Detection of soluble expression and *in vivo* interactions of the inner membrane protein OppC using green fluorescent protein

Q.J. Xiang[2,3], J.F. Zhai[1], M. Zhang[1] and B. Zhang[1]

[1]Department of Gynecology, Central Hospital of Xuzhou,
Affiliated Hospital of Southeast University, Xuzhou, China
[2]Department of Microbiology, College of Resource and Environmental Sciences,
Sichuan Agricultural University, Chengdu, China
[3]College of Life Sciences,
Sichuan Key Laboratory of Molecular Biology and Biotechnology,
Sichuan University, Chengdu, China

Corresponding authors: M. Zhang / B. Zhang
E-mail: aimeng860107@126.com / bettyzhang10@163.com

ABSTRACT. In this study, the *in vivo* interaction system of oligopeptide permease (Opp) proteins was analyzed, and a high expression system of inner membrane protein OppC was constructed by flexible usage of the green fluorescent protein (GFP). The *Escherichia coli OppC* gene, which encodes a transmembrane component of oligopeptide transporter, was cloned into different vectors. Recombinant plasmids were transformed into different *E. coli* strains, and the expression conditions were optimized. The effect of plasmids and expression strains on OppC production was evaluated by in-gel and western blot analyses. OppC produced by the pWaldo-GFPe vector, harboring the GFP reporter gene, transformed into *E. coli* C43(DE3) provided sufficient functional protein for biochemical and biophysical studies. *In vivo* protein-protein interactions were detected among oligopeptide permease proteins using a GFP fragment reassembly

protocol. The substrate binding protein OppA showed no interaction with the other components, while the ATP-binding component OppD did not interact with OppF. OppD and OppF interacted with the transmembrane components OppB and OppC. OppB also showed direct interaction with OppC. *In vivo* OppC functionality was determined by constructing an *OppC* gene deletion strain. OppC was shown to be essential for peptide uptake, and non-essential for cell viability. These results could help in elucidating the oligopeptide transport mechanism in bacteria.

Key words: Oligopeptide permease; Protein-protein interaction; Inner membrane protein; Green fluorescent protein

INTRODUCTION

A number of bacterial species cannot synthesize amino acids, fatty acids, and certain other essential elements (Matthews, 1975; Lin et al., 2001), and uptake essential nutrients from their environment. The ABC transporters mediate a large fraction of the nutrient uptake capacity in bacteria. Some species typically possess multiple peptide transporters that specifically facilitate the utilization of different peptides as amino acid and nitrogen sources (Payne and Smith, 1994; Juillard et al., 1995). The oligopeptide permease (Opp), or the oligopeptide uptake system, represents the major peptide uptake method adopted by archaea and bacteria.

Besides transporting oligopeptides, the Opp systems in several pathogenic bacteria (Wang et al., 2004; Moutran et al., 2004; Flores-Valdez et al., 2009; Moraes et al., 2014) participate in distinct life processes, including the binding of soluble host proteins (Cundell et al., 1995; Henrich et al., 1999; Fenno et al., 2000), cellular signal transudation (Ruhfel et al., 1993), recycling of cell-wall peptides released from the growing peptidoglycan (Goodell and Higgins, 1987), and initiation of sporulation (Perego et al., 1991). The Opp system plays a leading role in imparting cell sensitivity to aminoglycoside antibiotics (Kashiwagi et al., 1992), and imparts virulence and pathogenicity to the organisms (Moraes et al., 2014). The blocking of oligopeptide transport in pathogenic bacteria usually resulted in a reduction in virulence (Reniere et al., 2015). Therefore, oligopeptide transporters are believed to be effective targets in some multi-drug-resistant bacteria.

Opp permease is a multicomponent system that is usually composed of five subunits: two integral membrane subunits (OppB and OppC) forming the transmembrane pore and two cytoplasmic ATP-binding subunits (OppD and OppF) that can hydrolyze ATP, or work in conjunction with one periplasmic (or membrane-bound lipoprotein in Gram-positive bacteria) peptide-binding subunit (OppA), which confers high affinity (but low specificity) to oligopeptide uptake systems.

The mechanism of the Opp system might vary from that of the mammalian oligopeptide-proton symporters PepT1 and PepT2 (Newstead et al., 2010), the dipeptide permease (Dpp) system (Letoffe et al., 2006) and specific nutrient transporters (Suzuki et al., 2005) in bacteria. Further studies on Opp transporters would help clarify the transport mechanism, as well as the evolutionary pathway from the dipeptide transporter to oligopeptide transporter, in bacteria. Understanding the mechanism may also provide deeper insights into the oligopeptide transportation and drug-resistance mechanisms.

One major limitation to the study of membrane proteins is the non-availability of pure and sufficient functional protein. Over-expression and purification of OppC will be a challenge. In this study, the expression vectors and strains, and optimized expression conditions of *OppC* extraction from *Escherichia coli k-12* were screened. Green fluorescent protein (GFP), a fusion protein, was

used to facilitate the soluble expression of OppC, and to optimize the expression and purification protocols. Sufficient purified protein was obtained for solid-state nuclear magnetic resonance and other biophysical studies, such as crystallization for protein structure research.

Peptides have been previously utilized as nitrogen sources (Matthews, 1975; Juillard et al., 1995). *In vivo* function of OppC was tested using M9 medium (which was used to cultivate the *OppC* deletion strain), using a peptide mixture as the only nitrogen source. Through this, OppC was proven to be essential for peptide uptake in *E. coli*.

