SR & Martin JL (2007) The name's bond……disulfide bond. *Curr Opin Struct Biol* **17**: 691–698.
Heras B, Shouldice SR, Totsika M, Scanlon MJ, Schembri MA & Martin JL (2009) DSB proteins and bacterial pathogenicity. *Nat Rev Microbiol* **7**: 215–225.
Hilden I, Leggio LL, Larsen S & Poulsen P (2000) Characterization and crystallization of an active N-terminally truncated form of the *Escherichia coli* glycogen branching enzyme. *Eur J Biochem* **267**: 2150–2155.
Hilschmann N & Craig LC (1965) Amino acid sequence studies with Bence-Jones proteins. *P Natl Acad Sci USA* **53**: 1403–1409.
Hiniker A & Bardwell JC (2004) *In vivo* substrate specificity of periplasmic disulfide oxidoreductases. *J Biol Chem* **279**: 12967–12973.
Hofmann BE, Bender H & Schulz GE (1989) Three-dimensional structure of cyclodextrin glycosyltransferase from *Bacillus circulans* at 3.4 A resolution. *J Mol Biol* **209**: 793–800.
Holliger P & Hudson PJ (2005) Engineered antibody fragments and the rise of single domains. *Nat Biotechnol* **23**: 1126–1136.
Holliger P, Wing M, Pound JD, Bohlen H & Winter G (1997) Retargeting serum immunoglobulin with bispecific diabodies. *Nat Biotechnol* **15**: 632–636.
Holmgren A (1989) Thioredoxin and glutaredoxin systems. *J Biol Chem* **264**: 13963–13966.
Holmgren A & Branden CI (1989) Crystal structure of chaperone protein PapD reveals an immunoglobulin fold. *Nature* **342**: 248–251.
Holmgren A, Kuehn MJ, Branden CI & Hultgren SJ (1992) Conserved immunoglobulin-like features in a family of periplasmic pilus chaperones in bacteria. *EMBO J* **11**: 1617–1622.
Holt LJ, Basran A, Jones K, Chorlton J, Jespers LS, Brewis ND & Tomlinson IM (2008) Anti-serum albumin domain antibodies for extending the half-lives of short lived drugs. *Protein Eng Des Sel* **21**: 283–288.
Hoogenboom HR (2002) Overview of antibody phage-display technology and its applications. *Methods Mol Biol* **178**: 1–37.

Hoogenboom HR (2005) Selecting and screening recombinant antibody libraries. *Nat Biotechnol* **23**: 1105–1116.

Hoogenboom HR, de Bruine AP, Hufton SE, Hoet RM, Arends JW & Roovers RC (1998) Antibody phage display technology and its applications. *Immunotechnology* **4**: 1–20.

Horn SJ, Sorbotten A, Synstad B, Sikorski P, Sorlie M, Varum KM & Eijsink VG (2006) Endo/exo mechanism and processivity of family 18 chitinases produced by *Serratia marcescens*. *FEBS J* **273**: 491–503.

Horne SM & Young KD (1995) *Escherichia coli* and other species of the Enterobacteriaceae encode a protein similar to the family of Mip-like FK506-binding proteins. *Arch Microbiol* **163**: 357–365.

Hu S, Shively L, Raubitschek A *et al.* (1996) Minibody: a novel engineered anti-carcinoembryonic antigen antibody fragment (single-chain Fv-CH3) which exhibits rapid, high-level targeting of xenografts. *Cancer Res* **56**: 3055–3061.

Hu K, Galius V & Pervushin K (2006) Structural plasticity of peptidyl-prolyl isomerase sFkpA is a key to its chaperone function as revealed by solution NMR. *Biochemistry* **45**: 11983–11991.

Huang Y, Smith BS, Chen LX, Baxter RH & Deisenhofer J (2009) Insights into pilus assembly and secretion from the structure and functional characterization of usher PapC. *P Natl Acad Sci USA* **106**: 7403–7407.

Huffman DL, Huyett J, Outten FW, Doan PE, Finney LA, Hoffman BM & O'Halloran TV (2002) Spectroscopy of Cu (II)-PcoC and the multicopper oxidase function of PcoA, two essential components of *Escherichia coli* pco copper resistance operon. *Biochemistry* **41**: 10046–10055.

Hullmann J, Patzer SI, Romer C, Hantke K & Braun V (2008) Periplasmic chaperone FkpA is essential for imported colicin M toxicity. *Mol Microbiol* **69**: 926–937.

Humphreys DP, Weir N, Lawson A, Mountain A & Lund PA (1996) Co-expression of human protein disulphide isomerase (PDI) can increase the yield of an antibody Fab' fragment expressed in *Escherichia coli*. *FEBS Lett* **380**: 194–197.

Hung DL, Knight SD, Woods RM, Pinkner JS & Hultgren SJ (1996) Molecular basis of two subfamilies of immunoglobulin-like chaperones. *EMBO J* **15**: 3792–3805.

Hussack G, Hirama T, Ding W, MacKenzie R & Tanha J (2011) Engineered single-domain antibodies with high protease resistance and thermal stability. *PLoS One* **6**: e28218.

Huston JS, Levinson D, Mudgett-Hunter M *et al.* (1988) Protein engineering of antibody binding sites: recovery of specific activity in an anti-digoxin single-chain Fv analogue produced in *Escherichia coli*. *P Natl Acad Sci USA* **85**: 5879–5883.

Ieva R, Tian P, Peterson JH & Bernstein HD (2011) Sequential and spatially restricted interactions of assembly factors with an autotransporter beta domain. *P Natl Acad Sci USA* **108**: E383–E391.

Imlay KR & Imlay JA (1996) Cloning and analysis of sodC, encoding the copper-zinc superoxide dismutase of *Escherichia coli*. *J Bacteriol* **178**: 2564–2571.

Inaba K (2009) Disulfide bond formation system in *Escherichia coli*. *J Biochem* **146**: 591–597.

Isberg RR & Tran Van Nhieu G (1994) Binding and internalization of microorganisms by integrin receptors. *Trends Microbiol* **2**: 10–14.

Isberg RR, Voorhis DL & Falkow S (1987) Identification of invasin: a protein that allows enteric bacteria to penetrate cultured mammalian cells. *Cell* **50**: 769–778.

Isberg RR, Hamburger Z & Dersch P (2000) Signaling and invasin-promoted uptake via integrin receptors. *Microbes Infect* **2**: 793–801.

Ito K & Inaba K (2008) The disulfide bond formation (Dsb) system. *Curr Opin Struct Biol* **18**: 450–458.

Jacob-Dubuisson F, Pinkner J, Xu Z, Striker R, Padmanhaban A & Hultgren SJ (1994) PapD chaperone function in pilus biogenesis depends on oxidant and chaperone-like activities of DsbA. *P Natl Acad Sci USA* **91**: 11552–11556.

Jacobson RH, Zhang XJ, DuBose RF & Matthews BW (1994) Three-dimensional structure of beta-galactosidase from *E. coli*. *Nature* **369**: 761–766.

Jacso T, Grote M, Daus ML, Schmieder P, Keller S, Schneider E & Reif B (2009) Periplasmic loop P2 of the MalF subunit of the maltose ATP binding cassette transporter is sufficient to bind the maltose binding protein MalE. *Biochemistry* **48**: 2216–2225.

Jakob RP, Zoldak G, Aumuller T & Schmid FX (2009) Chaperone domains convert prolyl isomerases into generic catalysts of protein folding. *P Natl Acad Sci USA* **106**: 20282–20287.

Jarchow S, Luck C, Gorg A & Skerra A (2008) Identification of potential substrate proteins for the periplasmic *Escherichia coli* chaperone Skp. *Proteomics* **8**: 4987–4994.

Jedrzejczak R, Dauter Z, Dauter M *et al.* (2006) Structure of DraD invasin from uropathogenic *Escherichia coli*: a dimer with swapped beta-tails. *Acta Crystallogr D Biol Crystallogr* **62**: 157–164.

Jendeberg L, Persson B, Andersson R, Karlsson R, Uhlén M & Nilsson B (1995) Kinetic analysis of the interaction between protein a domain variants and human Fc using plasmon resonance detection. *J Mol Recognit* **8**: 270–278.

