



Hyper-production of porcine contagious pleuropneumonia subunit vaccine proteins in *Escherichia coli* by developing a bicistronic T7 expression system

Journal:	<i>Biotechnology Journal</i>
Manuscript ID	biot.202300187.R1
Wiley - Manuscript type:	Research Article
Date Submitted by the Author:	23-Nov-2023
Complete List of Authors:	Sun, Manman; Jiangnan University; Imperial College London - South Kensington Campus, Gao, Xiong Li, An Ledesma-Amaro, Rodrigo Wang, Peng; Key laboratory of high magnetic field and Ion beam physical biology, Hefei Institutes of Physical Science, Chinese Academy of Sciences Chen, Wenchao; Oil Crops Research Institute Chinese Academy of Agricultural Sciences Bai, Zhonghu; Jiangnan University, School of Biotechnology Liu, Xiuxia; Jiangnan University, School of Biotechnology
Primary Keywords:	Synthetic biology, Bioprocess engineering, Industrial biotechnology
Secondary Keywords:	Protein expression, Bioengineering, Bioprocess development, Biosynthesis, Recombinant proteins
Additional Keywords:	

SCHOLARONE™
Manuscripts

Author's response

Dear Editor and Reviewers,

We greatly appreciate the time and effort you invested in reviewing our manuscript (ID biot.202300187, "Hyper-production of porcine contagious pleuropneumonia subunit vaccine proteins in *Escherichia coli* by developing a bicistronic T7 expression system"). Your thoughtful comments and careful reading of our manuscript have been beneficial. In response to your comments, we have thoroughly revised and improved our manuscript. All alterations made in the revised manuscript are highlighted in red. Here are our responses to your comments.

Reviewer 1

General comment: The authors introduce bicistronic 5'UTRs in the commonly used T7 promoter pET system. They show increased production of GFP as well as three vaccine candidate proteins. For all (small increases) in production are seen. The antigens are also tested for immunization in mice. Overall a well written nice case study on BCD use for expression in *E. coli* Still some minor comments.

Response: We sincerely appreciate your positive feedback and constructive comments on our work. We acknowledge the significant enhancement in GFP expression using our BCD system, which could be attributed to the effective expression of GFP as a model protein. Although the yield increase for the three antigen proteins may not be as substantial as that of GFP, these small increases validate the effectiveness and value of our system.

Comment 1: the graphs showing improved production should also show if these increases are statistically significant, especially in figure 3.

Response: We appreciate your suggestion and have incorporated statistical analysis into the graphs to further solidify our data.

Comment 2: L52 replace 'exist' by 'has'

Response: Thanks for your suggestion, we have reorganized this paragraph and changed the description to "Despite these advancements, *E. coli* remains the primary expression system for PCP subunit vaccines, and the production of PCP subunit vaccines needs further improvement." in the revised manuscript (L49-51).

Comment 3: L79, 'sequences that allow efficient translation are not always suitable as fore-cistrons'. Please give some more explanation on this. And I wonder why the authors did not use the fore-cistrons' already demostrated in *E. coli* by Mutalik et al and several other authors? Good to briefly discuss-, there has been done more work on using BCDs for expression *E. coli* that I suggest the authors to look at and where appropriate cite <https://doi.org/10.1016/j.jmb.2014.11.002><https://doi.org/10.1111/1751-7915.1333210.1021/acssynbio.9b0010110.1186/s13568-022-01478-8>

Response: Your insightful comment is much appreciated. In principle, any sequence that permits efficient ribosome binding and translation initiation should be suitable as a fore-cistron.

However, our previous studies have shown that only a subset of sequences, chosen from the top 12 highly-translated genes, exhibited improved expression intensity when used as fore-cistrons in bicistronic expression systems (<https://doi.org/10.1186/s12934-020-01370-9>, <https://pubmed.ncbi.nlm.nih.gov/27580890/>). A study by Mukhopadhyay et al. describes a similar phenomenon wherein the use of a highly expressible gene (GST) as a fore-cistron sequence did not result in successful translation initiation due to the presence of internal negative control elements (<https://pubmed.ncbi.nlm.nih.gov/12089471/>). This demonstrates that sequences promoting efficient translation are not universally suitable as fore-cistrons. In our review on bicistronic design, we summarized key considerations for constructing a bicistronic expression cassette to enhance recombinant protein production (<https://doi.org/10.1007/s00253-021-11611-5>). We have revised our manuscript to clarify that the presence of internal negative control elements, such as potential complementary sequences or Shine-Dalgarno (SD)-like sequences, can affect the suitability of sequences as fore-cistrons (L72-76).

Actually, the fore-cistrons used in this study have already demonstrated good performance in *E. coli*. These fore-cistrons were selected based on our previous data in P_{btac} system. We selected 8 well-performing fore-cistrons in *E. coli* as candidates to construct P_{bT7} system. We realized that our previous description might not have been clear enough. Upon revision, we've ensured to provide a clearer description (L189-190, L290-291). The fore-cistrons demonstrated by Mutalik et al. and other researchers in *E. coli* indeed present viable options that are worth exploring in future studies.

Certainly, there are numerous related studies in *E. coli*. In our previous review, we provided a comprehensive summary of these studies (<https://doi.org/10.1007/s00253-021-11611-5>). Based on your suggestion, we have now included additional related content and references in the discussion section of this article (L283-286).

Comment 4: L87 before ' appeared in China ' I suggest to add 'which'

Response: We appreciate your suggestion and have revised the description into "Oml1 and Oml7 are the outer membrane lipoprotein of *A. pleuropneumoniae* serotypes 1 and 7, which commonly appeared in China." (L80-81).

Comment 5: L119 'activated' I would replace by 'inoculated'

Response: Thanks for pointing this out, we have made the correction according to your suggestion (L114).

Comment 6: L191 change 'well-performed' to 'well-performing'

Response: We appreciate your suggestion and have revised the phrase accordingly. (L189-190).

Comment 7: 192 change 'on' to 'from'

Response: Thanks for pointing it out, we have correct it according to your suggestion.