The transport mechanism was analyzed by evaluating the *in vivo* protein-protein interactions among oligopeptide permease proteins using a previously described GFP fragment reassembly protocol (Wilson et al., 2004; Magliery et al., 2005). OppB and OppC were found to be located at the center of this interaction map. Studies on the function and interaction of OppC would help gain deeper insights into the peptide transport mechanisms, in addition to assisting in the development of novel drugs against multi-drug-resistant pathogenic bacteria (Moraes et al., 2014).

MATERIAL AND METHODS

Strains and plasmids

Genomic DNA of *E. coli k-12* was used as the template for cloning *OppC*. *E. coli BW25113* was used to knock out *OppC*; *E. coli* JM109 was used for gene cloning, and the BL21(DE3), Rosetta (DE3), C41(DE3), C43(DE3) strains were used for protein expression. pACYC-duet1, pET28, and pET28-derived pWaldo-GFPe were employed as gene cloning and expression vectors. pET11a-link-NGFP and pMRBAD-link-CGFP, kindly gifted by Professor Lynne Regan of Yale University, were used for protein interaction analyses. *KOD-Fx* polymerase and restriction enzymes were purchased from ToYoBo (Shanghai, China) and Fermentas (Shenzhen, China).

Construction of expression and interaction plasmids

The OppC encoding sequence was amplified by standard PCR, using the *KOD-Fx* DNA polymerase. The expression primer sequences are depicted in **Table S1**. The PCR products were gel-purified, digested with restriction enzymes, and ligated into the pre-digested expression pACYC-duet1, pET28, or pWaldo-GFPe vectors. Coding sequences of OppA, OppB, OppC, OppD, and OppF were amplified by interaction primers (**Table S1**). PCR products were gel-purified, digested with restriction enzymes, and subsequently ligated into the pre-digested interaction plasmids pET11a-link-NGFP and pMRBAD-link-CGFP. The ligation mix was transformed into *E. coli* JM109 competent cells. The sequences of each transformant were confirmed by sequencing in a BGI analyzer (Shenzhen, China), and analyzed using a Basic Local Alignment Search Tool (BLAST) algorithm on the National Center for Biotechnology Information (NCBI) website. Multiple-sequence alignment of proteins was performed using the Clustal X2 program (Thompson et al., 1994). The sequence properties were analyzed using the DNAMAN 6.0 program.

Recombinant expression of OppC

Optmization of OppC-GFP expression

pACYC-OppC, pET28-OppC, and pWaldo-OppC were transformed and expressed in

BL21(DE3), Rosetta(DE3), C41(DE3), and C43(DE3) cells. The cells of fresh pET28-OppC and pWaldo-OppC transformants were cultured in 2 mL Luria-Bertani (LB) medium supplemented with 50 µg/mL kanamycin [BL21(DE3), C41(DE3), C43(DE3)], or 50 µg/mL kanamycin and 34 µg/mL chloramphenicol [Rosetta(DE3)], and cultured overnight at 37°C and 200 rpm. The pACYC-OppC transformants were cultured overnight in medium containing chloramphenicol (34 µg/mL) at 37°C and 200 rpm. The overnight culture was transferred to 10 mL fresh LB medium and incubated at 37°C until an OD_{600} of 0.5-0.6 was reached. The OppC expression was subsequently induced with isopropyl-thio-β-D-thiogalactopyranoside (IPTG) at 18°, 25°, and 30°C. The cells were harvested 20 h after induction by centrifugation, and washed with Tris-buffered saline (TBS, pH 8.0). The pellet could be used immediately for following experiments, or stored at -20°C for at least 1 week.

In-gel analysis of OppC-GFP expression

The cells were treated for sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). The recombinant strains were induced with 0.1 mM IPTG and cultured at 18°C and 200 rpm for 20 h, and subsequently harvested by centrifugation. The cells were resuspended, mixed with solubilization buffer (0.02% bromophenol blue, 20% glycerol, 5 mM EDTA, pH 8.0, 200 mM Tris-HCl, pH 8.8, 4% SDS, and 0.05 M DTT), and incubated at 37°C for 5 min. The samples were centrifuged at 13,000 g at 4°C for 15 min, and the resultant suspension analyzed by SDS-PAGE. The gel was exposed to ultraviolet light, and in-gel fluorescent images were captured using a CCD camera system (Bio-Rad, Shanghai, China). The molecular weight of the fluorescent band was confirmed by staining the gel with Coomassie brilliant blue R250. The cells were also disrupted with 5X SDS loading buffer (10% SDS, 0.5% bromophenol blue, 50% glycerol, 250 mM Tris-HCl, pH 6.8, and 5% β-mercaptoethanol). The protein was thoroughly denatured by incubation at 95°C for 5 min, and analyzed by SDS-PAGE. Following electrophoresis, the gel was stained with Coomassie brilliant blue R250, in order to visualize the protein bands.

Over-expression and purification of OppC-GFP

Overnight culture of C43(DE3) containing pWaldo-OppC (50 mL) was inoculated into 1 L LB media. At an OD_{600} of 0.5-0.6, 0.1 mM IPTG was added to induce protein expression. The temperature was reduced from 37° to 18°C, and incubation continued for 20 h at 200 rpm. The cells were harvested by centrifugation, and disrupted in TBS, pH 8.0, containing 1 µM DNase I (Sigma-Aldrich), 1 µM lysozyme (Sigma-Aldrich), and 1 mM phenylmethylsulfonyl fluoride (Sigma-Aldrich) by French press. The cell debris was removed by centrifugation at 5,000 g for 15 min at 4°C.