Jespers L, Schon O, Famm K & Winter G (2004) Aggregation-resistant domain antibodies selected on phage by heat denaturation. *Nat Biotechnol* **22**: 1161–1165.

Jespersen HM, MacGregor EA, Sierks MR & Svensson B (1991) Comparison of the domain-level organization of starch hydrolases and related enzymes. *Biochem J* **280**(Pt 1): 51–55.

Jiang J, Zhang X, Chen Y, Wu Y, Zhou ZH, Chang Z & Sui SF (2008) Activation of DegP chaperone-protease via formation of large cage-like oligomers upon binding to substrate proteins. *P Natl Acad Sci USA* **105**: 11939–11944.

Jones CH, Danase PN, Pinkner JS, Silhavy TJ & Hultgren SJ (1997) The chaperone-assisted membrane release and folding pathway is sensed by two signal transduction systems. *EMBO J* **16**: 6394–6406.

Juers DH, Huber RE & Matthews BW (1999) Structural comparisons of TIM barrel proteins suggest functional and evolutionary relationships between beta-galactosidase and other glycohydrolases. *Protein Sci* **8**: 122–136.

Juers DH, Jacobson RH, Wigley D, Zhang XJ, Huber RE, Tronrud DE & Matthews BW (2000) High resolution refinement of beta-galactosidase in a new crystal form reveals multiple metal-binding sites and provides a structural basis for alpha-complementation. *Protein Sci* **9**: 1685–1699.

Jung ST, Reddy ST, Kang TH, Borrok MJ, Sandlie I, Tucker PW & Georgiou G (2010) Aglycosylated IgG variants expressed in bacteria that selectively bind FcγRI potentiate tumor cell killing by monocyte-dendritic cells. *Proc Nat Acad Sci USA* **107**: 604–609.

Jung ST, Kang TH, Kelton W & Georgiou G (2011) Bypassing glycosylation: engineering aglycosylated full-length IgG antibodies for human therapy. *Curr Opin Biotechnol* **22**: 858–867.

Jurado P, Ritz D, Beckwith J, de Lorenzo V & Fernández LA (2002) Production of functional single-chain Fv antibodies in the cytoplasm of *Escherichia coli*. *J Mol Biol* **320**: 1–10.

Jurado P, Fernández LA & de Lorenzo V (2006a) *In vivo* drafting of single-chain antibodies for regulatory duty on the sigma54-promoter Pu of the TOL plasmid. *Mol Microbiol* **60**: 1218–1227.

Jurado P, de Lorenzo V & Fernández LA (2006b) Thioredoxin fusions increase folding of single chain fv antibodies in the cytoplasm of *Escherichia coli*: evidence that chaperone activity is the prime effect of thioredoxin. *J Mol Biol* **357**: 49–61.

Justice SS, Hunstad DA, Harper JR *et al.* (2005) Periplasmic peptidyl prolyl cis-trans isomerases are not essential for viability, but SurA is required for pilus biogenesis in *Escherichia coli*. *J Bacteriol* **187**: 7680–7686.

Kadokura H, Katzen F & Beckwith J (2003) Protein disulfide bond formation in prokaryotes. *Annu Rev Biochem* **72**: 111–135.

Kang HJ, Paterson NG, Gaspar AH, Ton-That H & Baker EN (2009) The Corynebacterium diphtheriae shaft pilin SpaA is built of tandem Ig-like modules with stabilizing isopeptide and disulfide bonds. *P Natl Acad Sci USA* **106**: 16967–16971.

Katzen F & Beckwith J (2000) Transmembrane electron transfer by the membrane protein DsbD occurs via a disulfide bond cascade. *Cell* **103**: 769–779.

Kawase T, Saito A, Sato T *et al.* (2004) Distribution and phylogenetic analysis of family 19 chitinases in Actinobacteria. *Appl Environ Microbiol* **70**: 1135–1144.

Kelly G, Prasannan S, Daniell S *et al.* (1999) Structure of the cell-adhesion fragment of intimin from enteropathogenic *Escherichia coli*. *Nat Struct Biol* **6**: 313–318.

Kennedy EP (1982) Osmotic regulation and the biosynthesis of membrane-derived oligosaccharides in *Escherichia coli*. *P Natl Acad Sci USA* **79**: 1092–1095.

Kenny B, DeVinney R, Stein M, Reinscheid DJ, Frey EA & Finlay BB (1997) Enteropathogenic *E. coli* (EPEC) transfers its receptor for intimate adherence into mammalian cells. *Cell* **91**: 511–520.

Khare SD, Ding F & Dokholyan NV (2003) Folding of Cu, Zn superoxide dismutase and familial amyotrophic lateral sclerosis. *J Mol Biol* **334**: 515–525.

Khoushab F & Yamabhai M (2010) Chitin research revisited. *Mar Drugs* **8**: 1988–2012.

Kipriyanov SM, Moldenhauer G, Strauss G & Little M (1998) Bispecific CD3 x CD19 diabody for T cell-mediated lysis of malignant human B cells. *Int J Cancer* **77**: 763–772.

Kirchhofer A, Helma J, Schmidthals K *et al.* (2010) Modulation of protein properties in living cells using nanobodies. *Nat Struct Mol Biol* **17**: 133–138.

Kisiela DI, Kramer JJ, Tchesnokova V, Aprikian P, Yarov-Yarovoy V, Clegg S & Sokurenko EV (2011) Allosteric catch bond properties of the FimH adhesin from *Salmonella enterica* serovar Typhimurium. *J Biol Chem* **286**: 38136–38147.

Klein J & Nikolaidis N (2005) The descent of the antibody-based immune system by gradual evolution. *P Natl Acad Sci USA* **102**: 169–174.

Klemm P & Schembri MA (2000) Fimbriae-assisted bacterial surface display of heterologous peptides. *Int J Med Microbiol* **290**: 215–221.

Knight SD & Bouckaert J (2009) Structure, function and assembly of Type I fimbriae. *Glycoscience and Microbial Adhesion*. Topics in Current Chemistry, Vol. 288, pp. 67–107. Springer, Berlin/Heidelberg.

Knight SD, Choudhury D, Hultgren S, Pinkner J, Stojanoff V & Thompson A (2002) Structure of the S pilus periplasmic chaperone SfaE at 2.2 A resolution. *Acta Crystallogr D Biol Crystallogr* **58**: 1016–1022.

Knowles TJ, Scott-Tucker A, Overduin M & Henderson IR (2009) Membrane protein architects: the role of the BAM complex in outer membrane protein assembly. *Nat Rev Microbiol* **7**: 206–214.

Knutton S, Rosenshine I, Pallen MJ *et al.* (1998) A novel EspA-associated surface organelle of enteropathogenic *Escherichia coli* involved in protein translocation into epithelial cells. *EMBO J* **17**: 2166–2176.

Koay M, Zhang L, Yang B, Maher MJ, Xiao Z & Wedd AG (2005) CopC protein from *Pseudomonas syringae*: intermolecular transfer of copper from both the copper (I) and copper(II) sites. *Inorg Chem* **44**: 5203–5205.

Kohler G & Milstein C (1975) Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature* **256**: 495–497.

Kontermann RE, Wing MG & Winter G (1997) Complement recruitment using bispecific diabodies. *Nat Biotechnol* **15**: 629–631.

Korndorfer IP, Dommel MK & Skerra A (2004) Structure of the periplasmic chaperone Skp suggests functional similarity with cytosolic chaperones despite differing architecture. *Nat Struct Mol Biol* **11**: 1015–1020.