Comment 8: L203 change ' standardized' to normalized

Response: Thanks for pointing it out, we have revised the phrase accordingly. (L201).

Comment 9: L310, this statement is a bit bold as the authors did not observed improved translation in all cases, please make this statement less strong and maybe discuss alternative influences of the changed 5 UTR in some cases (higher transcript stability) And as a last comment I suggest the authors to make the 4 well performing pET vectors available to the community via Addgene

Response: Thank you for your valuable comments and suggestions. We agree that our initial statement was a bit strong and did not fully represent the complexity of the observed effects. We have revised this statement to make it more accurate and have discussed other possible mechanisms behind this (L290-304). Furthermore, we appreciate your suggestion to make the four well performing pET vectors available to the community, and we are in the process of depositing these vectors to Addgene.

Reviewer 2

Comment 1: Page 3, Line 50, "as all 15 serotypes of *A. pleuropneumoniae* except serotype 10 express it." This sentence should be rewritten.

Response: Thanks for your suggestion, we have revise the sentence to "Among various Apx toxins, ApxII stands out as the most promising vaccine candidate due to its expression across most *A. pleuropneumoniae* serotypes, excluding serotype 10" (L44-45).

Comment 2: Page 3 "Introduction", the term "unprecedented" is not appropriate to describe the yield without comparing it to the work of other researchers. A brief description in the introduction is recommended.

Response: We appreciate your suggestion. We agree that providing a brief description would be more appropriate to give context to the yield achieved in our study. In the revised version, we have provided a comparative overview of the work carried out by other researchers in the field, this will ensure that readers have a better understanding of the significance of the yield achieved in our research. The description of 'So far, Oml and ApxII have been successfully cloned and expressed in *E. coli*, yeast, and *Corynebacterium glutamicum* (*C. glutamicum*).^[8-12] The highest yield of Oml has been achieved in *C. glutamicum*, with a secretion level of 1.01 g/L.^[12] ApxII primarily exists as inclusion bodies in *E. coli*, but soluble expression has been achieved via truncation of the protein sequence, yielding up to 10 mg/L.^[10]' has been added to the revised version. (L47-49)

Comment 3: Page 5, Line 99, "DNA manipulation and Plasmid construction" Plasmid should be in small letter.

Response: Thank you for catching this. We have corrected the word "Plasmid" to "plasmid" in the revised version (L93).

Comment 4: Page 6, Line 122, "diluted at about 0.5" means 0.5 fold or ?

Response: Thank you for inquiring for clarification. The phrase "diluted at about 0.5" in the original manuscript (Page 6, Line 122) was intended to mean that the cells were diluted to achieve a cell density (OD₆₀₀) of approximately 0.5. We have clarified this point in the revised

manuscript (L116-117).

Comment 5: Page 7, Line 136, “micro-well plates” was first appeared in Line 25, the abbreviation MWP should be written in Line 25.

Response: Thanks for your correction, we have corrected it in the revised version (L26, L139).

Comment 6: Page 12, Line 250, “lower” should be “Lower”. Line 253, “production” should be “production yield” or “yield”.

Response: Thank you for catching these errors. We have made the necessary corrections in the revised manuscript (L241-244).

Comment 7: The biological assays for the testing of vaccines activities were only conducted by mouse experiment, which is unconvincing. Is there any other method used to test the bioactivities of the vaccines?

Response: The utilization of mouse models is a standard approach in the preliminary testing of the protective efficacy of *A. pleuropneumoniae* protein-based vaccines. Numerous studies have shown that the mouse model is effective in conducting an initial screening of antigen, which aids in selecting promising candidates for further testing in the natural host (summarized in this review, <https://doi.org/10.1016/j.vetmic.2018.02.028>). In the subsequent stages of research, further testing in swine models may provide more specific information. Nevertheless, data from mouse model tests still contribute significantly to our understanding of vaccine activity against *A. pleuropneumoniae*. Therefore, despite some limitations, the mouse model offers a balance between cost-effectiveness, time efficiency, and the generation of valuable insights.

In addition to animal models, Enzyme-Linked Immunosorbent Assay (ELISA) can be used to test the immunogenicity of the vaccine, although it does not directly measure protective efficacy. We will consider implementing these additional testing methods in our future research.

Comment 8: The text requires rather extensive English revisions before it can be accepted for publication.

Response: Thank you for your comment, we have thoroughly revised the manuscript to improve the clarity and quality of the English language. We hope that these edits have addressed your concerns.

Comment 9: Page 13, Line 272, How to measure the purity of vaccine? Please describe in detail.

Response: We appreciate your suggestion. In response, we've included the following analysis steps in the methods section: “The purified protein was analyzed through SDS-PAGE, and the gel image was processed using Image J software to measure band intensities. The relative purity of the protein was estimated by comparing the intensity of the desired protein band to the total protein band intensities in the same lane.” in the methods section (L173-176).

Comment 10: Page 13 “3.6 Protective immunity induced by recombinant PCP antigen against *A. pleuropneumoniae*”, I think it is questionable to explain its protective effect by stating that the activity of Oml7 is lost during purification or storage. Positive control or optimization of

purification and storage methods may explain this issue.

Response: We thank you for your suggestion and realize our initial description was inadequate. We've updated the description in the revised manuscript into "In contrast, Oml7, which showed a high protective effect when previously expressed in *C. glutamicum*,^[12] exhibited the lowest protective effect among the three antigens tested, with a survival rate of 50%. However, this was still significantly higher than the 20% survival rate observed in the control group, suggesting that Oml7 retains some protective potential against *A. pleuropneumoniae*." in the revised manuscript (L269-272).

Comment 11: Authors said they use AKTA purification system in methods, however, no related results were shown in the results.

Response: In our manuscript, the related SDS-PAGE results are presented in Figure 6D-F. We understand that our description may not have been sufficiently clear, leading to an oversight. In response to your comment, we've revised the manuscript to ensure the data presentation is more direct and clear (L259-261).

Comment 12: Part 2.6 is better put before part 2.5.

Response: We appreciate your suggestion. We agree that this rearrangement would improve the logical flow of the paper and reorganize these parts in the revised version (L132-139).