The supernatant was then ultracentrifuged at 100,000 g for 1 h to pellet the whole membrane, which was subsequently solubilized in 20 mM Tris-HCl, pH 7.8, containing 10 mM imidazole, 100 mM NaCl, and 1% (w/v) DDM (n-dodecyl-β-D-maltoside) for 2 h at room temperature. The solubilized membrane was ultracentrifuged at 100,000 g for 1 h to remove membrane debris prior to being loaded onto a nickel-nitrilotriacetate affinity resin (Ni-NTA; Qiagen). The protein bound to Ni^{2+} was washed with 30 mM imidazole, 20 mM Tris-HCl, pH 7.8, 300 mM NaCl, and 0.016% DDM, and eluted with 300 mM imidazole, 20 mM Tris-HCl, pH 7.8, 300 mM NaCl, and 0.016% DDM. The eluted protein was desalted using a HiPrep 26/10 Desalting column (GE Healthcare), and digested overnight with Tobacco Etch Virus (TEV) protease at room temperature. The aggregated protein

was removed by centrifugation at 5000 *g* for 15 min. The supernatant was loaded onto an Ni-NTA to remove the GFP-His8 fragment. The effluent containing purified protein was pooled and further purified by size-exclusion chromatography, using a HiLoad 16/60 Superdex 200 prep grade column (GE Healthcare). The purified protein was subsequently pooled and concentrated. The desalt and gel-filtration buffer consisted of 20 mM Tris-HCl, pH 7.8, 100 mM NaCl, and 0.016% DDM. The protein purity was monitored on a 12% SDS-PAGE. The protein concentration was determined using the Bradford method, using bovine serum albumin as the standard.

Western blot analysis

The protein expression was confirmed by western blot analysis. OppC-GFP fusion protein was resolved on a 12% SDS-PAGE which was electro-blotted onto a nitrocellulose membrane using an iBlot™ Dry Blotting System (Invitrogen). The membrane was blocked with PBS, pH 7.4, and 0.1% Tween 20 (PBST) containing 5% (w/v) non-fat dry milk for 1 h at room temperature. The blocked membrane was then washed thrice with PBST and incubated with anti-His monoclonal antibody (anti-6His, 400 µg/mL; 1:1000 dilution) for 1 h at room temperature. The membrane was washed again and incubated with a horseradish peroxidase-labeled antibody (goat anti-mouse, 0.8 mg/mL; 1:5000 dilution) for 1 h at room temperature. The blot signal was revealed by enhanced chemiluminescence, and the protein bands visualized by autoradiography. The OppC expression was also confirmed by mass spectrometry.

Functional assays

OppC del-F and OppC del-R primers (**Table S1**) were used to generate the *OppC* knockout *E. coli* strain, using the Red system. *E. coli BW25113* cells, containing the Red helper plasmid was grown in 5 mL SOB medium supplemented with ampicillin and L-arabinose at 30°C, until an OD_{600} of 0.5-0.6 was reached. Competent cells were obtained by washing with ice-cold 10% glycerol. The kanamycin fragment in pKD13 was cloned using OppC del-F and OppC del-R by standard PCR. The PCR product (upstream 55 bp and downstream 54 bp), homologous to the adjacent regions of *OppC*, was digested with *Dpn*I and gel-purified. The cells were subjected to electroporation using a MicroPulser (Bio-Rad) with a 0.2-cm chamber according to the manufacturer protocols, using 80 ng PCR product and 100 µL competent cells. The *OppC* deletion was confirmed by PCR, using OppC expression F1 and R1 primers (**Table S1**).

To determine if *OppC* was essential for oligopeptide transport, we tested the ability of *OppC* deleted *E. coli* strain in acquiring nitrogen through peptide uptake. A mixture of two to five amino acid peptides (Lys-Leu, Lys-Leu-Gly, Glu-Cys-Gly, Lys-Leu-Leu-Gly, Lys-Leu-Gly-Leu, and Lys-Leu-Leu-Leu-Gly) served as the only nitrogen source in M9 medium. In order to determine whether the *OppC* could compensate this defection, plasmid pACYC-OppC was transformed into the *OppC* deleted strain.

OppC interacts with other components of the oligopeptide permease system

The interaction between OppC and the other four components of the oligopeptide permease system was detected using a GFP fragment reassembly protocol. The coding sequences of OppA, OppB, OppC, OppD, and OppF were cloned into pET11a-link-NGFP and pMRBAD-link-CGFP.

Both plasmids were simultaneously transformed into BL21(DE3), using 1-10 ng of each construct. The doubly transformed cells were incubated at 37°C and 200 rpm for 1 h, and subsequently plated onto selective LB agar plates containing 35 µg/mL kanamycin and 100 µg/mL ampicillin. Single colonies were selected and incubated overnight in LB liquid medium with shaking at 37°C. A 1:10 dilution of fresh overnight culture (5-10 µL) was plated onto screening media containing 10 µM IPTG and 0.2% arabinose, in order to screen for GFP reassembly.

As GFP reassembly is effectively irreversible, in-gel analysis was also performed to verify the interactions. Fresh overnight culture (500 µL) was inoculated into 20 mL fresh LB medium, and incubated at 37°C, 200 rpm until an OD_{600} of 0.5-0.6 was attained. The temperature was shifted to 18°C, and 10 µM IPTG and 0.2% arabinose were added to induce the expression of genes in pET11a-link-NGFP and pMRBAD-link-CGFP, respectively. The cells were harvested after 24 h, and treated with SB for in-gel SDS-PAGE analysis.