Koronakis V, Koronakis E & Hughes C (1989) Isolation and analysis of the C-terminal signal directing export of

*Escherichia coli* hemolysin protein across both bacterial membranes. *EMBO J* **8**: 595–605.
Koronakis V, Eswaran J & Hughes C (2004) Structure and function of tolc: the bacterial exit duct for proteins and drugs. *Annu Rev Biochem* **73**: 467–489.
Korotkova N, Le Trong I, Samudrala R *et al.* (2006) Crystal structure and mutational analysis of the DaaE adhesin of *Escherichia coli*. *J Biol Chem* **281**: 22367–22377.
Korshunov S & Imlay JA (2006) Detection and quantification of superoxide formed within the periplasm of *Escherichia coli*. *J Bacteriol* **188**: 6326–6334.
Kos V & Ford RC (2009) The ATP-binding cassette family: a structural perspective. *Cell Mol Life Sci* **66**: 3111–3126.
Kretzschmar T & von Ruden T (2002) Antibody discovery: phage display. *Curr Opin Biotechnol* **13**: 598–602.
Krogfelt KA, Bergmans H & Klemm P (1990) Direct evidence that the FimH protein is the mannose-specific adhesin of *Escherichia coli* type 1 fimbriae. *Infect Immun* **58**: 1995–1998.
Krojer T, Garrido-Franco M, Huber R, Ehrmann M & Clausen T (2002) Crystal structure of DegP (HtrA) reveals a new protease-chaperone machine. *Nature* **416**: 455–459.
Krojer T, Sawa J, Schafer E, Saibil HR, Ehrmann M & Clausen T (2008) Structural basis for the regulated protease and chaperone function of DegP. *Nature* **453**: 885–890.
Kuehn MJ, Heuser J, Normark S & Hultgren SJ (1992) P pili in uropathogenic *E. coli* are composite fibres with distinct fibrillar adhesive tips. *Nature* **356**: 252–255.
Kuehn MJ, Ogg DJ, Kihlberg J, Slonim LN, Flemmer K, Bergfors T & Hultgren SJ (1993) Structural basis of pilus subunit recognition by the PapD chaperone. *Science* **262**: 1234–1241.
Lacroix JM & Bohin JP (2010) Osmoregulated periplasmic glucan polymerization requires constant protein synthesis in *Escherichia coli*. *Curr Microbiol* **61**: 396–400.
Lander ES, Linton LM, Birren B *et al.* (2001) Initial sequencing and analysis of the human genome. *Nature* **409**: 860–921.
Lauwereys M, Arbabi Ghahroudi M, Desmyter A *et al.* (1998) Potent enzyme inhibitors derived from dromedary heavy-chain antibodies. *EMBO J* **17**: 3512–3520.
Le Trong I, Aprikian P, Kidd BA *et al.* (2010) Structural basis for mechanical force regulation of the adhesin FimH via finger trap-like beta sheet twisting. *Cell* **141**: 645–655.
Leo JC, Grin I & Linke D (2012) Type V secretion: mechanism(s) of autotransport through the bacterial outer membrane. *Philos Trans R Soc Lond B Biol Sci* **367**: 1088–1101.
Leong JM, Fournier RS & Isberg RR (1990) Identification of the integrin binding domain of the *Yersinia pseudotuberculosis* invasin protein. *EMBO J* **9**: 1979–1989.
Leong JM, Morrissey PE & Isberg RR (1993) A 76-amino acid disulfide loop in the *Yersinia pseudotuberculosis* invasin protein is required for integrin receptor recognition. *J Biol Chem* **268**: 20524–20532.
Lequette Y, Odberg-Ferragut C, Bohin JP & Lacroix JM (2004) Identification of mdoD, an mdoG paralog which encodes a twin-arginine-dependent periplasmic protein that controls osmoregulated periplasmic glucan backbone structures. *J Bacteriol* **186**: 3695–3702.
Lesk AM & Chothia C (1982) Evolution of proteins formed by beta-sheets. II. The core of the immunoglobulin domains. *J Mol Biol* **160**: 325–342.
Levy R, Weiss R, Chen G, Iverson BL & Georgiou G (2001) Production of correctly folded Fab antibody fragment in the cytoplasm of *Escherichia coli* trxB gor mutants via the coexpression of molecular chaperones. *Protein Expr Purif* **23**: 338–347.
Leyton DL, Rossiter AE & Henderson IR (2012) From self sufficiency to dependence: mechanisms and factors important for autotransporter biogenesis. *Nat Rev Microbiol* **10**: 213–225.
Li SL, Liang SJ, Guo N, Wu AM & Fujita-Yamaguchi Y (2000) Single-chain antibodies against human insulin-like growth factor I receptor: expression, purification, and effect on tumor growth. *Cancer Immunol Immunother* **49**: 243–252.
Li Q, Ng TW, Dodson KW So SS, Bayle KM, Pinkner JS, Scarlata S, Hultgren SJ & Thanassi DG (2010) The differential affinity of the usher for chaperone-subunit complexes is required for assembly of complete pili. *Mol Microbiol* **76**: 159–172.
Lo Leggio L, Ernst HA, Hilden I & Larsen S (2002) A structural model for the N-terminal N1 module of *E. coli* glycogen branching enzyme. *Biologia* **57**: 108–118.
Lonberg N (2008) Fully human antibodies from transgenic mouse and phage display platforms. *Curr Opin Immunol* **20**: 450–459.
Luo Y, Frey EA, Pfuetzner RA *et al.* (2000) Crystal structure of enteropathogenic *Escherichia coli* intimin-receptor complex. *Nature* **405**: 1073–1077.
Lynch M & Kuramitsu H (2000) Expression and role of superoxide dismutases (SOD) in pathogenic bacteria. *Microbes Infect* **2**: 1245–1255.
Makino T, Skretas G, Kang TH & Georgiou G (2011) Comprehensive engineering of *Escherichia coli* for enhanced expression of IgG antibodies. *Metab Eng* **13**: 241–251.
Mapingire OS, Henderson NS, Duret G, Thanassi DG & Delcour AH (2009) Modulating effects of the plug, helix, and N- and C-terminal domains on channel properties of the PapC usher. *J Biol Chem* **284**: 36324–36333.
Marín E, Bodelón G & Fernández LA (2010) Comparative analysis of the biochemical and functional properties of C-terminal domains of autotransporters. *J Bacteriol* **192**: 5588–5602.
Marks JD, Hoogenboom HR, Bonnert TP, McCafferty J, Griffiths AD & Winter G (1991) By-passing immunization. Human antibodies from V-gene libraries displayed on phage. *J Mol Biol* **222**: 581–597.
Marra A & Isberg RR (1997) Invasin-dependent and invasin-independent pathways for translocation of *Yersinia*