Comment 13: In discussion part, authors simply stated the obvious results and no deep discussion. For example, there is too much background introduction, Figure 2B and the activity of Oml7 are not explained clearly. Please revise the discussion section carefully.

Response: Thank you for your suggestion, we realized the discussion section lacked depth and we appreciate your specific feedback. We have revised this section accordingly. We have condensed the extended background introduction and shifted our focus more towards the interpretations and implications of our findings. We've included a discussion on the potential reasons for the increased transcriptional level seen in Figure 2B, and also discussed the possible reasons for Oml7's lower activity. We believe these revisions will enhance understanding of our study. (L276-324)

Comment 14: Page 5 Line 100, "Plasmid extraction and PCR purification used kits from CWBIO" should be change to "Plasmid extraction and polymerase chain reaction (PCR) purification used kits from CWBIO". And "The polymer chain reaction" in the next sentence can be abbreviated as "PCR".

Response: We have corrected this and thank you for pointing it out (L94-95).

Comment 15: Page 6 Line 105, "EGFP" should be change to "Enhanced Green Fluorescent Protein".

Response: We have corrected this and thank you for pointing it out (L99).

Comment 16: Page 6 Line 123, "a 507 emission wavelength" should be change to "a 507 nm emission wavelength".

Response: We acknowledge your correction and have revised the description to "a 507 nm

emission wavelength" (L118-119).

Comment 17: Page 25 Line 519, "DO level" should be change to "Dissolved oxygen (DO) level". Alternatively, change the "Dissolved oxygen" on Page 9, Line 172 to "Dissolved oxygen (DO)".

Response: Thanks for your correction, we have revised the "Dissolved oxygen" on Page 9, Line 172 to "Dissolved oxygen (DO)" (L168).

Comment 18: Page 27 Fig 1B, "Pet28-EGFP" should be change to "pET28-EGFP". It is necessary to proofread "pET28" in this paper.

Response: Thanks for your correction, we have corrected it in Fig. 1 and proofread all "pET28" in the manuscript.

For Peer Review

1
2
3
4 1 **Hyper-production of porcine contagious pleuropneumonia subunit vaccine proteins in *Escherichia coli* by**
5
6 2 **developing a bicistronic T7 expression system**
7
8
9 3 Manman Sun^{1,2,3} #, Alex Xiong Gao⁴ #, An Li¹ #, Rodrigo Ledesma-Amaro², Peng Wang³, Wenchao Chen⁵,
10
11 4 Zhonghu Bai¹, Xiuxia Liu^{1*}
12
13
14 5
15
16
17 6 #Co-first author
18
19 7 ¹National Engineering Research Center of Cereal Fermentation and Food Biomanufacturing, Jiangnan University,
20
21 8 Wuxi 214112, China
22
23
24 9 ² Department of Bioengineering and Imperial College Centre for Synthetic Biology, Imperial College London,
25
26 10 London SW7 2AZ, UK
27
28
29 11 ³ Key laboratory of high magnetic field and Ion beam physical biology, Hefei Institutes of Physical Science,
30
31 12 Chinese Academy of Sciences
32
33
34 13 ⁴Division of Life Science, The Hong Kong University of Science and Technology, Hong Kong, China
35
36
37 14 ⁵Oil Crops Research Institute of the Chinese Academy of Agricultural Sciences, Wuhan, 430062, China
38
39
40 15 * Correspondence: Xiuxia Liu, liuxiuxia@jiangnan.edu.cn
41
42
43 16
44

45 17 **Abstract**
46
47
48 18 The ApxII toxin and the outer membrane lipoprotein (Oml) of *Actinobacillus pleuropneumoniae* (*A.*
49
50 19 *pleuropneumoniae*) are important vaccine antigens against porcine contagious pleuropneumonia (PCP), a
51
52 20 prevalent infectious disease affecting the swine industry worldwide. Previous studies have reported the
53
54 21 recombinant expression of ApxII and Oml in *Escherichia coli* (*E. coli*), however, their yields were not satisfactory.
55
56 22 Here, we aimed to enhance the production of ApxII and Oml by constructing a bicistronic expression system based
57
58
59
60

on the widely used T7 promoter. To create efficient T7 bicistronic expression cassettes, 16 different fore-cistron sequences were introduced downstream of the T7 promoter. The expression of three vaccine antigens Oml1, Oml7, and ApxII in the four strongest bicistronic vectors were enhanced compared to the monocistronic control. Further optimization of the fermentation conditions in micro-well plates (MWP) led to improved production. Finally, the production yields reached unprecedented levels of 2.43 g/L of Oml1, 2.59 g/L of Oml7, and 1.21 g/L of ApxII, in a 5 L bioreactor. These three antigens also demonstrated well-protective immunity against *A. pleuropneumoniae* infection. In conclusion, this study establishes an efficient bicistronic T7 expression system that can be used to express recombinant proteins in *E. coli* and achieves the hyper-production of PCP vaccine proteins.

Keywords: bicistronic T7 expression system, PCP subunit vaccine protein, hyper-production, *Escherichia coli*, recombinant expression

1. Introduction

Porcine contagious pleuropneumonia (PCP) caused by *Actinobacillus pleuropneumoniae* (*A. pleuropneumoniae*) is a highly contagious and often fatal disease in pigs.^[1,2] This disease has led to significant economic losses to the global pig-rearing industry and is recognized as one of the five most critical diseases threatening this industry. Antibiotics have been traditionally used to prevent PCP outbreaks, but their overuse has led to antibiotic resistance in pathogenic bacteria. Furthermore, numerous countries have implemented restrictive policies on antibiotics use in livestock and poultry feed, thereby escalating the demand for alternative solutions such as PCP vaccines.^[3,4] However, current PCP vaccines based on inactivated whole cells from several pathogenic serotypes offer limited cross-protection.^[5] Subunit vaccines derived from immune-protective antigens, such as Apx toxins and outer membrane lipoprotein (Oml), which are common to many serotypes, have demonstrated strong cross-immune effects.^[5,6] Among various Apx toxins, ApxII stands out as the most promising vaccine candidate due to its expression across most *A. pleuropneumoniae* serotypes, excluding serotype 10.^[6,7] So far, Oml and ApxII have been successfully cloned and expressed in *E. coli*, yeast, and *Corynebacterium glutamicum* (*C. glutamicum*).^[8–12] The highest yield of Oml has been achieved in *C. glutamicum*, with a secretion level of 1.01 g/L.^[12] ApxII primarily exists as inclusion bodies in *E. coli*, but soluble expression has been achieved via truncation of the protein sequence, yielding up to 10 mg/L.^[10] Despite these advancements, *E. coli* remains the primary expression system for PCP subunit vaccines, and the production of PCP subunit vaccines needs further improvement.