RESULTS

Protein interaction and OppC expression plasmid construction

The coding sequence of OppA, OppB, OppC, OppD, and OppF was amplified using *E. coli k-12* genomic DNA as the template. OppA, OppB, OppC, and OppD were cloned into pET11a-link-NGFP vectors, while OppB, OppC, and OppF were inserted into a pMRBAD-link-CGFP. The resulting plasmids were called pN-OppA, pN-OppB, pN-OppC, and pN-OppD, and pC-OppB, pC-OppC, and pC-OppF, respectively. The constructions were confirmed by DNA sequencing.

The coding sequence of OppC was also cloned into different expression vectors, which were named pACYC-OppC, pET28-OppC, and pWaldo-OppC. The insertion was confirmed by DNA sequencing.

Optimizing strains for OppC expression

The optimal expression conditions for pWaldo-OppC were determined by monitoring bacterial growth and protein fluorescence under varying temperature and IPTG concentrations. The OppC-GFP fusion protein was analyzed by SDS-PAGE and western blotting, using an antibody against the His-tag.

The OppC-GFP expression and solubility levels were significantly different among the four strains when induced with 0.1 mM IPTG at 18°C for 20 h. The yield of OppC-GFP was much higher in the BL21(DE3) and Rosetta(DE3) strains than in C43(DE3) and C41(DE3) (Figure 1B). However, visual inspection of the strains and the in-gel GFP bands showed better expression of OppC-GFP in C43(DE3) and C41(DE3) than in the BL21(DE3) and Rosetta(DE3) strains (Figure 1A). Although we observed a higher yield of OppC-GFP in the BL21(DE3) and Rosetta(DE3) strains, the protein could not fold properly. Well-folded OppC-GFP was fluorescent, which could be enhanced by UV light. Therefore, C43(DE3) was selected to over-express and purify the OppC-GFP protein. The expression of OppC-GFP in C43(DE3) was not affected by IPTG concentration (varying from 0.1 to 1.0 mM (Figure 1C and D); however, increasing the incubation temperature from 18° to 30°C slightly decreased the expression level and folding of OppC-GFP.

Finally, the optimized expression conditions were determined to be as follows: C43(DE3) containing pWaldo-OppC, induced with 0.1 mM IPTG, and cultured at 18°C for 20 h prior to harvesting.
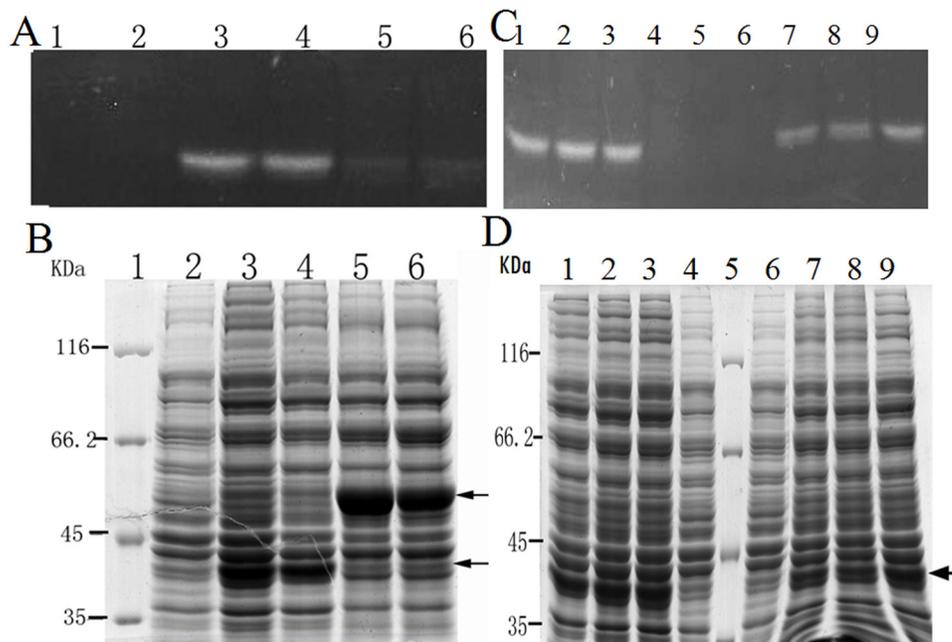
Figure 1. In-gel fluorescence analysis of OppC-GFP expression. **A.** and **B.** OppC-GFP expression in diverse strains. *Lane 1* = marker; *lane 2* = C43(DE3) strain without IPTG induction; *lanes 3-6* = C43(DE3), C41(DE3), BL21(DE3), and Rosetta(DE3) strains grown at 18°C and induced with 0.1 mM IPTG. **C.** and **D.** OppC-GFP expression in C43(DE3) under different culture conditions. *Lane 5* = marker; *lanes 1-3* grown at 18°C and induced with 0.1, 0.5, and 1.0 mM IPTG; *lanes 4 and 6* = not induced with IPTG; *lanes 7-9* = cells induced with 0.1 mM IPTG and grown at 30°, 25°, and 18°C. **A.** and **C.** depict in-gel fluorescence results. **B.** and **D.** depict the same gel stained with Coomassie brilliant blue. The arrow indicates OppC-GFP fluorescent fusion.