*pseudotuberculosis* across the Peyer's patch intestinal epithelium. *Infect Immun* **65**: 3412–3421.
Martin JL, Bardwell JC & Kuriyan J (1993) Crystal structure of the DsbA protein required for disulphide bond formation *in vivo*. *Nature* **365**: 464–468.
Martin F, Volpari C, Steinkuhler C *et al.* (1997) Affinity selection of a camelized V(H) domain antibody inhibitor of hepatitis C virus NS3 protease. *Protein Eng* **10**: 607–614.
Martsev SP, Tsybovsky YI, Stremovskiy OA *et al.* (2004) Fusion of the antiferritin antibody VL domain to barnase results in enhanced solubility and altered pH stability. *Protein Eng Des Sel* **17**: 85–93.
Matsunaga J, Barocchi MA, Croda J *et al.* (2003) Pathogenic *Leptospira* species express surface-exposed proteins belonging to the bacterial immunoglobulin superfamily. *Mol Microbiol* **49**: 929–945.
Matthews BW (2005) The structure of *E. coli* beta-galactosidase. *C R Biol* **328**: 549–556.
Maynard J & Georgiou G (2000) Antibody engineering. *Annu Rev Biomed Eng* **2**: 339–376.
Mazor Y, Van Blarcom T, Mabry R, Iverson BL & Georgiou G (2007) Isolation of engineered, full-length antibodies from libraries expressed in *Escherichia coli*. *Nat Biotechnol* **25**: 563–565.
Mazor Y, Van Blarcom T, Iverson BL & Georgiou G (2008) E-clonal antibodies: selection of full-length IgG antibodies using bacterial periplasmic display. *Nat Protoc* **3**: 1766–1777.
Mazor Y, Van Blarcom T, Iverson BL & Georgiou G (2009) Isolation of full-length IgG antibodies from combinatorial libraries expressed in *Escherichia coli*. *Methods Mol Biol* **525**: 217–239, xiv.
Mazor Y, Van Blarcom T, Carroll S & Georgiou G (2010) Selection of full-length IgGs by tandem display on filamentous phage particles and *Escherichia coli* fluorescence-activated cell sorting screening. *FEBS J* **277**: 2291–2303.
McCord JM & Fridovich I (1969) Superoxide dismutase. An enzymic function for erythrocuprein (hemocuprein). *J Biol Chem* **244**: 6049–6055.
McDaniel TK, Jarvis KG, Donnenberg MS & Kaper JB (1995) A genetic locus of enterocyte effacement conserved among diverse enterobacterial pathogens. *P Natl Acad Sci USA* **92**: 1664–1668.
Merdanovic M, Clausen T, Kaiser M, Huber R & Ehrmann M (2011) Protein quality control in the bacterial periplasm. *Annu Rev Microbiol* **65**: 149–168.
Missiakas D & Raina S (1997) Protein folding in the bacterial periplasm. *J Bacteriol* **179**: 2465–2471.
Missiakas D, Georgopoulos C & Raina S (1994) The *Escherichia coli* dsbC (xprA) gene encodes a periplasmic protein involved in disulfide bond formation. *EMBO J* **13**: 2013–2020.
Missiakas D, Schwager F & Raina S (1995) Identification and characterization of a new disulfide isomerase-like protein (DsbD) in *Escherichia coli*. *EMBO J* **14**: 3415–3424.
Mogensen JE & Otzen DE (2005) Interactions between folding factors and bacterial outer membrane proteins. *Mol Microbiol* **57**: 326–346.
Moon HW, Whipp SC, Argenzio RA, Levine MM & Giannella RA (1983) Attaching and effacing activities of rabbit and human enteropathogenic *Escherichia coli* in pig and rabbit intestines. *Infect Immun* **41**: 1340–1351.
Mori M, Jimenez B, Piccioli M, Battistoni A & Sette M (2008) The solution structure of the monomeric copper, zinc superoxide dismutase from *Salmonella enterica*: structural insights to understand the evolution toward the dimeric structure. *Biochemistry* **47**: 12954–12963.
Moussatova A, Kandt C, O'Mara ML & Tieleman DP (2008) ATP-binding cassette transporters in *Escherichia coli*. *Biochim Biophys Acta* **1778**: 1757–1771.
Munera D, Hultgren S & Fernandez LA (2007) Recognition of the N-terminal lectin domain of FimH adhesin by the usher FimD is required for type 1 pilus biogenesis. *Mol Microbiol* **64**: 333–346.
Munera D, Palomino C & Fernandez LA (2008) Specific residues in the N-terminal domain of FimH stimulate type 1 fimbriae assembly in *Escherichia coli* following the initial binding of the adhesin to FimD usher. *Mol Microbiol* **69**: 911–925.
Munson GP, Lam DL, Outten FW & O'Halloran TV (2000) Identification of a copper-responsive two-component system on the chromosome of *Escherichia coli* K-12. *J Bacteriol* **182**: 5864–5871.
Mutuberria R, Satijn S, Huijbers A *et al.* (2004) Isolation of human antibodies to tumor-associated endothelial cell markers by *in vitro* human endothelial cell selection with phage display libraries. *J Immunol Methods* **287**: 31–47.
Muyldermans S, Atarhouch T, Saldanha J, Barbosa JA & Hamers R (1994) Sequence and structure of VH domain from naturally occurring camel heavy chain immunoglobulins lacking light chains. *Protein Eng* **7**: 1129–1135.
Muyldermans S, Cambillau C & Wyns L (2001) Recognition of antigens by single-domain antibody fragments: the superfluous luxury of paired domains. *Trends Biochem Sci* **26**: 230–235.
Muyldermans S, Baral TN, Retamozzo VC *et al.* (2009) Camelid immunoglobulins and nanobody technology. *Vet Immunol Immunopathol* **128**: 178–183.
Nesta B, Spraggon G, Alteri C *et al.* (2012) FdeC, a novel broadly conserved *Escherichia coli* adhesin eliciting protection against urinary tract infections. *mBio* **3**: e00010–12.
Ng TW, Akman L, Osisami M & Thanassi DG (2004) The usher N terminus is the initial targeting site for chaperone-subunit complexes and participates in subsequent pilus biogenesis events. *J Bacteriol* **186**: 5321–5331.
Niemann HH, Schubert WD & Heinz DW (2004) Adhesins and invasins of pathogenic bacteria: a structural view. *Microbes Infect* **6**: 101–112.
Nimmerjahn F & Ravetch JV (2008) Fcgamma receptors as regulators of immune responses. *Nat Rev Immunol* **8**: 34–47.
Nishiyama M & Glockshuber R (2010) The outer membrane usher guarantees the formation of functional pili by

selectively catalyzing donor-strand exchange between subunits that are adjacent in the mature pilus. *J Mol Biol* **396**: 1–8.

Nishiyama M, Vetsch M, Puorger C, Jelesarov I & Glockshuber R (2003) Identification and characterization of the chaperone-subunit complex-binding domain from the type 1 pilus assembly platform FimD. *J Mol Biol* **330**: 513–525.

Nishiyama M, Horst R, Eidam O *et al.* (2005) Structural basis of chaperone-subunit complex recognition by the type 1 pilus assembly platform FimD. *EMBO J* **24**: 2075–2086.

Nishiyama M, Ishikawa T, Rechsteiner H & Glockshuber R (2008) Reconstitution of pilus assembly reveals a bacterial outer membrane catalyst. *Science* **320**: 376–379.

Nissim A, Hoogenboom HR, Tomlinson IM, Flynn G, Midgley C, Lane D & Winter G (1994) Antibody fragments from a 'single pot' phage display library as immunochemical reagents. *EMBO J* **13**: 692–698.

O'Connell D, Becerril B, Roy-Burman A, Daws M & Marks JD (2002) Phage versus phagemid libraries for generation of human monoclonal antibodies. *J Mol Biol* **321**: 49–56.

Oh M-Y, Joo H-Y, Hur B-U, Jeong Y-H & Cha S-H (2007) Enhancing phage display of antibody fragments using gIII-amber suppression. *Gene* **386**: 81–89.

Oldham ML & Chen J (2011) Crystal structure of the maltose transporter in a pretranslocation intermediate state. *Science* **332**: 1202–1205.

Oldham ML, Khare D, Quiocho FA, Davidson AL & Chen J (2007) Crystal structure of a catalytic intermediate of the maltose transporter. *Nature* **450**: 515–521.

Oomen CJ, van Ulsen P, van Gelder P, Feijen M, Tommassen J & Gros P (2004) Structure of the translocator domain of a bacterial autotransporter. *EMBO J* **23**: 1257–1266.

Orengo CA, Jones DT & Thornton JM (1994) Protein superfamilies and domain superfolds. *Nature* **372**: 631–634.

Orengo CA, Michie AD, Jones S, Jones DT, Swindells MB & Thornton JM (1997) CATH–a hierarchic classification of protein domain structures. *Structure* **5**: 1093–1108.

Ortega J, Iwanczyk J & Jomaa A (2009) *Escherichia coli* DegP: a structure-driven functional model. *J Bacteriol* **191**: 4705–4713.

Otey CA, Dixon R, Stack C & Goicoechea SM (2009) Cytoplasmic Ig-domain proteins: cytoskeletal regulators with a role in human disease. *Cell Motil Cytoskeleton* **66**: 618–634.

Pack P & Pluckthun A (1992) Miniantibodies: use of amphipathic helices to produce functional, flexibly linked dimeric FV fragments with high avidity in *Escherichia coli*. *Biochemistry* **31**: 1579–1584.

Padlan EA (1996) X-ray crystallography of antibodies. *Adv Protein Chem* **49**: 57–133.

Pallesen L, Poulsen LK, Christiansen G & Klemm P (1995) Chimeric FimH adhesin of type 1 fimbriae: a bacterial surface display system for heterologous sequences. *Microbiology* **141**(Pt 11): 2839–2848.

Palomino C, Marín E & Fernández LA (2011) The fimbrial usher FimD follows the sura-bamb pathway for its assembly in the outer membrane of *Escherichia coli*. *J Bacteriol* **193**: 5222–5230.

Palomo M, Kralj S, van der Maarel MJ & Dijkhuizen L (2009) The unique branching patterns of *Deinococcus* glycogen branching enzymes are determined by their N-terminal domains. *Appl Environ Microbiol* **75**: 1355–1362.