To enhance the production of recombinant proteins in *E. coli*, several strategies have been adopted. These include the development of strong promoters and ribosome binding sites (RBSs), modification of host cells, and optimization of fermentation conditions.^[13–16] Promoters are deemed the most effective way of controlling gene expression and can determine nearly 80% of the protein levels. Promoters from different origins, such as bacteria

(Lac, tac, trp, and araBAD) and bacteriophage (T7, T5, and SP6), are commonly utilized in *E. coli*.^[14] The pET expression vectors, based on the T7 promoter, have gained popularity due to their superior performance,^[14,17] with over 90% of the PDB proteins in 2003 produced using this promoter.^[17]

Despite the benefits of strong promoters in enhancing gene expression, the expression level of target genes can be limited by other genetic elements. One such limitation arises from unfavorable mRNA secondary structures in the translation initiation region (TIR), which can impede translation initiation even with a strong promoter controlling the target gene.^[18] While optimizing the TIR region for each target gene is a potential solution, it is a tedious and time-consuming process.^[19] An alternative approach involves using a bicistronic design (BCD), where a short coding sequence (CDS) is inserted as the fore-cistron upstream of the target gene.^[20] The intrinsic helicase activity of the ribosome recruited by the translation of fore-cistron can disrupt the mRNA structure around the target gene TIR, facilitating its translation initiation.^[20,21] Since the fore-cistronic peptide serves no functional purpose in BCD, it can be modified and replaced independently. A well-translated fore-cistron sequence can be introduced as an “enhancer” to further improve the expression level of the target protein through translation coupling. For instance, in our previous study, the expression of seven proteins was greatly enhanced in *C. glutamicum* by using bicistronic P_{tac} systems containing strong fore-cistron sequences.^[19] However, it is worth noting that constructing a new bicistronic system to enhance recombinant protein expression is not a straightforward task. Due to the existence of internal negative control elements such as potential complementary sequences or shine-dalgarno (SD)-like sequences, not all sequences that allow efficient ribosomal translation are suitable for use as fore-cistrons.^[19,22,23] Consequently, to ensure the effectiveness of the newly constructed BCD system in enhancing protein production, it becomes necessary to optimize the genetic elements involved, particularly the fore-cistron sequence.

In this study, we aim to enhance the production of PCP subunit vaccine proteins in *E. coli* by constructing a

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

bicistronic T7 expression system. We first evaluated the impact of various fore-cistron sequences on BCD expression intensity and selected the top four strongest BCD vectors for the expression of three PCP vaccine proteins: Oml1, Oml7, and ApxII. **Oml1 and Oml7 are the outer membrane lipoproteins of *A. pleuropneumoniae* serotypes 1 and 7, which commonly appeared in China.** We optimized culture conditions, induction patterns, and medium composition to further improve protein yields. Remarkably, fed-batch cultivation in a 5 L bioreactor resulted in unprecedented high yields of Oml1 (2.43 g/L), Oml7 (2.59 g/L), and ApxII (1.21 g/L). Furthermore, we demonstrated that these recombinant antigens display excellent immune protective efficacy in mouse models.

2. Materials and methods

2.1 Strains and culture condition

The strains and plasmids used in this study are listed in **Table S1**. *E. coli* JM109 was used for plasmid construction, and BL21 was used for protein expression. *E. coli* JM109 was cultured in LB medium or on an LB plate containing 1.5% (w/v) agar at 37 °C. *E. coli* BL21 was cultured in LB medium or TBSB medium (yeast extract 24 g, tryptone 12 g, KH₂PO₄ 2.31 g, K₂HPO₄·3H₂O 12.55 g, glycerin 10 g, and 1 L H₂O) at 37 °C. The final concentration of kanamycin for *E. coli* JM109 and BL21 was 50 mg/L.

2.2 DNA manipulation and plasmid construction

Plasmid extraction and polymerase chain reaction (PCR) purification used kits from CWBIO. PCR was carried out using EsTaq polymerase (CWBIO, China) or PrimerSTAR (TaKaRa, China). The restriction endonuclease was purchased from TaKaRa. Gene sequences were inserted into target vectors through ligation reaction (TaKaRa, China) or homologous recombination (Abclone, China). All experiments were performed following the standard procedure.

Table S2 lists all the primers used in this study. **First, the enhanced green fluorescent protein (EGFP)**

sequence was inserted into pET28a through *Nco* I and *Hind* III to obtain pET28-EGFP. This plasmid was used as a backbone to construct bicistronic T7 expression vectors containing different fore-cistronic sequences (**Fig. 1A**). The construction process was as follows: A 62 bp fore-cistron sequence ending with a conserved SD2 (AAAGGAGGACAATAATG) and a TAATG translation coupling frame was obtained by primer annealing. Then, pET28-EGFP was digested with *Nco* I, and these sixteen different sequences were inserted upstream of EGFP through homologous recombination, respectively. The *Nco* I cleavage site was removed from the final plasmid (**Fig. 1B**).

To construct expression plasmids for PCP vaccine protein production, the *Oml1*, *Oml7*, and *Apx II* genes were codon-optimized and synthesized by Genewiz (Suzhou, China). The sequences are listed in **Supplementary material**, and all three genes were appended with a 6×his tag at the 3' end. Then, these gene sequences and the corresponding fore-cistron sequences were inserted into the pET28a to obtain BCD expression vectors. The expression plasmid containing a monocistronic T7 promoter was used as a control.