In-gel mobility shift of the OppC-GFP fusion protein

The OppC-GFP fusion protein obtained from different expression strains was analyzed by SDS-PAGE and in-gel detection of GFP fluorescence. The protein sample was not denatured for in-gel analysis. Each sample displayed a single prominent fluorescent band. The gel was also analyzed by western blot using an anti-His6 antibody. Surprisingly, two obvious bands were observed. One band was located at the same position as the fluorescent band, while the other band was approximately 10 kDa higher, which was consistent with the expected molecular weight of OppC-GFP (Figure 2A and B). The intensity ratios of the two bands were different. The anomalous migration of fluorescent bands might be a result of preservation of the GFP moiety structure.

The origin of the dual migration was analyzed by preparing protein samples by denaturing and non-denaturing methods. The non-fluorescent band with the correct molecular mass was found to be denatured OppC-GFP. The OppC-GFP protein represented by the GFP band was in the folded form. The intensity of the non-denatured bands was closely related to the protein expression level, and was not affected by the sample preparation process (Wang et al., 2011). The fluorescent bands were proposed to represent the soluble portion of the target protein (Geertsma et al., 2008).
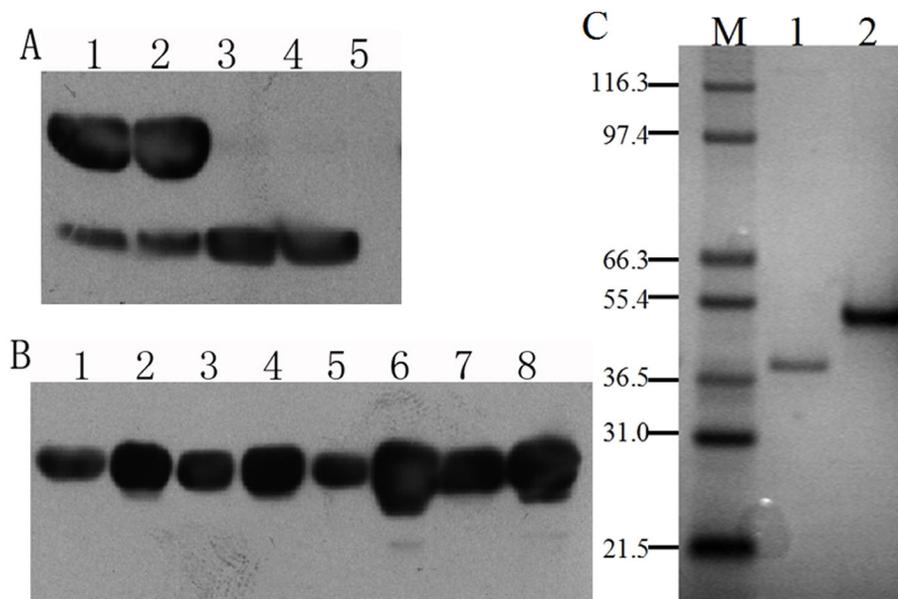
Figure 2. Western blot confirmation and purification of OppC. **A.** Western blot analysis of OppC-GFP expression in different host strains. Samples were mixed with solubilization buffer and incubated at 37°C for 5 min. Lower and upper bands indicate the positions of fluorescent and non-fluorescent species of OppC-GFP, respectively. *Lane 1* = Rosetta(DE3); *lane 2* = BL21(DE3); *lane 3* = C41(DE3); *lane 4* = C43(DE3); *lane 5* = uninduced. **B.** Western blot analysis of OppC-GFP expression in different strains. The cells were disrupted by sonication, and denatured by boiling for 5 min. *Lanes 1* and *2* = C43(DE3) soluble fraction and insoluble fraction; *lanes 3* and *4* = C41(DE3) soluble and insoluble fractions; *lanes 5* and *6* = BL21(DE3) soluble and insoluble fractions; *lanes 7* and *8* = Rosetta(DE3) soluble and insoluble fraction. **C.** SDS-PAGE analysis of purified OppC. M = protein marker; *lane 1* = purified OppC protein; *lane 2* = eluted GFP moiety.

OppC-GFP produced by the BL21(DE3) and Rosetta(DE3) strains showed two bands. The intensity of denatured bands was much higher than that of the lower bands. C41(DE3) and C43(DE3) strains produced proteins that expressed only one obvious fluorescent band (Figure 2A). According to the above observation, OppC-GFP must be soluble in C41(DE3) and C43(DE3) and insoluble (almost) in BL21(DE3) and Rosetta(DE3). This conjecture was tested by western blot analysis of the solubility of OppC-GFP produced by the four bacterial strains. A greater portion of the protein expressed by C41(DE3) and C43(DE3) was found to be insoluble (Figure 2B). The intensities of the soluble fractions in BL21(DE3) and Rosetta(DE3) were found to be similar as in C41(DE3) and C43(DE3). However the fluorescence of the cell and in-gel protein bands was found to be obscure. This showed that BL21(DE3) and Rosetta(DE3) might be able to express some OppC-GFP as soluble, but improperly folded fusion proteins.

Purification of the OppC-GFP protein

C43(DE3) was used to over-express OppC-GFP as this strain could promote the solubility of the protein. Detergent DDM was used to extract and solubilize OppC-GFP from the membrane. Ni^{2+}-NTA column enriched the protein from the cell lysate. TEV digested OppC-GFP overnight and could separate OppC and GFP. Subsequently, the GFP moiety containing the 8-His tag

was removed by Ni^{2+}-NTA. The OppC was further purified by gel filtration, in order to obtain a homogeneous protein. The purified OppC protein was found to be quite pure (Figure 2C).

