ISSN 2229-5518

Bacterial Surface Display using Outer Membrane Proteins as Anchoring Sites

Sharadwata Pan and Michael K. Danquah

Abstract— Surface display of heterologous proteins or polypeptides on the surface of bacteria has gained momentum in recent years. Until recently, arrays of anchors or carriers have been identified for displaying diverse passenger proteins on the surface of *Escherichia coli*, majority of these involving the outer membrane proteins (OMPs). The reason for opting outer membrane proteins lies mainly in its ability to withstand the incorporation of large libraries of novel polypeptides, without a significant loss in steadiness. In this context, a thorough understanding of the underlying genetic mechanism of the OMP-mediated surface display is necessary. In this mini-review, we attempt to do the same and also compare the OMPs from two commonly used and available bacteria. The far reaching consequence of this lies in efficient surface display of imaginative and innovative polypeptides with applications ranging from bioremediation, immunology to vaccine development.

Index Terms— Anchoring site, bacteria, metalloprotein, outer membrane protein, surface display

1 Introduction

Today, it is accepted that for displaying heterologous proteins / polypeptides, the Gram-negative bacterium *Escherichia coli* (*E.coli*) is the preferred host mainly due to its ability to synthesize recombinant proteins in high amounts, and that it can be easily genetically manipulated ([1], [2]). A display system has to pass through the complex *E. coli* cell envelope to achieve surface display, which involves a periplasm that separates an inner membrane (IM) and an outer membrane (OM) [3]. By genetically combining the carrier and the passenger protein (protein of interest), it is comparatively easier to export the passenger protein across the cell envelope. This will affix the passenger to the surface of the bacterial cell, resulting in efficient surface display ([1], [2], [3]). For this, it is important that the carrier protein should have a stout surface anchor to fix the passenger to the cell surface , to ensure proper surface exposure and that the passenger can admit any externally added substrate. This has opened up possibilities to several biotechnological applications, ranging from protein library screening, vaccine development to the production of biofuels [4]. A detailed account on the applications of bacterial cell surface display is given in an earlier review [1].

Till date, different carrier proteins have been efficiently used for the surface display of passenger proteins on *E. coli* (for a very recent and comprehensive summary, see Table 1 in [3]). However, these systems have limitations in the form of use of peptides and small proteins. By contrast, large and intricate passengers may be effectively displayed using systems based on autotransporter proteins (ATs) and ice nucleation protein (INP) [3]. Also, it is noteworthy that diverse protein scaffolds have been used as carrier proteins [3], of which, the enormous majority is based on *E. coli* β-barrel Outer Membrane Proteins (OMPs) (i.e. LamB, FhuA), and the porins (OmpA, OmpC and OmpX) ([5], [6], [7], [8]). Integral membrane proteins (like β-Barrel proteins) span the outer membrane with antiparallel β-sheets, thus forming a barrel-shaped structure [9]. It has been reported that majority of the OMP-based display systems can endure the incorporation of petite peptides in surface-exposed loops, without a major loss of steadiness [3]. Recently, a new display system based on the OmpX- eCPX (the passanger peptides may attach to its N or C terminus, or both) has been reported [7]. Of late, several systems have been developed that endorse the exhibit of reasonably large passenger proteins [3]. For example, a shortened alternative of OmpC and a system based on the imaginary OMP Omp1 from *Zymomonas mobilis* support the display of C-terminally attached passengers of 50 kDa and 56 kDa, respectively ([6], [10]). Considering all these, it is important to compare the OMPs for a deeper understanding. In this mini-review, we look at and compare the OMPs in common bacteria like *E. coli* and *Salmonella typhimurium*, which enriches our understanding of the underlying genetic mechanisms behind efficient surface display.