2.3 EGFP intensity measurement

After successful construction in *E. coli* JM109, all plasmids were transformed into *E. coli* BL21. Three transformants were picked up for each expression strain and inoculated in a 24-deep-well plate containing 2 mL LB medium and 50 mg/L kanamycin at 37 °C, 220 rpm for 12 h. Then, the cultures were transferred to 2 mL fresh LB medium with a 10% (v/v) inoculation rate, and 1 mM IPTG was directly added after inoculation. After 24 h of induction, the cells were harvested and diluted to an OD₆₀₀ of approximately 0.5. The fluorescence intensity was measured using a fluorescence spectrophotometer with a 488 nm excitation wavelength and a 507 nm emission wavelength. The unit fluorescence intensity of each sample was calculated by normalizing the fluorescence intensity by OD₆₀₀.

2.4 Real-time quantitative PCR (qPCR) and calculation of relative translation efficiency

To compare the EGFP transcriptional level of different plasmids, *E. coli* strains were in the LB medium as described above. 1 mL of cultures was collected by centrifugation and washed twice with ice-cooled PBS. Total RNA extraction (CWBIO, China), reverse transcription (Vazyme, China), and qPCR (Vazyme, China) were then performed according to the manufacturer's instructions. qPCR was carried out using the StepOnePlus (Applied Biosystem, USA) system and the condition was set as follows: 95 °C for 30 s and 40 cycles at 95 °C for 15 s, 62 °C for 30 s, and 72 °C for 20 s. Each sample contained three biological repeats and three duplicated wells, the relative EGFP transcription level was analyzed by the $2^{-\Delta\Delta C_t}$ method.^[19] The 16S rRNA was used as the endogenous control, and the EGFP transcription level of pET28-EGFP was defined as 1. The relative EGFP translation efficiency is calculated by dividing the EGFP fluorescence intensity by mRNA abundance as previously reported.^[24,25]

2.5 SDS-PAGE analysis and protein quantitation

After 24 h of cultivation, 10 OD₆₀₀ of cells harboring PCP vaccine proteins were harvested by centrifugation at 12,000 g and 4 °C for 10 min. The cells were washed and resuspended in 1 mL PBS, followed by sonication on ice for disruption. After sonication, the lysates were centrifuged at 12,000 g for 15 min, and the supernatants were collected for 12% (w/v) SDS-PAGE analysis. The concentration of the target protein band was determined by measuring and comparing its strength with that of standard bovine serum albumin (BSA) at a known concentration loaded in the same gel using Image J software.

2.6 Optimization of PCP vaccine protein production in MWP

Optimization of expression vector

PCP vaccine protein genes were cloned into different expression vectors. To compare the expression performance of these vectors, three transformants were selected for each strain and activated in MWP containing 2 mL of LB medium and 50 mg/L of kanamycin at 37 °C and 220 rpm for 12 h. Subsequently, the cultures were

transferred to 2 mL fresh TBSB medium with a 10% (v/v) inoculation rate, and 1 mM IPTG was directly added after inoculation. After 24 h of induction, the cells were collected and disrupted for SDS-PAGE analysis.

Optimization of cultivation temperature

The cultivation conditions were set as described above. After the addition of IPTG, the MWP was kept at 25°C, 30°C, 37°C, and 220 rpm for 24 h, respectively.

Optimization of induction conditions

For the optimization of IPTG concentration, after the culture was transferred to TBSB, IPTG with final concentrations of 0, 0.05, 0.1, 0.2, 0.4, 0.6, 0.8, and 1.0 mM was added to the culture medium. The MWP was then kept at 25°C and 220 rpm for 24 h.

For the optimization of the pre-induction period, three transformants were selected for each strain and activated in MWP containing 2 mL LB medium and 50 mg/L kanamycin at 37 °C, 220 rpm for 12 h. Then, the cultures were transferred to a 2 mL fresh TBSB medium with a 10% (v/v) inoculation rate. 0.1 mM IPTG was added to the culture medium at 0, 1, 2, 3, and 4 h after inoculation, respectively. All MWPs were kept at 25°C, 220 rpm for 24 h.

Optimization of medium components

TBSB was used as the initial medium for medium components optimization. For carbon source optimization, 10 g/L of glucose, fructose, xylose, and sucrose were added into the TBSB medium to replace the original glycerol. The cell growth and protein production using glucose carbon source were used as the control. For nitrogen source optimization, an additional 10 g/L of urea, (NH₄)₂SO₄, and NH₄Cl were added to the TBSB medium, and the protein yield and OD₆₀₀ in the TBSB medium were used as the control. For all medium optimization experiments, the cultivation and induction conditions were set at the previously determined optimal level.

2.7 Fed-batch cultivation and purification of PCP antigen proteins

The fed-batch culture was carried out in a 5 L bioreactor (Applikon, Holland) with a working volume of 2 L. After overnight activation in LB medium, 200 mL *E. coli* seed solution was transferred to a 1.8 L TBSB medium. Throughout the fermentation process, the dissolved oxygen (DO) was maintained at 30% (v/v). The pH was controlled at 7.0, and the stirring speed was set at 400-1000 rpm. The temperature was initially maintained at 37 °C for 2 h and then changed to 25 °C after IPTG addition. The induction conditions refer to the optimum level determined in MWP. To avoid starvation, a feed medium (50 g/L yeast extract, 714 g/L glycerin, and 25 g/L (NH₄)₂SO₄) was added with a feeding rate of 3.5 mL/h/L after 12 h of inoculation. The antigen protein was purified using an AKTA purifier system (GE, Sweden) and a HisTrap HP affinity column. The purified protein was analyzed through SDS-PAGE, and the gel image was processed using Image J software to measure band intensities. The relative purity of the protein was estimated by comparing the intensity of the desired protein band to the total protein band intensities in the same lane.