The protein could be traced during the purification process using GFP as the reporter. Fractions containing OppC were pooled and concentrated to approximately 10 mg/mL. Approximately 0.5 mg OppC protein was purified from 1 L of the bacterial culture. This volume and concentration could support protein function and crystallography research.

OppC is essential for peptide transport

OppC was replaced by a kanamycin-resistant gene, and the resultant *E. coli* strain named Δ*OppC*. *OppC* gene deletion was confirmed by PCR (data not shown). In order to test if OppC is essential for peptide transport in *E. coli*, M9 medium was used to cultivate the Δ*OppC* strain, and a mixture of 2-5 peptides was used as the only nitrogen source. The growth conditions for wild-type *E. coli,* using peptides as the sole nitrogen source, were applied as the control. Wild-type *E. coli* growth in peptide-containing M9 medium was not affected compared to cultivation in LB broth (peptide-rich medium), which showed that the peptide mixture could work as the nitrogen source.

The *OppC* gene in pACYC-OppC could compensate for the Δ*OppC* defection. The wild-type *E. coli* and pACYC-OppC-transformed Δ*OppC* showed normal growth. The *OppC*-depleted strain containing the pACYC-duet1 empty vector showed normal, but depleted growth (Figure 3A).
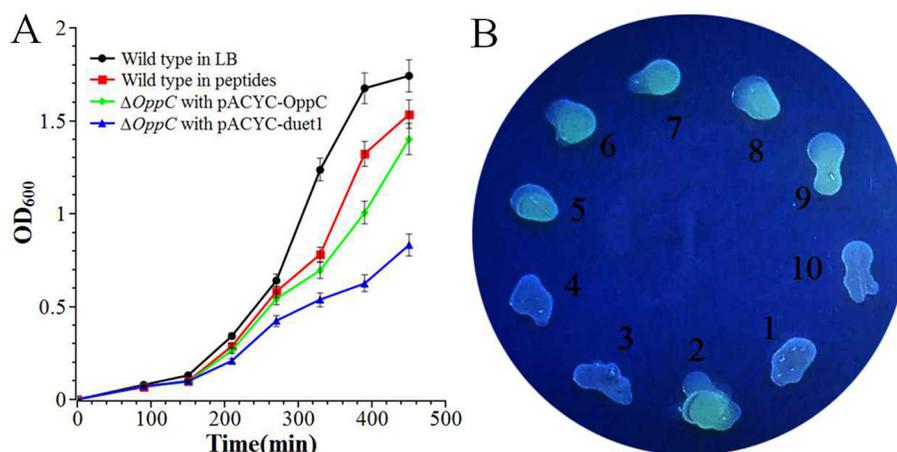

Figure 3. Functional assay of OppC *in vivo*. **A.** Cell growth of the Δ*OppC* strain under different cultivation conditions. Error bars represent means ± SD (N = 3). **B.** Protein-protein interactions among oligopeptide permease proteins. The BL21(DE3) strain, transformed with interaction plasmids, was induced with 10 μM IPTG and 0.2% arabinose. The fluorescent property of GFP was activated by UV light. 1-10: pET11a-link-NGFP and pMRBAD-link-CGFP; pN-Z and pC-Z; pN-OppA and pC-OppB; pN-OppA and pC-OppC; pN-OppB and pC-OppC; pN-OppB and pC-OppF; pN-OppD and pC-OppB; pN-OppD and pC-OppC; pN-OppC and pC-OppF; and pN-OppD and pC-OppF.

OppC interacts with other oligopeptide permease proteins

pN-OppA and pC-OppB, pN-OppA and pC-OppC, pN-OppB and pC-OppC, pN-OppB and pC-OppF, pN-OppC and pC-OppF, pN-OppD and pC-OppB, pN-OppD and pC-OppC, and

pN-OppD and pC-OppF were simultaneously transformed into BL21(DE3). The protein-protein interactions were analyzed using screening media containing 10 µM IPTG and 0.2% arabinose, in order to screen for GFP reassembly. The assembled GFP displayed in-cell fluorescence. This showed that the substrate binding protein OppA did not interact with the transmembrane components OppB and OppC. On the other hand, the ATP-binding component OppD had no direct interaction with OppF. OppB and OppC could interact with each other and also with OppD and OppF (Figure 3B). In-gel analysis of GFP reassembly confirmed the *in vivo* results (**Figure S1**).

DISCUSSION

Many powerful vectors could be used to express foreign genes in *E. coli*, and the proteins could be easily purified by tags. pET28 and pACYC-duet1 were used to express the inner membrane protein OppC. The yield was not very good (data not shown); however, the yield of OppC in pWaldo-GFPe was much higher. pWaldo-GFPe used GFP as the reporter, which was convenient for the optimization of the expression and purification protocols. GFP allowed for real-time monitoring of the target protein during the expression and purification period.

As the GFP gene was located at the C-terminal of the *OppC* gene, the fluorescence could be a reporter of the expressed full-length OppC. The GFP moiety was exceptionally stable, and the fluorescence could be detected even in SDS-PAGE with a detection limit of less than 5 ng GFP per protein band (Drew et al., 2006). Therefore, pWaldo-GFPe was selected for OppC expression.