2 Metalloproteins and Role of Histidine and Cysteine in Metal Binding

A metalloprotein, in Biochemistry, is a standard term for a protein that has a metal cofactor. The metal may be an isolated ion or may be coordinated with a nonprotein organic compound, such as the porphyrin found in hemoproteins [11]. In some cases, the metal is co-coordinated with a side chain of the protein and an inorganic nonmetallic ion. This kind of protein-metal-nonmetal structure is seen in iron-sulfur clusters. Histidine is one of the 20 most common natural

- *Sharadwata Pan (Corresponding author) is currently pursuing PhD in Chemical Engineering at Indian Institute of Technology Bombay, Powai, Mumbai-400076, India, PH-918879347206. E-mail: sharadwata@gmail.com*
- *Michael K. Danquah is currently an adjunct Faculty inChemical Engineering at Monash University, Melbourne, VIC 3800, Australia. E-mail: kobinadanquah@yahoo.com*

amino acids present in proteins. In the nutritional sense, in humans, histidine is considered an essential amino acid, but mostly only in children [12]. The imidazole side chains and the relatively neutral pKa of histidine mean that relatively small shifts in cellular pH will change its charge. For this reason, this amino acid side chain finds its way into considerable use as a co-ordinating ligand in metalloproteins.

Cysteine is a naturally occurring, sulfur-containing amino acid that is found in most proteins, although only in small quantities. Cysteine is unique amongst the twenty natural amino acids as it contains a thiol group. The cysteine thiol group is also a nucleophile and can undergo addition and substitution reactions. The thiol groups become much more reactive when they are ionised, and cysteine residues in proteins have pKa values close to neutrality, so are often in their reactive thiolate form in the cell. The thiol group also has a high affinity for heavy metals [13] and proteins containing cysteine will bind metals such as mercury, lead and cadmium tightly. A variety of metal ions are known to interact with a lot of proteins. Coordinating metal ions interact with amino acid residues that contain electron-donating atoms (S, O or N) such as aspartate, glutamate, cysteine or histidine. In a number of metalloproteins, histidine side chains are able to adopt favorable conformations to form metal binding sites [14]. Examination of 3D structures of metalloproteins has allowed the identification of structural features of metal-protein interactions that can be used to incorporate metal binding sites in recombinant proteins for various purposes.

3 Outer Membrane Proteins in *E. coli* cells including Omp C

Outer Membrane Protein C (OmpC) is one of the most abundant OMPs in *E. coli* cells (up to 105 molecules per cell may be present). This protein is one of the three classical porins of *E. coli* K-12 strains (the other two are OmpF and PhoE) and consists of 367 amino acids, including a signal peptide consisting of 21 amino acids. Three OmpC molecules form a pore structure on the outer membrane of an *E. coli* cell, which allows small hydrophilic molecules to pass through. One OmpC molecule consists of 16 transmembrane, antiparallel β-strands, which produce a β-barrel structure surrounding a large channel and are connected by seven internal loops and eight external loops [15]. In general, the amino acid sequences of the external loops are less conservative, and thus these loops may be relatively tolerant to insertion and deletion. Therefore, most of the scientists had used one of the external loops as the point of insertion for foreign peptides for cell surface display.

This porin is a trimeric protein integrated in the outer membrane of *E. coli* that forms relatively nonspecific pores which allows the passive diffusion of small hydrophilic molecules. It has also been reported that OmpC is structurally important for stabilizing the outer membrane. The ompC gene from *E. coli* has been cloned, the sequence of the 1,101 base pair coding-region has been determined and it has been demonstrated that its expression is controlled by the osmolarity of the culture medium. This protein, with a molecular weight of 37,083 Daltons, is synthesized in a high osmolarity medium such as LB or after salt addition (at least 0.15 M) to a low osmolarity medium [16]. The high copy level of OmpC (2 - 10^5 molecules per cell) suggests that this outer membrane protein could be an excellent candidate for peptide surface expression. In Salmonella typhi, it has been reported that the OmpC porin could be used to successfully display a rotavirus epitope. Many scientists had used the OmpC porin from *E. coli* as an alternative system to display heterologous peptides. They had introduced into OmpC a metal-binding epitope and studied its metal-binding capacity on the surface of *E. coli*.

The matrix proteins, OmpC and OmpF, are major outer membrane proteins of Escherichia coli. In addition to these porins, a new porin, PhoE protein, has been recently reported, which forms a passive diffusion pore across the outer membrane preferentially for organic and inorganic phosphate. These proteins also share a common property of remaining associated with the peptidoglycan layer after extraction of the outer membrane with sodium dodecyl sulfate (SDS) in the presence of Mg^{2+} and probably exist in the outer membrane as trimers. Although the genes for these porins are completely unlinked on the *E. coli* chromosome, these porin proteins share a few common features [17]: (i) they have similar amino acid compositions and molecular weights; (ii) they are immunologically cross-reactive; and (iii) they form passive diffusion pores of similar diameters (1.3 nm for OmpC protein, 1.4 nm for OmpF protein, and 1.2 nm for PhoE protein).