2.8 Challenge experiment with *A. pleuropneumoniae* serotype 7

The *A. pleuropneumoniae* serotype 7 strain used in this study was provided by Tecon Biology Co., Ltd and was cultured in BHI medium supplemented with 10 mg/L β-nicotinamide adenine dinucleotide at 37 °C.[26] For the challenge experiment, forty mice were divided into four groups: three experimental groups were immunized with 20 μg purified Oml1, Oml7, or ApxII, respectively. The control group was immunized with PBS. After 14 days of immunization, all mice were intraperitoneally injected with 0.5 mL of *A. pleuropneumoniae* serotype 7 (6×10⁸ CFU). The survival of mice was recorded for the following 7 days.

3. Results

3.1 Design and construction of the bicistronic T7 expression system

Artificial bicistronic structures in prokaryotes have demonstrated the capability to enhance the production of

recombinant proteins.^[20] To further improve the expression performance of the T7 promoter in *E. coli*, we engineered this promoter in a bicistronic manner. We selected eight well-performing fore-cistrons from the P_{btac} system based on our previous data in *E. coli*,^[19] and optimized their sequences according to the *E. coli* codon preference, thereby generating 8 additional sequences. This yielded a total of 16 fore-cistron sequences, which were used to construct bicistronic T7 promoter systems (**Table 1**). Following the characteristics of the BCD expression cassette, we inserted a conserved SD sequence (the second SD motif, SD2) and a translation coupling frame between the fore-cistron and the reporter gene, EGFP (**Fig. 1A and 1B**).^[20] The LacI gene, located upstream of the T7 promoter, confers an inducible characteristic to the BCD system, thus requiring the addition of IPTG for the expression of recombinant proteins using the bicistronic T7 systems.

3.2 Effect of different fore-cistron sequences on the performance of bicistronic T7 expression systems

The expression plasmid was initially constructed in *E. coli* JM109 and subsequently transferred into *E. coli* BL21, a strain known for its efficiency in expressing recombinant proteins. To evaluate the performance of different fore-cistron sequences in the bicistronic T7 expression system, EGFP fluorescence intensity was measured and normalized to OD₆₀₀ after 24 h of cultivation in MWP. The EGFP expression level controlled by a monocistronic T7 in pET28-EGFP served as a control. The results in **Fig. 2A** demonstrated substantial fluorescence intensity in the bicistronic T7 systems, with 10 out of the 16 vectors showing enhanced levels compared to the control. This indicates that the fore-cistron sequence significantly impacts the expression of the target gene and could either enhance or diminish its expression. Four vectors-pET28-HT5, pET28-HT8, pET28-HT12Y, and pET28-HP11Y-were identified as the most potent, increasing expression levels by 68%, 65%, 76%, and 80%, respectively.

The EGFP transcription level of the four most potent BCD vectors was tested, and their translation efficiency was calculated to understand the reasons for the increased EGFP expression level in the bicistronic T7 system.

Unlike the generally improved translation efficiency we previously observed in *C. glutamicum*,^[19] different effects of fore-cistron sequences were observed across the four vectors (**Fig. 2B**). For instance, pET28-HT5-EGFP exhibited increased EGFP expression due to enhanced translation efficiency, while pET28-HP11Y-EGFP showed decreased translation efficiency but improved EGFP fluorescence intensity due to an increased transcriptional level.

3.3 Enhanced expression of PCP vaccine proteins Oml1, Oml7, and Apx II

To determine if the expression intensity of the four bicistronic T7 vectors is influenced by the target gene sequences and to evaluate their effectiveness in expressing other recombinant proteins, three PCP vaccine antigens: Oml1, Oml7, and ApxII, were expressed using the four best-performing T7 BCD systems. As the full-length ApxII protein was observed to be expressed as an inclusion body in *E. coli* (data not shown), a truncated ApxII fragment 5 (439-801 aa), known for its good immunogenicity in previous studies, was selected for expression.^[10,11] SDS-PAGE confirmed the successful expression of all three antigens. Additionally, the protein yield per OD₆₀₀ of all four BCD vectors exceeded that of the monocistronic control (**Fig. 3A-3C**). The relative protein yield per liter of cells for each vector was also calculated, revealing that the use of pET28-HT5, pET28-HT8, pET28-HP11Y, and pET28-HT12Y increased the yield of Oml1 by 38%, 29%, 29%, and 24%; of Oml7 by 13%, 8%, 18%, and 17%; and of ApxII by 24%, 18%, 22%, and 17%, respectively (**Fig. 3D-3F**). These results suggest that all four bicistronic T7 vectors have good compatibility and can enhance the expression of various proteins. Based on these results, we selected the strongest vector for each protein (pET28-HT5-Oml1, pET28-HP11Y-Oml7, and pET28-HT5-ApxII) for subsequent protein expression experiments.

3.4 Optimization of PCP vaccine protein expression conditions in MWP

Apart from the expression element, cultivation temperature, induction conditions, and medium components (such as carbon and nitrogen sources) are also important factors, which can influence cell growth, cell wall

structure, energy metabolism, and consequently, protein expression.^[27–29] To optimize the production of PCP vaccine proteins, we first examined temperature effects. As shown in **Fig. 4A–4C**, 25°C was the most favorable temperature for the expression of the three tested proteins: Oml1, Oml7, and ApxII. The beneficial effect of lower temperature on protein expression has also been observed in previous studies.^[30,31] Thus, 25°C was used as the cultivation temperature in subsequent PCP vaccine protein expression experiments.

We also optimized the IPTG concentration and pre-induction period. Increasing the IPTG concentration from 0 to 0.2 mM resulted in the appearance of Oml1 and Oml7 bands on SDS-PAGE and ApxII expression was first observed with 0.1 mM IPTG addition (data not shown). No significant effect on cell growth was observed within the IPTG concentration range tested (data not shown). 0.2 mM IPTG was optimal for the three proteins tested (**Fig. 4D–F**). Therefore, 0.2 mM IPTG was used for PCP antigen production. The effect of the pre-induction period on protein production is shown in **Fig. 4G–4I**, a 2 h pre-induction period was found to be optimal for the expression of Oml1, Oml7, and ApxII. Shorter pre-induction periods led to lower protein expression levels, likely due to an increased metabolic burden during early growth phases. Longer pre-induction periods similarly decreased production yield for all three proteins. Therefore, a 2-h pre-induction period was chosen for subsequent experiments.