High-level OppC-GFP expression strains with greater solubility were determined by testing a panel of genetically well-characterized *E. coli* strains with different phenotypes (Terpe, 2006). All of the *E. coli* strains [Rosetta(DE3), BL21(DE3), C41(DE3), C43(DE3)] have previously successfully functioned as expression systems for the purification and crystallization of a number of membrane proteins, each of the host strains had specific advantages pertaining to protein expression. *E. coli* BL21(DE3), lacking the lon and ompT proteases, which could degrade abnormal and outer membrane proteins, respectively, was suitable for high expression of normal and membrane proteins. *E. coli* Rosetta(DE3), which carried the pRARE plasmid that provided tRNA for *E. coli* and rarely used codons, was suitable for the expression of eukaryotic proteins. *E. coli* C41(DE3) and C43(DE3) were mutants designed to promote the expression of toxic and membrane proteins.

The two bands of OppC-GFP protein were not mere artifacts, but different folding states. Although the intensity ratios of the two bands depended on the protein expression levels, the ratios of the two bands could somewhat represent the expression quality; however, the exact quantities of the soluble and insoluble protein fractions could not be determined.

The OppC-GFP protein yield was much higher in the BL21(DE3) and Rosetta(DE3) strains than in C41(DE3) and C43(DE3) under the same induction conditions. The expression situations contribute to the intensity of the fluorescent band. Overall, high protein expression levels usually tended to produce aggregated proteins. Decreased speed of expression could sometimes give the strain more time for proper protein folding, and promoting the expression of fluorescent bands.

The gene deletion results indicated that *OppC* in *E. coli* was essential for the uptake of peptides, and not essential for cell viability. Protein-protein interaction results confirmed the essential role of OppC in oligopeptide permease system.

CONCLUSIONS

Membrane proteins are usually very important, as a majority are perfect drug targets. However, research on membrane proteins is limited according to the quality and quantity of properly folded proteins. Organisms such as *E. coli* are usually very sensitive to the massive membrane protein insertion. Therefore, in most cases, membrane proteins are expressed as inclusion bodies. Refolding of membrane proteins from the aggregated inclusion bodies is a huge challenge, and is not suitable for all membrane proteins, especially the inner membrane proteins, with intrinsic lower stability. This is because inner membrane proteins are mainly composed of α-helices, which are merely connected by side chain interactions.

The expression and purification protocol for a protein is usually unique. It depends mainly on the target protein sequence. However, some fusion proteins could obviously promote the expression level and solubility of target proteins. GFP is a very popular fusion protein, which helps in the proper folding of the target protein, and also works as a reporter. GFP is convenient for proper optimization of protein over-expression and purification. GFP is very stable, which tends to form well-folded fluorescent proteins. The stable structure of GFP might help the N-terminal-fused target protein to fold properly. On the other hand, proper folding of GFP usually depends on the correct folding of the target protein. Proper folding of the fusion protein results in fluorescence of the GFP. Therefore, the fluorescent GFP might represent the functionally expressed fusion proteins. The observed dual in-gel migration of the GFP fusion proteins could somewhat represent the folded and aggregated protein populations.

In this study, the expression and purification conditions of OppC were optimized using GFP. Noticeably, in-gel analysis showed that OppC-GFP expression was very obscure in the BL21(DE3) and Rosetta(DE3) strains, and brighter in the C41(DE3) and C43(DE3) strains. SDS-PAGE analysis revealed the aggregation of the OppC-GFP protein in BL21(DE3) and Rosetta(DE3). This phenomenon might be a result of too high expression levels. Optimization of the expression strain and induction conditions led to over-expression of the protein. Functional and pure OppC protein could be recovered easily from GFP fusion using the TEV protease digest.

The inner membrane protein OppC and OppB could form the transmembrane pore. Peptides have been proven to function as nitrogen sources; deletion of *OppC* in *E. coli* will impair the oligopeptide uptake. However, Opp is not the only peptide transporter in Gram-negative bacteria. Dipeptide transporters and glutathione transporters might contribute to the survival of the ΔOppC strain in the peptide-containing medium. Therefore, OppC is essential for the oligopeptide permease system, and not for the cell viability.

The interaction proteins might put insights into the functional mechanism of a protein, and reflect their functions. The protein-protein interaction of OppC with other components was determined in *E. coli* based on the irreversible reassembly process of GFP fragments. OppA did not interact with other Opp proteins. This could be a result of non-binding to the oligopeptide. Binding with oligopeptide is speculated to cause changes in the OppA conformation, which may initiate the oligopeptide transport by interacting with the channel proteins OppB and OppC. OppB and OppC could interact with OppD and OppF, which could hydrolyze ATP to support the transport process. Interestingly, OppD and OppF do not directly interact with each other, despite serving as ATP-binding components. This could be a result of the inactivation state of the oligopeptide permease system.

Using as a fusion tag, GFP promotes the soluble expression of OppC and improved the time-effectiveness of the purification process. Availability of the functional protein makes it possible

for future structural and functional studies. GFP, which is used as a reporter, directly reports on the interactions. The detection of protein-protein interactions is critical to our understanding of biological processes. In conclusion, effective and flexible use of GFP is greatly advantageous for the study of membrane proteins.

Conflicts of interest

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

Supplementary material

REFERENCES

Cundell DR, Pearce BJ, Sandros J, Naughton AM, et al. (1995). Peptide permeases from *Streptococcus pneumoniae* affect adherence to eucaryotic cells. *Infect. Immun.* 63: 2493-2498.