Pellecchia M, Guntert P, Glockshuber R & Wuthrich K (1998) NMR solution structure of the periplasmic chaperone FimC. *Nat Struct Biol* **5**: 885–890.

Peréz-Martínez D, Tanaka T & Rabbitts TH (2010) Intracellular antibodies and cancer: new technologies offer therapeutic opportunities. *BioEssays* **32**: 589–598.

Perisic O, Webb PA, Holliger P, Winter G & Williams RL (1994) Crystal structure of a diabody, a bivalent antibody fragment. *Structure* **2**: 1217–1226.

Perrakis A, Tews I, Dauter Z, Oppenheim AB, Chet I, Wilson KS & Vorgias CE (1994) Crystal structure of a bacterial chitinase at 2.3 A resolution. *Structure* **2**: 1169–1180.

Perrakis A, Ouzounis C & Wilson KS (1997) Evolution of immunoglobulin-like modules in chitinases: their structural flexibility and functional implications. *Fold Des* **2**: 291–294.

Perry JJ, Shin DS, Getzoff ED & Tainer JA (2010) The structural biochemistry of the superoxide dismutases. *Biochim Biophys Acta* **1804**: 245–262.

Pesce A, Capasso C, Battistoni A, Folcarelli S, Rotilio G, Desideri A & Bolognesi M (1997) Unique structural features of the monomeric Cu,Zn superoxide dismutase from *Escherichia coli*, revealed by X-ray crystallography. *J Mol Biol* **274**: 408–420.

Pesce A, Battistoni A, Stroppolo ME *et al.* (2000) Functional and crystallographic characterization of *Salmonella typhimurium* Cu,Zn superoxide dismutase coded by the sodCI virulence gene. *J Mol Biol* **302**: 465–478.

Pettigrew D, Anderson KL, Billington J *et al.* (2004) High resolution studies of the Afa/Dr adhesin DraE and its interaction with chloramphenicol. *J Biol Chem* **279**: 46851–46857.

Phan G, Remaut H, Wang T *et al.* (2011) Crystal structure of the FimD usher bound to its cognate FimC-FimH substrate. *Nature* **474**: 49–53.

Piatek R, Bruzdziak P, Wojciechowski M, Zalewska-Piatek B & Kur J (2010) The non-canonical disulfide bond as the important stabilizing element of immunoglobuline fold of Dr Fimbrial DraE subunit. *Biochemistry* **49**: 1460–1468.

Plückthun A, Krebber C, Krebber U *et al.* (1996) Producing antibodies in *Escherichia coli*: from PCR to fermentation. *Antibody Engineering: A Practical Approach* (McCafferty J, Hoogenboom HR & Chiswell DJ, eds), pp. 203–252. IRL Press, Oxford.

Pohlner J, Halter R, Beyreuther K & Meyer TF (1987) Gene structure and extracellular secretion of *Neisseria gonorrhoeae* IgA protease. *Nature* **325**: 458–462.

Poljak RJ, Amzel LM, Avey HP, Chen BL, Phizackerley RP & Saul F (1973) Three-dimensional structure of the Fab' fragment of a human immunoglobulin at 2,8-A resolution. *P Natl Acad Sci USA* **70**: 3305–3310.

Ponnuraj K, Bowden MG, Davis S *et al.* (2003) A "dock, lock, and latch" structural model for a staphylococcal adhesin binding to fibrinogen. *Cell* **115**: 217–228.

Prelli F & Frangione B (1992) Franklin's disease: Ig gamma 2 H chain mutant BUR. *J Immunol* **148**: 949–952.

Proba K, Worn A, Honegger A & Pluckthun A (1998) Antibody scFv fragments without disulfide bonds made by molecular evolution. *J Mol Biol* **275**: 245–253.

Procko E, O'Mara ML, Bennett WF, Tieleman DP & Gaudet R (2009) The mechanism of ABC transporters: general lessons from structural and functional studies of an antigenic peptide transporter. *FASEB J* **23**: 1287–1302.

Proft T & Baker EN (2009) Pili in Gram-negative and Gram-positive bacteria – structure, assembly and their role in disease. *Cell Mol Life Sci* **66**: 613–635.

Puig S & Thiele DJ (2002) Molecular mechanisms of copper uptake and distribution. *Curr Opin Chem Biol* **6**: 171–180.

Puorger C, Vetsch M, Wider G & Glockshuber R (2011) Structure, folding and stability of FimA, the main structural subunit of type 1 pili from uropathogenic *Escherichia coli* strains. *J Mol Biol* **412**: 520–535.

Purdy GE, Fisher CR & Payne SM (2007) IcsA surface presentation in *Shigella flexneri* requires the periplasmic chaperones DegP, Skp, and SurA. *J Bacteriol* **189**: 5566–5573.

Qi H, Lu H, Qiu H-J, Petrenko V & Liu A (2012) Phagemid vectors for phage display: properties, characteristics and construction. *J Mol Biol* **417**: 129–143.

Quinternet M, Tsan P, Selme-Roussel L, Jacob C, Boschi-Muller S, Branlant G & Cung MT (2009) Formation of the complex between DsbD and PilB N-terminal domains from *Neisseria meningitidis* necessitates an adaptability of nDsbD. *Structure* **17**: 1024–1033.

Raetz CR & Whitfield C (2002) Lipopolysaccharide endotoxins. *Annu Rev Biochem* **71**: 635–700.

Raju TS (2008) Terminal sugars of Fc glycans influence antibody effector functions of IgGs. *Curr Opin Immunol* **20**: 471–478.

Rakonjac J, Jovanovic G & Model P (1997) Filamentous phage infection-mediated gene expression: construction and propagation of the gIII deletion mutant helper phage R408d3. *Gene* **198**: 99–103.

Ramm K & Plückthun A (2000) The periplasmic *Escherichia coli* peptidylprolyl cis,trans-isomerase FkpA. II. Isomerase-independent chaperone activity *in vitro*. *J Biol Chem* **275**: 17106–17113.

Rao S, Hu S, McHugh L *et al.* (2005) Toward a live microbial microbicide for HIV: commensal bacteria secreting an HIV fusion inhibitor peptide. *P Natl Acad Sci USA* **102**: 11993–11998.

Reilly DE & Yansura DG (2010) Production of monoclonal antibodies in *E. coli*. *Current Trends in Monoclonal Antibody Development and Manufacturing*, Vol. XI (Shire SJ, Gombotz W, Bechtold-Peters K & Andya J, eds), pp. 295–308. Springer, New York.

Reiter Y, Schuck P, Boyd LF & Plaksin D (1999) An antibody single-domain phage display library of a native heavy chain variable region: isolation of functional single-domain VH molecules with a unique interface. *J Mol Biol* **290**: 685–698.

Remaut H & Waksman G (2006) Protein-protein interaction through beta-strand addition. *Trends Biochem Sci* **31**: 436–444.

Remaut H, Rose RJ, Hannan TJ, Hultgren SJ, Radford SE, Ashcroft AE & Waksman G (2006) Donor-strand exchange in chaperone-assisted pilus assembly proceeds through a concerted beta strand displacement mechanism. *Mol Cell* **22**: 831–842.

Remaut H, Tang C, Henderson NS *et al.* (2008) Fiber formation across the bacterial outer membrane by the chaperone/usher pathway. *Cell* **133**: 640–652.

Ricci DP & Silhavy TJ (2011) The Bam machine: a molecular cooper. *Biochim Biophys Acta* **1818**: 1067–1084.

Richardson JS, Richardson DC, Thomas KA, Silverton EW & Davies DR (1976) Similarity of three-dimensional structure between the immunoglobulin domain and the copper, zinc superoxide dismutase subunit. *J Mol Biol* **102**: 221–235.

Rietsch A, Belin D, Martin N & Beckwith J (1996) An *in vivo* pathway for disulfide bond isomerization in *Escherichia coli*. *P Natl Acad Sci USA* **93**: 13048–13053.

Rietsch A, Bessette P, Georgiou G & Beckwith J (1997) Reduction of the periplasmic disulfide bond isomerase, DsbC, occurs by passage of electrons from cytoplasmic thioredoxin. *J Bacteriol* **179**: 6602–6608.

Ritz D & Beckwith J (2001) Roles of thiol-redox pathways in bacteria. *Annu Rev Microbiol* **55**: 21–48.