We assessed the impact of different carbon and nitrogen sources on PCP antigen production. As shown in **Fig. 5A–5C**, Glycerol emerged as the most effective carbon source, increasing the yield of Oml1, Oml7, and ApxII by 30%, 15%, and 30% respectively, compared to glucose. Unexpectedly, the addition of extra nitrogen sources (urea, (NH₄)₂SO₄, and NH₄Cl) to TBSB led to a decrease in cell growth and protein yield (**Fig. 5D–5F**), suggesting that the original TBSB nitrogen source was adequate for cell growth and protein expression. Thus, in subsequent experiments, the nitrogen source was kept as in the original culture medium.

3.5 High-level production of Oml1, Oml7, and Apx II by fed-batch cultivation

To achieve high-level production of Oml1, Oml7, and ApxII in *E. coli*, fed-batch cultivation was performed for each protein in a 5 L bioreactor with a 2 L working volume. As shown in **Fig. 6A-6C**, cell growth and protein accumulation displayed similar trends for all three antigens, with the highest cell density and maximum yield observed at 30 h after inoculation. The maximum OD₆₀₀ of Oml1, Oml7, and ApxII were 98, 110, and 80, respectively, correlating with protein yields of 2.43 g/L, 2.59 g/L, and 1.21 g/L. Unlike the continuous protein accumulation observed in secretory expression systems,^[12,32] protein accumulation ceased when cell density plateaued during intracellular expression. The recombinant antigens Oml1, Oml7, and ApxII were then purified using a HisTrap HP affinity column. As shown in **Fig. 6D-6F**, after simple purification steps, all three proteins were successfully purified with high purity (>90%). This allows us to further explore whether these three recombinant antigens are biologically active.

3.6 Protective immunity induced by recombinant PCP antigen against *A. pleuropneumoniae*

Mice injected intraperitoneally with *A. pleuropneumoniae* serotype 7 (APP7) were used as protective models to examine the protective immunity of PCP antigens produced in this work.^[5,10] Given the small molecular weight and neutral charge of the 6×His tag under physiological conditions, it was hypothesized that its presence would not affect the immunogenicity of the recombinant antigens.^[12] As shown in **Fig. 6G**, mice immunized with Oml1 and ApxII were all well protected against a lethal APP7 challenge, 90% and 80% of the mice were still alive even 168 h after infection, respectively. In contrast, Oml7, which showed a high protective effect when previously expressed in *C. glutamicum*,^[12] exhibited the lowest protective effect among the three antigens tested, with a survival rate of 50%. However, this was still significantly higher than the 20% survival rate observed in the control group, suggesting that Oml7 retains some protective potential against *A. pleuropneumoniae*. The robust immune protective effect of Oml1 observed in this study reaffirms the strong cross-protection effect of the subunit vaccine against different serotypes, underscoring its potential utility in PCP prevention.

4. Discussion

To date, numerous expression systems for recombinant protein production have been developed, including mammalian cells, insects, plants, yeast, and bacteria.^[13,33,34] Among them, *E. coli* has been widely utilized as a 'workhorse' for the production of recombinant proteins due to its ease of manipulation, rapid growth, well-defined genetic background, and high yield capacity.^[35,36] Typically, strong promoters are favored to achieve satisfactory protein yields, which is why the T7 promoter and its derived pET expression vectors are the most commonly used systems in *E. coli*.^[14] However, the use of strong promoters does not always guarantee high levels of gene expression due to the influence of other genetic elements such as 5'UTR and TIR on mRNA stability and secondary structure.^[37,38] To overcome this issue, the bicistronic design has been employed. This design, derived from the polycistron in archaea and bacteria, has been applied in various prokaryotic systems such as *E. coli*, *C. glutamicum*, and *Bacillus subtilis*, for enhanced protein expression, membrane protein production, gene expression control, protein translation, and abundance monitoring, etc.^[20,39–41] In this study, we further applied the bicistronic strategy to the T7-based pET28 expression system to further enhance its expression intensity and performance. Previous work has shown that the arbitrary introduction of a fore-cistron may not necessarily boost promoter strength, highlighting the necessity of fore-cistron screening when constructing new BCD systems.^[19,20]

To ensure the successful enhancement of expression performance, we constructed a series of bicistronic T7 systems based on eight well-performing fore-cistrons identified from the P_{btac} system in *E. coli*.^[19] Interestingly, unlike most previous studies that reported an increase in translation efficiency with BCD, our research observed a general increase in transcription levels in the four selected bicistronic T7 systems (**Figure 2**). This phenomenon aligns with a previous study of ours on leaderless BCD.^[42] Possible reasons for the increased transcription level may include the following: the fore-cistron sequence could contain regulatory elements such as operators or activators that can bind to transcription factors and modulate the rate of transcription. Small fore-cistron peptides

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

might affect cell growth status and consequently regulate transcription levels.^[20] Additionally, the bicistronic design might introduce structural changes, enhancing the gene's accessibility and thereby promoting transcription.^[43] We also observed good compatibility and enhancement effects across different genes within our bicistronic T7 systems (**Figure 3**). This compatibility could potentially be attributed to the translation of the fore-cistron, which might minimize unfavorable mRNA secondary structures and consequently mitigate the effect of the target gene sequence on the gene expression level.^[20,44] Furthermore, the introduction of the fore-cistron could also alter the 5' UTR, possibly increasing mRNA stability, promoting gene expression, and reducing sequence-specific impact.^[45]

Despite these promising results, there is still room for improvement in the T7 bicistronic expression system. For instance, the expression intensity of the T7 BCD system can be further enhanced by optimizing other genetic elements, such as SD2 and the length of the fore-cistron.^[20] In this study, three recombinant antigens induced a protective immune response against *A. pleuropneumoniae*. However, the protective effect of Oml7 was lower than expected, which contrasts with its high protective effect observed when previously expressed in *C. glutamicum*.^[12] This variation could be attributed to differences in protein folding and protein stability. Intracellular conditions, such as varying pH, ion concentrations, and redox potential, potentially influence protein activity between different host cells. Further studies are required to explore the exact causes. In addition, the protective effect of conjugate antigens Oml1/ApxII, Oml7/ApxII, and Oml1/Oml7/ApxII remains to be evaluated, previous studies have suggested that combining immunization with Apx toxin and Oml can lead to an improved survival rate.^[4] Lastly, the application of the T7 BCD systems should be extended to the expression of other proteins to evaluate its expression performance comprehensively.