These strongly suggest that the genes for these proteins may have evolved from a common ancestral gene. The DNA sequences for both the ompF and the phoE genes have been recently determined. It is a common perception that comparison of the DNA sequences of these genes as well as the primary amino acid sequences will provide insight as to how they are related to each other, how they are evolved and how the expressions of these genes are regulated. It is now known that the ompC and ompF genes are regulated by another independent operon, ompB, consisting of the ompR and enuZ genes, in such a way that the ompC gene is preferentially expressed in cells grown in a media of high osmolarity, while the ompF gene is expressed in cells grown in a media of low osmolarity [17]. It is a known fact that the DNA sequence encompassing the entire ompB operon has been determined. In contrast to the osmoregulation of the ompC and ompF genes, the phoE gene is known to be controlled by phosphate concentrations in a culture medium via the phoB gene.

4 Comparison of Outer Membrane Porin Proteins

As an example, let's consider outer membrane proteins produced by two well known and well studied bacteria: *E. coli* and *Salmonella typhimurium*. Wild-type *E. coli* K-12 produces two porin species, the OmpC protein and the OmpF protein. The structural genes for these proteins are located at ompC and ompF, and the expression of both proteins is controlled by

one or more genes at the ompB locus. The single porin species (termed matrix protein) produced by *E. coli* B appears to be almost identical to the OmpF protein of *E. coli* K-12, and *E. coli* B has ompB and ompF loci which are functionally and genetically equivalent to those of *E. coli* K-12.

Wild-type *S. typhimurium* LT2 produces three porin species, the OmpC (36K) protein, the OmpF (35K) protein and the OmpD (34K) protein. The expression of OmpC and OmpF proteins by this species is also determined by a gene at the ompB locus and the map positions of *S. typhimurium* LT2 ompB, ompC, and ompF are approximately the same as those of *E. coli* K-12. The ompD locus maps at 34 min on the *S. typhimurium* LT2 map and no corresponding locus has been identified in *E. coli* K-12. In addition to the similarity of genetic loci noted above, the OmpC and OmpF proteins of the two species share other similarities. The OmpF proteins of both species are repressed by high salt concentrations. The production of the OmpC protein is enhanced by high salt concentrations in *S. typhimurium* LT2, as it is in *E. coli* K-12. Although the ompB, ompC, and ompF loci of these two species are equivalent both in terms of map location and function, only about half of the peptides of either the OmpC or OmpF proteins are similar or identical between these species. Thus, there has evidently been considerable evolution of these genes since they diverged from common ancestral genes.

The OmpC, OmpF, and Lc (NmpC) porin proteins of *E. coli* K-12 have been shown to be similar to the OmpC (36K), OmpF (35K) and OmpD (34K) porin proteins of Salmonella typhimurium LT2 in terms of function, regulation of expression, and, in the case of OmpC and OmpF proteins, equivalence of the genetic loci determining their production [18]. However, the corresponding pairs of proteins from these two species showed only limited similarity in peptide maps and no similarity in terms of migration on polyacrylamide gels.

OmpC protein consists of a total of 367 amino acid residues with a signal peptide (21 amino acids) at its NH_2-terminal end. The 5′end noncoding region including the promoter of the *ompC* gene is extremely [A-T]-rich and the codon usage in the *ompC* gene is unusual as are those in genes for other abundant outer membrane proteins. The DNA sequence shows that an open translational reading frame can be ex- tended to 1101 nucleotides from the translation initiation codon ATG (nucleotides 404-406) to the termination codon TAA (nucleotides 1505-1507). This is the only possible open reading frame which could sufficiently encode a protein the size of OmpC protein [11]. There are as many as 42 out of phase termination codons (14 TAA, 6 TAG and 22 TGA) scattered through the entire coding region. Open Reading Frame (ORF) of the *ompC* gene (total 1104 base pairs):