Drew DLM, Kunji E, Slotboom DJ and de Gier JW (2006). Optimization of membrane protein overexpression and purification using GFP fusions. *Nat. Methods* 3: 303-313.

Fenno JC, Tamura M, Hannam PM, Wong GW, et al. (2000). Identification of a *Treponema denticola* OppA homologue that binds host proteins present in the subgingival environment. *Infect. Immun.* 68: 1884-1892.

Flores-Valdez MA, Morris RP, Laval F, Daffe M, et al. (2009). *Mycobacterium tuberculosis* modulates its cell surface via an oligopeptide permease (Opp) transport system. *FASEB J.* 23: 4091-4104.

Geertsma ER, Groeneveld M, Slotboom DJ and Poolman B (2008). Quality control of overexpressed membrane proteins. *PNAS* 105: 5722-5727.

Goodell EW and Higgins CF (1987). Uptake of cell wall peptides by *Salmonella typhimurium* and *Escherichia coli*. *J. Bacteriol*. 169: 3861-3865.

Henrich B, Hopfe M, Kitzerow A and Hadding U (1999). The adherence-associated lipoprotein P100, encoded by an opp operon structure, functions as the oligopeptide-binding domain OppA of a putative oligopeptide transport system in *Mycoplasma hominis*. *J. Bacteriol.* 181: 4873-4878.

Juillard V, Le Bars D, Kunji ER, Konings WN, et al. (1995). Oligopeptides are the main source of nitrogen for *Lactococcus lactis* during growth in milk. *Appl. Environ. Microbiol.* 61: 3024-3030.

Kashiwagi K, Miyaji A, Ikeda S, Tobe T, et al. (1992). Increase of sensitivity to aminoglycoside antibiotics by polyamine-induced protein (oligopeptide-binding protein) in *Escherichia coli*. *J. Bacteriol.* 174: 4331-4337.

Letoffe S, Delepelaire P and Wandersman C (2006). The housekeeping dipeptide permease is the *Escherichia coli* heme transporter and functions with two optional peptide binding proteins. *Proc. Natl. Acad. Sci. U. S. A.* 103: 12891-12896.

Lin B, Short SA, Eskildsen M, Klempner MS, et al. (2001) Functional testing of putative oligopeptide permease (Opp) proteins of *Borrelia burgdorferi*: a complementation model in opp(-) *Escherichia coli*. *Biochim. Biophys. Acta* 1499: 222-231.

Magliery TJ, Wilson CGM, Pan W, Mishler D, et al. (2005). Detecting protein-protein interactions with a green fluorescent protein fragment reassembly trap: scope and mechanism. *J. Am. Chem. Soc.* 127: 146-157.

Matthews DM (1975). Intestinal absorption of peptides. *Phys. Rev.* 55: 537-608.

Moraes PM, Seyffert N, Silva WM, Castro TL, et al. (2014). Characterization of the Opp peptide transporter of *Corynebacterium pseudotuberculosis* and its role in virulence and pathogenicity. *BioMed Res. Int.* 2014: 489782.

Moutran A, Quaggio RB, Balan A, Ferreira LC, et al. (2004). The oligopeptide permease (Opp) of the plant pathogen *Xanthomonas axonopodis* pv. citri. *Curr. Microbiol.* 48: 354-359.

Newstead SDD, Cameron AD, Postis VLG, Xia XB, et al. (2010). Crystal structure of a prokaryotic homologue of the mammalian oligopeptide-proton symporters, PepT1 and PepT2. *EMBO J.* 20: 417-426.

Payne JW and Smith MW (1994). Peptide transport by micro-organisms. *Adv. Microb. Physiol.* 36: 1-80.

Perego M, Higgins CF, Pearce SR, Gallagher MP, et al. (1991). The oligopeptide transport system of *Bacillus subtilis* plays a role in the initiation of sporulation. *Mol. Microbiol.* 5: 173-185.

Reniere ML, Whiteley AT, Hamilton KL, John SM, et al. (2015). Glutathione activates virulence gene expression of an intracellular pathogen. *Nature* 517: 170-173.

Ruhfel RE, Manias DA and Dunny GM (1993). Cloning and characterization of a region of the *Enterococcus faecalis* conjugative plasmid, pCF10, encoding a sex pheromone-binding function. *J. Bacteriol.* 175: 5253-5259.

Suzuki HTK, Izuka S, Onishi A and Kumagai H (2005). The yliA, -B, -C, and -D genes of *Escherichia coli* K-12 encode a novel glutathione importer with an ATP-binding cassette. *J. Bacteriol.* 187: 5861-5867.

Terpe K (2006). Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. *Appl. Microbiol. Biotechnol.* 72: 211-222.

Thompson JD, Higgins DG and Gibson TJ (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22: 4673-4680.

Wang XG, Kidder JM, Scagliotti JP, Klempner MS, et al. (2004). Analysis of differences in the functional properties of the substrate binding proteins of the *Borrelia burgdorferi* oligopeptide permease (Opp) operon. *J. Bacteriol.* 186: 51-60.

Wang Z, Xiang Q, Wang G, Wang H, et al. (2011). Optimizing expression and purification of an ATP-binding gene gsiA from *Escherichia coli* k-12 by using GFP fusion. *Genet. Mol. Biol.* 34: 661-668.

Wilson CG, Magliery TJ and Regan L (2004). Detecting protein-protein interactions with GFP-fragment reassembly. *Nat. Methods* 1: 255-262.