Rondot S, Koch J, Breitling F & Dubel S (2001) A helper phage to improve single-chain antibody presentation in phage display. *Nat Biotechnol* **19**: 75–78.

Roopenian DC & Akilesh S (2007) FcRn: the neonatal Fc receptor comes of age. *Nat Rev Immunol* **7**: 715–725.

Roovers RC, Laeremans T, Huang L *et al.* (2007) Efficient inhibition of EGFR signaling and of tumour growth by antagonistic anti-EFGR Nanobodies. *Cancer Immunol Immunother* **56**: 303–317.

Rosen DA, Pinkner JS, Jones JM, Walker JN, Clegg S & Hultgren SJ (2008) Utilization of an intracellular bacterial community pathway in *Klebsiella pneumoniae* urinary tract infection and the effects of FimK on type 1 pilus expression. *Infect Immun* **76**: 3337–3345.

Rossiter AE, Leyton DL, Tveen-Jensen K *et al.* (2011) The essential beta-barrel assembly machinery complex components BamD and BamA are required for autotransporter biogenesis. *J Bacteriol* **193**: 4250–4253.

Rozhkova A, Stirnimann CU, Frei P *et al.* (2004) Structural basis and kinetics of inter- and intramolecular disulfide exchange in the redox catalyst DsbD. *EMBO J* **23**: 1709–1719.

Ruiz N, Kahne D & Silhavy TJ (2006) Advances in understanding bacterial outer-membrane biogenesis. *Nat Rev Microbiol* **4**: 57–66.

Ruiz N, Kahne D & Silhavy TJ (2009) Transport of lipopolysaccharide across the cell envelope: the long road of discovery. *Nat Rev Microbiol* **7**: 677–683.

Saerens D, Ghassabeh GH & Muyldermans S (2008) Single-domain antibodies as building blocks for novel therapeutics. *Curr Opin Pharmacol* **8**: 600–608.

Salih O, Remaut H, Waksman G & Orlova EV (2008) Structural analysis of the Saf pilus by electron microscopy and image processing. *J Mol Biol* **379**: 174–187.

Sathaliyawala T, Islam MZ, Li Q, Fokine A, Rossmann MG & Rao VB (2010) Functional analysis of the highly antigenic outer capsid protein, Hoc, a virus decoration protein from T4-like bacteriophages. *Mol Microbiol* **77**: 444–455.

Sauer FG, Futterer K, Pinkner JS, Dodson KW, Hultgren SJ & Waksman G (1999) Structural basis of chaperone function and pilus biogenesis. *Science* **285**: 1058–1061.

Sauer FG, Pinkner JS, Waksman G & Hultgren SJ (2002) Chaperone priming of pilus subunits facilitates a topological transition that drives fiber formation. *Cell* **111**: 543–551.

Sauer FG, Remaut H, Hultgren SJ & Waksman G (2004) Fiber assembly by the chaperone-usher pathway. *Biochim Biophys Acta* **1694**: 259–267.

Saul FA, Arie JP, Vulliez-le Normand B, Kahn R, Betton JM & Bentley GA (2004) Structural and functional studies of FkpA from *Escherichia coli*, a cis/trans peptidyl-prolyl isomerase with chaperone activity. *J Mol Biol* **335**: 595–608.

Saulino ET, Thanassi DG, Pinkner JS & Hultgren SJ (1998) Ramifications of kinetic partitioning on usher-mediated pilus biogenesis. *EMBO J* **17**: 2177–2185.

Sawa J, Heuck A, Ehrmann M & Clausen T (2010) Molecular transformers in the cell: lessons learned from the DegP protease-chaperone. *Curr Opin Struct Biol* **20**: 253–258.

Sazinsky SL, Ott RG, Silver NW, Tidor B, Ravetch JV & Wittrup KD (2008) Aglycosylated immunoglobulin G1 variants productively engage activating Fc receptors. *P Natl Acad Sci USA* **105**: 20167–20172.

Schafer U, Beck K & Muller M (1999) Skp, a molecular chaperone of gram-negative bacteria, is required for the formation of soluble periplasmic intermediates of outer membrane proteins. *J Biol Chem* **274**: 24567–24574.

Schiefner A, Chatwell L, K√∂rner J *et al.* (2011) A disulfide-free single-domain VL intrabody with blocking activity towards huntingtin reveals a novel mode of epitope recognition. *J Mol Biol* **414**: 337–355.

Schlapschy M & Skerra A (2011) Periplasmic chaperones used to enhance functional secretion of proteins in *E. coli*. *Methods Mol Biol* **705**: 211–224.

Schmidt M & Wels W (1996) Targeted inhibition of tumour cell growth by a bispecific single-chain toxin containing an antibody domain and TGF alpha. *Br J Cancer* **74**: 853–862.

Schroeder HW Jr & Cavacini L (2010) Structure and function of immunoglobulins. *J Allergy Clin Immunol* **125**: S41–S52.

Seligmann M, Mihaesco E, Preud'homme JL, Danon F & Brouet JC (1979) Heavy chain diseases: current findings and concepts. *Immunol Rev* **48**: 145–167.

Selkrig J, Mosbahi K, Webb CT *et al.* (2012) Discovery of an archetypal protein transport system in bacterial outer membranes. *Nat Struct Mol Biol* **19**: 506–510.

Seo MJ, Jeong KJ, Leysath CE, Ellington AD, Iverson BL & Georgiou G (2009) Engineering antibody fragments to fold in the absence of disulfide bonds. *Protein Sci* **18**: 259–267.

Sevier CS & Kaiser CA (2006) Conservation and diversity of the cellular disulfide bond formation pathways. *Antioxid Redox Signal* **8**: 797–811.

Shapiro L, Kwong PD, Fannon AM, Colman DR & Hendrickson WA (1995) Considerations on the folding topology and evolutionary origin of cadherin domains. *P Natl Acad Sci USA* **92**: 6793–6797.

Simmons LC, Reilly D, Klimowski L *et al.* (2002) Expression of full-length immunoglobulins in *Escherichia coli*: rapid and efficient production of aglycosylated antibodies. *J Immunol Methods* **263**: 133–147.

Skerra A (1993) Bacterial expression of immunoglobulin fragments. *Curr Opin Immunol* **5**: 256–262.

Skerra A & Pluckthun A (1991) Secretion and *in vivo* folding of the Fab fragment of the antibody McPC603 in *Escherichia coli*: influence of disulphides and cis-prolines. *Protein Eng* **4**: 971–979.

Skillman KM, Barnard TJ, Peterson JH, Ghirlando R & Bernstein HD (2005) Efficient secretion of a folded protein domain by a monomeric bacterial autotransporter. *Mol Microbiol* **58**: 945–958.

Smith DK & Xue H (1997) Sequence profiles of immunoglobulin and immunoglobulin-like domains. *J Mol Biol* **274**: 530–545.

So SS & Thanassi DG (2006) Analysis of the requirements for pilus biogenesis at the outer membrane usher and the function of the usher C-terminus. *Mol Microbiol* **60**: 364–375.

Soltes G, Barker H, Marmai K, Pun E, Yuen A & Wiersma EJ (2003) A new helper phage and phagemid vector system improves viral display of antibody Fab fragments and avoids propagation of insert-less virions. *J Immunol Methods* **274**: 233–244.

Somerville JE Jr, Goshorn SC, Fell HP & Darveau RP (1994) Bacterial aspects associated with the expression of a single-chain antibody fragment in *Escherichia coli*. *Appl Microbiol Biotechnol* **42**: 595–603.

Soto GE & Hultgren SJ (1999) Bacterial adhesins: common themes and variations in architecture and assembly. *J Bacteriol* **181**: 1059–1071.

Spagnolo L, Toro I, D'Orazio M *et al.* (2004) Unique features of the sodC-encoded superoxide dismutase from *Mycobacterium tuberculosis*, a fully functional copper-containing enzyme lacking zinc in the active site. *J Biol Chem* **279**: 33447–33455.

Sperandeo P, Deho G & Polissi A (2009) The lipopolysaccharide transport system of Gram-negative bacteria. *Biochim Biophys Acta* **1791**: 594–602.