ATGAAAGTTAAAGTACTGTCCCTCCTGGTCCCAGCTCTG
CTGGTAGCAGGCGCAGCAAACGCTGCTGAAGTTTAC
AACAAAGACGGCAACAAATTAGATCTGTACGGTAA
AGTAGACGGCCTGCACTATTTCTCTGACAACAAAGA
TGTAGATGGCGACCAGACCTACATGCGTCTTGGCTT
CAAAGGTGAAACTCAGGTTACTGACCAGCTGACCG
GTTACGGCCAGTGGGAATATCAGATCCAGGGCAAC
AGCGCTGAAAACGAAAACAACTCCTGGACCCGTGT
GGCATTCGCAGGTCTGAAATTCCAGGATGTGGGTTC
TTTCGACTACGGTCGTAACTACGGCGTTGTTTATGAC
GTAACTTCCTGGACCGACGTACTGCCAGAATTCGGT
GGTGACACCTACGGTTCTGACAACTTCATGCAGCAG
CGTGGTAACGGCTTCGCGACCTACCGTAACACTGAC
TTCTTCGGTCTGGTTGACGGCCTGAACTTTGCTGTTC
AGTACCAGGGTAAAAACGGCAACCCATCTGGTGAA
GGCTTTACTAGTGGCGTAACTAACAACGGTCGTGAC
GCACTGCGTCAAAACGGCGACGGCGTCGGCGGTTCT
ATCACTTATGATTACGAAGGTTTCCGTATCGGTGGTG
CGATCTCCAGCTCCAAACGTACTGATGCTCAGAACA
CCGCTGCTTACATCGGTAACGGCGACCGTGCTGAAA
CCTACACTGGTGGTCTGAAATACGACGCTAACAACA
TCTACCTGGCTGCTCAGTACACCCAGACCTACAACG
CAACTCGCGTAGGTTCCCTGGGTTGGGCGAACAAAG
CACAGAACTTCGAAGCTGTTGCTCAGTACCAGTTCG
ACTTCGGTCTGCGTCCGTCCCTGGCTTACCTGCAGTC
TAAAGGTAAAAACCTGGGTCGTGGCTACGACGACG
AAGATATCCTGAAATATGTTGATGTTGGTGCTACCT
ACTACTTCAACAAAAACATGTCCACCTACGTTGACT
ACAAAATCAACCTGCTGGACGACAACCAGTTCACTC
GTGACGCTGGCATCAACACTGATAACATCGTAGCTC
TGGGTCTGGTTTACCAGTTC**TAA**

Amino acid sequence of the pro-OmpC protein (as deduced by translation using Expasy Proteomics Server- Reverse Translation tool) [19]:

MKVKVLSLLVPALLVAGAANAAEVYNKDGNKLDLYG
KVDGLHYFSDNKDVDGDQTYMRLGFKGETQVTDQLT
GYGQWEYQIQGNSAENENNSWTRVAFAGLKFQDVGS
FDYGRNYGVVYDVTSWTDVLPEFGGDTYGSDNFMQQ
RGNGFATYRNTDFFGLVDGLNFAVQYQGKNGNPSGE
GFTSGVTNNGRDALRQNGDGVGGSITYDYEGFRIGGAI
SSSKRTDAQNTAAYIGNGDRAETYTGGLKYDANNIYL
AAQYTQTYNATRVGSLGWANKAQNFEAVAQYQFDFG
LRPSLAYLQSKGKNLGRGYDDEDILKYVDVGATYYFNK
NMSTYVDYKINLLDDNQFTRDAGINTDNIVALGLVYQF

5 Conclusions

In this mini-review, we have analyzed and compare the OMPs of two common bacteria which lead us to a clearer understanding of the mechanisms and specific stuructural details involved in the surface display of target passenger proteins on their surface, attached to the carriers. With efficient knowledge of protein translocation pathways, assembly of β-barrel OMPs and the current structural understanding (for example, the details of the ORF and amino acid sequence of *omp*C gene), it may be feasible to redesign and optimize protein assembly machineries. The far reaching consequence of this lies in fine-tuning the amount of the surface-displayed passenger protein.

REFERENCES

[1] S. Pan and M.K. Danquah, "Cell Surface Display," *Biodiversity,* A. Sofo, ed., Croatia: InTech, pp. 63-74, 2011.

[2] S.Y. Lee, J.H. Choi and Z. Xu, "Microbial Cell-Surface Display," *Trends Biotechnol*, vol. 21, no. 1, pp. 45–52, 2003.

[3] E. van Bloois, R.T. Winter, H. Kolmar and M.W. Fraaije, "Decorating Microbes: Surface Display of Proteins on *Escherichia coli*," *Trends Biotechnol*, vol. 29, no. 2, pp. 79-86, 2011.