Srinivasan M & Roeske RW (2005) Immunomodulatory peptides from IgSF proteins: a review. *Curr Protein Pept Sci* **6**: 185–196.

Stam MR, Danchin EG, Rancurel C, Coutinho PM & Henrissat B (2006) Dividing the large glycoside hydrolase family 13 into subfamilies: towards improved functional annotations of alpha-amylase-related proteins. *Protein Eng Des Sel* **19**: 555–562.

Stanfield RL, Dooley H, Flajnik MF & Wilson IA (2004) Crystal structure of a shark single-domain antibody V region in complex with lysozyme. *Science* **305**: 1770–1773.

Steiner LA (1996) Immunoglobulin evolution, 30 years on. *Glycobiology* **6**: 649–656.

Steiner D, Forrer P, Stumpp MT & Pluckthun A (2006) Signal sequences directing cotranslational translocation expand the range of proteins amenable to phage display. *Nat Biotechnol* **24**: 823–831.

Stevens FJ (2008) Homology versus analogy: possible evolutionary relationship of immunoglobulins, cupredoxins, and Cu,Zn-superoxide dismutase. *J Mol Recognit* **21**: 20–29.

Stewart EJ, Katzen F & Beckwith J (1999) Six conserved cysteines of the membrane protein DsbD are required for the transfer of electrons from the cytoplasm to the periplasm of *Escherichia coli*. *EMBO J* **18**: 5963–5971.

Stijlemans B, Conrath K, Cortez-Retamozo V *et al.* (2004) Efficient targeting of conserved cryptic epitopes of infectious agents by single domain antibodies. African trypanosomes as paradigm. *J Biol Chem* **279**: 1256–1261.

Stirnimann CU, Rozhkova A, Grauschopf U, Grutter MG, Glockshuber R & Capitani G (2005) Structural basis and kinetics of DsbD-dependent cytochrome c maturation. *Structure* **13**: 985–993.

Stirnimann CU, Grutter MG, Glockshuber R & Capitani G (2006) nDsbD: a redox interaction hub in the *Escherichia coli* periplasm. *Cell Mol Life Sci* **63**: 1642–1648.

Strindelius L, Folkesson A, Normark S & Sjoholm I (2004) Immunogenic properties of the *Salmonella* atypical fimbriae in BALB/c mice. *Vaccine* **22**: 1448–1456.

Subrini O & Betton JM (2009) Assemblies of DegP underlie its dual chaperone and protease function. *FEMS Microbiol Lett* **296**: 143–148.

Suzuki K, Suzuki M, Taiyoji M, Nikaidou N & Watanabe T (1998) Chitin binding protein (CBP21) in the culture supernatant of *Serratia marcescens* 2170. *Biosci Biotechnol Biochem* **62**: 128–135.

Tainer JA, Getzoff ED, Beem KM, Richardson JS & Richardson DC (1982) Determination and analysis of the 2 A-structure of copper, zinc superoxide dismutase. *J Mol Biol* **160**: 181–217.

Tapia MI, Mourez M, Hofnung M & Dassa E (1999) Structure-function study of MalF protein by random mutagenesis. *J Bacteriol* **181**: 2267–2272.

Tchesnokova V, Aprikian P, Kisiela D, Gowey S, Korotkova N, Thomas W & Sokurenko E (2011) Type 1 fimbrial adhesin FimH elicits an immune response that enhances cell adhesion of *Escherichia coli*. *Infect Immun* **79**: 3895–3904.

Teichmann SA & Chothia C (2000) Immunoglobulin superfamily proteins in *Caenorhabditis elegans*. *J Mol Biol* **296**: 1367–1383.

Thanassi DG, Saulino ET & Hultgren SJ (1998) The chaperone/usher pathway: a major terminal branch of the general secretory pathway. *Curr Opin Microbiol* **1**: 223–231.

Thanassi DG, Stathopoulos C, Dodson K, Geiger D & Hultgren SJ (2002) Bacterial outer membrane ushers contain distinct targeting and assembly domains for pilus biogenesis. *J Bacteriol* **184**: 6260–6269.

Thanassi DG, Bliska JB & Christie PJ (2012) Surface organelles assembled by secretion systems of Gram-negative bacteria: diversity in structure and function. *FEMS Microbiol Rev* in press. doi: 10.1111/j.1574-6976.2012.00342.x

Tijink BM, Laeremans T, Budde M *et al.* (2008) Improved tumor targeting of anti-epidermal growth factor receptor Nanobodies through albumin binding: taking advantage of modular Nanobody technology. *Mol Cancer Ther* **7**: 2288–2297.

Tilson MD & Rzhetsky A (2000) A novel hypothesis regarding the evolutionary origins of the immunoglobulin fold. *Curr Med Res Opin* **16**: 88–93.

Tinsley CR, Voulhoux R, Beretti JL, Tommassen J & Nassif X (2004) Three homologues, including two membrane-bound proteins, of the disulfide oxidoreductase DsbA in *Neisseria meningitidis*: effects on bacterial growth and biogenesis of functional type IV pili. *J Biol Chem* **279**: 27078–27087.

Totsika M, Heras B, Wurpel DJ & Schembri MA (2009) Characterization of two homologous disulfide bond systems involved in virulence factor biogenesis in uropathogenic *Escherichia coli* CFT073. *J Bacteriol* **191**: 3901–3908.

Touze T, Hayward RD, Eswaran J, Leong JM & Koronakis V (2004) Self-association of EPEC intimin mediated by the beta-barrel-containing anchor domain: a role in clustering of the Tir receptor. *Mol Microbiol* **51**: 73–87.

Transue TR, De Genst E, Ghahroudi MA, Wyns L & Muyldermans S (1998) Camel single-domain antibody inhibits enzyme by mimicking carbohydrate substrate. *Proteins* **32**: 515–522.

Tsai JC, Yen MR, Castillo R, Leyton DL, Henderson IR & Saier MH Jr (2010) The bacterial intimins and invasins: a large and novel family of secreted proteins. *PLoS One* **5**: e14403.

Turski ML & Thiele DJ (2009) New roles for copper metabolism in cell proliferation, signaling, and disease. *J Biol Chem* **284**: 717–721.

Uchiyama T, Katouno F, Nikaidou N, Nonaka T, Sugiyama J & Watanabe T (2001) Roles of the exposed aromatic residues in crystalline chitin hydrolysis by chitinase A from *Serratia marcescens* 2170. *J Biol Chem* **276**: 41343–41349.

UniProt Consortium (2010) The universal protein resource (UniProt) in 2010. *Nucleic Acids Res* **38**: D142–D148.

Ureta AR, Endres RG, Wingreen NS & Silhavy TJ (2007) Kinetic analysis of the assembly of the outer membrane protein LamB in *Escherichia coli* mutants each lacking a

secretion or targeting factor in a different cellular compartment. *J Bacteriol* **189**: 446–454.
Vaaje-Kolstad G, Houston DR, Riemen AH, Eijsink VG & van Aalten DM (2005a) Crystal structure and binding properties of the *Serratia marcescens* chitin-binding protein CBP21. *J Biol Chem* **280**: 11313–11319.
Vaaje-Kolstad G, Horn SJ, van Aalten DM, Synstad B & Eijsink VG (2005b) The non-catalytic chitin-binding protein CBP21 from *Serratia marcescens* is essential for chitin degradation. *J Biol Chem* **280**: 28492–28497.
van den Berg B (2010) Crystal structure of a full-length autotransporter. *J Mol Biol* **396**: 627–633.
van den Beucken T, van Neer N, Sablon E, Desmet J, Celis L, Hoogenboom HR & Hufton SE (2001) Building novel binding ligands to b7.1 and b7.2 based on human antibody single variable light chain domains. *J Mol Biol* **310**: 591–601.
van der Linden RH, Frenken LG, de Geus B *et al.* (1999) Comparison of physical chemical properties of llama VHH antibody fragments and mouse monoclonal antibodies. *Biochim Biophys Acta* **1431**: 37–46.
van der Linden R, de Geus B, Stok W, Bos W, van Wassenaar D, Verrips T & Frenken L (2000) Induction of immune responses and molecular cloning of the heavy chain antibody repertoire of Lama glama. *J Immunol Methods* **240**: 185–195.
Van Molle I, Buts L, Coppens F *et al.* (2005) Crystallization of the FaeE chaperone of *Escherichia coli* F4 fimbriae. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **61**: 427–431.
Vaughn DE & Bjorkman PJ (1996) The (Greek) key to structures of neural adhesion molecules. *Neuron* **16**: 261–273.
Veiga E, de Lorenzo V & Fernández LA (1999) Probing secretion and translocation of a beta-autotransporter using a reporter single-chain Fv as a cognate passenger domain. *Mol Microbiol* **33**: 1232–1243.
Veiga E, De Lorenzo V & Fernández LA (2004) Structural tolerance of bacterial autotransporters for folded passenger protein domains. *Mol Microbiol* **52**: 1069–1080.
Vercruysse T, Pardon E, Vanstreels E, Steyaert J & Daelemans D (2010) An intrabody based on a llama single-domain antibody targeting the N-terminal alpha-helical multimerization domain of HIV-1 Rev prevents viral production. *J Biol Chem* **285**: 21768–21780.
Verger D, Miller E, Remaut H, Waksman G & Hultgren S (2006) Molecular mechanism of P pilus termination in uropathogenic *Escherichia coli*. *EMBO Rep* **7**: 1228–1232.
Verger D, Bullitt E, Hultgren SJ & Waksman G (2007) Crystal structure of the P pilus rod subunit PapA. *PLoS Pathog* **3**: e73.
Verger D, Rose RJ, Paci E *et al.* (2008) Structural determinants of polymerization reactivity of the P pilus adaptor subunit PapF. *Structure* **16**: 1724–1731.
Vetsch M, Puorger C, Spirig T, Grauschopf U, Weber-Ban EU & Glockshuber R (2004) Pilus chaperones represent a new type of protein-folding catalyst. *Nature* **431**: 329–333.
Vetsch M, Erilov D, Moliere N, Nishiyama M, Ignatov O & Glockshuber R (2006) Mechanism of fibre assembly through the chaperone-usher pathway. *EMBO Rep* **7**: 734–738.
Vincke C, Loris R, Saerens D, Martinez-Rodriguez S, Muyldermans S & Conrath K (2009) General strategy to humanize a camelid single-domain antibody and identification of a universal humanized nanobody scaffold. *J Biol Chem* **284**: 3273–3284.
Vu KB, Ghahroudi MA, Wyns L & Muyldermans S (1997) Comparison of llama VH sequences from conventional and heavy chain antibodies. *Mol Immunol* **34**: 1121–1131.
Waksman G & Hultgren SJ (2009) Structural biology of the chaperone-usher pathway of pilus biogenesis. *Nat Rev Microbiol* **7**: 765–774.
Walton TA & Sousa MC (2004) Crystal structure of Skp, a prefoldin-like chaperone that protects soluble and membrane proteins from aggregation. *Mol Cell* **15**: 367–374.
Walton TA, Sandoval CM, Fowler CA, Pardi A & Sousa MC (2009) The cavity-chaperone Skp protects its substrate from aggregation but allows independent folding of substrate domains. *P Natl Acad Sci USA* **106**: 1772–1777.
Ward ES, Gussow D, Griffiths AD, Jones PT & Winter G (1989) Binding activities of a repertoire of single immunoglobulin variable domains secreted from *Escherichia coli*. *Nature* **341**: 544–546.
Wentzel A, Christmann A, Adams T & Kolmar H (2001) Display of passenger proteins on the surface of *Escherichia coli* K-12 by the enterohemorrhagic *E. coli* intimin EaeA. *J Bacteriol* **183**: 7273–7284.
Wernimont AK, Huffman DL, Finney LA, Demeler B, O'Halloran TV & Rosenzweig AC (2003) Crystal structure and dimerization equilibria of PcoC, a methionine-rich copper resistance protein from *Escherichia coli*. *J Biol Inorg Chem* **8**: 185–194.
Wesolowski J, Alzogaray V, Reyelt J *et al.* (2009) Single domain antibodies: promising experimental and therapeutic tools in infection and immunity. *Med Microbiol Immunol* **198**: 157–174.
Williams AF (1984) The immunoglobulin superfamily takes shape. *Nature* **308**: 12–13.
Williams AF & Barclay AN (1988) The immunoglobulin superfamily–domains for cell surface recognition. *Annu Rev Immunol* **6**: 381–405.
Winter G, Griffiths AD, Hawkins RE & Hoogenboom HR (1994) Making antibodies by phage display technology. *Annu Rev Immunol* **12**: 433–455.
Worn A & Pluckthun A (1998) Mutual stabilization of VL and VH in single-chain antibody fragments, investigated with mutants engineered for stability. *Biochemistry* **37**: 13120–13127.
Worn A & Pluckthun A (2001) Stability engineering of antibody single-chain Fv fragments. *J Mol Biol* **305**: 989–1010.
Wörn A & Plückthun A (1999) Different equilibrium stability behavior of ScFv fragments: identification, classification,

and improvement by protein engineering. *Biochemistry* **38**: 8739–8750.

Xu X, Wang S, Hu Y-X & McKay DB (2007) The periplasmic bacterial molecular chaperone SurA adapts its structure to bind peptides in different conformations to assert a sequence preference for aromatic residues. *J Mol Biol* **373**: 367–381.

Yi Y, Ma Y, Gao F *et al.* (2010) Crystal structure of EHEC intimin: insights into the complementarity between EPEC and EHEC. *PLoS One* **5**: e15285.

Yu X, Visweswaran GR, Duck Z, Marupakula S, MacIntyre S, Knight SD & Zavialov AV (2009) Caf1A usher possesses a Caf1 subunit-like domain that is crucial for Caf1 fibre secretion. *Biochem J* **418**: 541–551.

Zavialov AV, Kersley J, Korpela T, Zav'yalov VP, MacIntyre S & Knight SD (2002) Donor strand complementation mechanism in the biogenesis of non-pilus systems. *Mol Microbiol* **45**: 983–995.

Zavialov AV, Berglund J, Pudney AF, Fooks LJ, Ibrahim TM, MacIntyre S & Knight SD (2003) Structure and biogenesis of the capsular F1 antigen from *Yersinia pestis*: preserved folding energy drives fiber formation. *Cell* **113**: 587–596.

Zavialov AV, Tischenko VM, Fooks LJ *et al.* (2005) Resolving the energy paradox of chaperone/usher-mediated fibre assembly. *Biochem J* **389**: 685–694.

Zavialov A, Zav'yalova G, Korpela T & Zav'yalov V (2007) FGL chaperone-assembled fimbrial polyadhesins: anti-immune armament of Gram-negative bacterial pathogens. *FEMS Microbiol Rev* **31**: 478–514.

Zav'yalov V, Zavialov A, Zav'yalova G & Korpela T (2010) Adhesive organelles of Gram-negative pathogens assembled with the classical chaperone/usher machinery: structure and function from a clinical standpoint. *FEMS Microbiol Rev* **34**: 317–378.

Zhai W, Glanville J, Fuhrmann M *et al.* (2011) Synthetic antibodies designed on natural sequence landscapes. *J Mol Biol* **412**: 55–71.

Zhang HZ & Donnenberg MS (1996) DsbA is required for stability of the type IV pilin of enteropathogenic *Escherichia coli*. *Mol Microbiol* **21**: 787–797.

Zhang Z, Li ZH, Wang F *et al.* (2002) Overexpression of DsbC and DsbG markedly improves soluble and functional expression of single-chain Fv antibodies in *Escherichia coli*. *Protein Expr Purif* **26**: 218–228.

Zhang J, Tanha J, Hirama T *et al.* (2004) Pentamerization of single-domain antibodies from phage libraries: a novel strategy for the rapid generation of high-avidity antibody reagents. *J Mol Biol* **335**: 49–56.

Zhang L, Koay M, Maher MJ, Xiao Z & Wedd AG (2006) Intermolecular transfer of copper ions from the CopC protein of *Pseudomonas syringae*. Crystal structures of fully loaded Cu(I)Cu(II) forms. *J Am Chem Soc* **128**: 5834–5850.

## Supporting Information

Additional Supporting Information may be found in the online version of this article:

**Table S1.** Examples of *E. coli* and enterobacterial fimbrial adhesins with Ig domains.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.