[4] C.H. Wu, A. Mulchandani and W. Chen, "Versatile Microbial Surface Display for Environtal Remediation and Biofuel Production," *Trends Microbiol,* vol. 16, no. 4, pp. 181-188, 2008.

[5] G.S. Verhoeven, S. Alexeeva, M. Dogterom and T. den Blaauwen, "Differential Bacterial Surface Display of Peptides by The Transmembrane Domain of OmpA.," *PLoS One*, vol. 4, no. 8, e6739, 2009.

[6] J.H. Baek, M.-J. Han, S.H. Lee and S.Y. Lee, "Enhanced Display of Lipase on The *Escherichia coli* Cell Surface, Based On Transcriptome Analysis," *Appl. Environ. Microbiol*, vol. 76, no. 3, pp. 971–973, 2009.

[7] J.J. Rice and P.S. Daugherty, "Directed Evolution of a Biterminal Bacterial Display Scaffold Enhances The Display of Diverse Peptides," *Protein Eng. Des. Sel*, vol. 21, no. 7, pp. 435–442, 2008.

[8] P. Samuelson, E. Gunneriusson, P.-Å. Nygren and S. Ståhl, "Display of proteins on bacteria," *J. Biotechnol*, vol. 96, no. 2, pp. 129–154, 2002.

[9] G.E. Schulz, "The Structure of Bacterial Outer Membrane Proteins," *Biochim. Biophys. Acta*, vol. 1565, no. 2, pp. 308–317, 2002.

[10] M.-X. He, H. Feng and Y.-Z. Zhang, "Construction of a Novel Cell-Surface Display System for Heterologous Gene Expression in *Escherichia coli* by Using an Outer Membrane Protein of *Zymomonas mobilis* as Anchor Motif," *Biotechnol. Lett*, vol. 30, no. 12, pp. 2111–2117, 2008.

[11] D.F. Shriver and P.W. Atkins, *Inorganic Chemistry* (Chapter 19, Bioinorganic chemistry), Third edition, UK: Oxford University Press, 1999.

[12] J.D. Kopple and M.E. Swendseid, "Evidence That Histididne is an Essential Amino Acid in Normal and Chronically Uremic Man," *J. Clin. Invest,* vol. 55, no. 5, pp. 881-891, 1975.

[13] D.H. Baker and G.L. Czarnecki-Maulden, "Pharmacologic Role of Cysteine in Amelietoring or Exacerbating Mineral Toxicities," *J. Nutr*, vol. 117, no. 6, pp. 1003-1010, 1987.

[14] W.S. Willett, S.A. Gillmor, J.J. Perona, R.J. Fletterick and C.S. Craik, "Engineered Metal Regulation of Trypsin Specificity," *Biochemistry*, vol. 34, no. 7, pp. 2172–2180, 1995.

[15] Z. Xu, and S.Y. Lee, "Display Of Polyhistidine Peptides On The *Escherichia coli* Cell Surface By Using Outer Membrane Protein C as an Anchoring Motif," *Appl. Environ. Microb*, vol. 65, no. 11, pp. 5142–5147, 1999.

[16] N. Cruz, S.L. Borgne, G.H. Chavez, G. Gosset, F. Valle and F. Bolivar, "Engineering The *Escherichia coli* Outer Membrane Protein OmpC for Metal Bioadsorption," *Biotechnol. Lett*, vol. 22, no. 7, pp. 623-629, 2000.

[17] T. Mizunol, M.Y. Chou and M. Inouye, "A Comparative Study on the Genes for Three Porins of the *Escherichia coli* Outer Membrane," *J. Biol. Chem*, vol. 258, no. 11, pp. 6932-6940, 1983.

[18] D.R. Lee and C.A. Schnaitman, "Comparison of Outer Membrane Porin Proteins Produced by *Escherichia coli* and *Salmonella typhimurium*," *J. Bacteriol*, vol. 142, no. 3, pp. 1019-1022, 1980.

[19] E. Gasteiger, A. Gattiker, C. Hoogland, I. Ivanyi, R.D. Appel and A. Bairoch, "ExPASy: The Proteomics Server for In-depth Protein Knowledge and Analysis," *Nucleic Acids Res*, vol. 31, no. 13, pp. 3784-3788, 2003.