In conclusion, this study successfully establishes a bicistronic T7 expression system for enhancing recombinant protein production in *E. coli*, as a proof of concept, this system led to hyper-production of PCP subunit

vaccine proteins Oml1, Oml7, and ApxII. The unprecedented yields of these proteins were achieved through the combination of the T7 BCD system and fermentation condition optimization in MWP, with final yields of 2.43 g/L for Oml1, 2.59 g/L for Oml7, and 1.21 g/L for ApxII in a 5-L bioreactor. This study provided a powerful protein expression system for the enhanced production of recombinant proteins and overexpression of key enzymes in *E. coli*. Additionally, it promotes the development of subunit vaccines against PCP and provides an excellent strain for the large-scale production of PCP vaccine antigens.

Supplementary Information

Supplementary material 1 Gene sequence of Oml1, Oml7, and Apx II. **Table S1** Bacteria strains and plasmids used in this study. **Table S2** Primers used in this study.

Funding

This work was supported by National Natural Science Foundation of China (No.22078128, 21938004, 32001620), the 111 Project (No. 111-2-06), the Fundamental Research Funds for the Central Universities (No. JUSRP221032), Postgraduate Research & Practice Innovation Program of Jiangsu Province KYCX21-2028, National First-Class Discipline Program of Light Industry Technology and Engineering (LITE2018-24), the Opening Project of the Key Laboratory of Industrial Biotechnology, Ministry of Education (KLIBKF201802), British Council 527429894 and BBSRC BB/T013176/1.

Author's contribution

M.S., X.G, and A.L. performed all experiments and wrote this manuscript with assistance from X. L. and R.L-A.. P.W., W.C., and Z.B. revised this manuscript.

1
2
3
4 342
5
6 343 **Statements and Declarations**
7
8

9 344 **Ethics approval** All procedures for handling the mice used in this study were reviewed and approved by the
10
11
12 345 Institute of Tecon Biology Co., Ltd (SYXK2021-0006) and performed ethically and humanely.
13

14 346 **Conflict of interest** The authors declare no competing interests.
15
16

17 347 **Consent for publication** All authors approved the consent for publishing the manuscript to Biotechnology Journal.
18
19

20 348
21
22 349 **Data availability**
23
24

25 350 All data in this study are included in this published article and its supplementary files.
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

References

1. Bossé, J. T., Janson, H., Sheehan, B. J., Beddek, A. J., Rycroft, A. N., Simon Kroll, J., & Langford, P. R. (2002). *Actinobacillus pleuropneumoniae*: pathobiology and pathogenesis of infection. *Microbes and Infection*, 4(2), 225–235. [https://doi.org/10.1016/S1286-4579\(01\)01534-9](https://doi.org/10.1016/S1286-4579(01)01534-9)
2. Haesebrouck, F., Chiers, K., Van Overbeke, I., & Ducatelle, R. (1997). *Actinobacillus pleuropneumoniae* infections in pigs: the role of virulence factors in pathogenesis and protection. *Veterinary Microbiology*, 58(4), 239–249. [https://doi.org/10.1016/S0378-1135\(97\)00162-4](https://doi.org/10.1016/S0378-1135(97)00162-4)
3. Archambault, M., Harel, J., Gouré, J., Tremblay, Y. D. N., & Jacques, M. (2012). Antimicrobial susceptibilities and resistance genes of Canadian isolates of *Actinobacillus pleuropneumoniae*. *Microbial Drug Resistance*, 18(2), 198–206. <https://doi.org/10.1089/mdr.2011.0150>
4. Loera-Muro, A., & Angulo, C. (2018). New trends in innovative vaccine development against *Actinobacillus pleuropneumoniae*. *Veterinary Microbiology*, 217, 66–75. <https://doi.org/10.1016/j.vetmic.2018.02.028>
5. Ramjeet, M., Deslandes, V., Gouré, J., & Jacques, M. (2008). *Actinobacillus pleuropneumoniae* vaccines: from bacterins to new insights into vaccination strategies. *Animal Health Research Reviews*, 9(1), 25–45. <https://doi.org/10.1017/S1466252307001338>
6. Frey, J. (1995). Virulence in *Actinobacillus pleuropneumoniae* and RTX toxins. *Trends in Microbiology*, 3(7), 257–261. [https://doi.org/10.1016/S0966-842X\(00\)88939-8](https://doi.org/10.1016/S0966-842X(00)88939-8)
7. Tascón, R. I., Vázquez-Boland, J. A., Gutiérrez-Martín, C. B., Rodríguez-Barbosa, I., & Rodríguez-Ferri, E. Z. F. (1994). The RTX haemolysins ApxI and ApxII are major virulence factors of the swine pathogen *Actinobacillus pleuropneumoniae*: evidence from mutational analysis. *Molecular Microbiology*, 14(2), 207–216. <https://doi.org/10.1111/j.1365-2958.1994.tb01282.x>
8. Gerlach, G. F., Anderson, C., Klashinsky, S., Rossi-Campos, A., Potter, A. A., & Willson, P. J. (1993).

- 373 Molecular characterization of a protective outer membrane lipoprotein (OmlA) from *Actinobacillus*
374 *pleuropneumoniae* serotype 1. *Infection and Immunity*, 61(2), 565–572. [https://doi.org/10.1128/iai.61.2.565-](https://doi.org/10.1128/iai.61.2.565-572.1993)
375 572.1993
- 376 9. Bunka, S., Christensen, C., Potter, A. A., Willson, P. J., & Gerlach, G. F. (1995). Cloning and characterization
377 of a protective outer membrane lipoprotein of *Actinobacillus pleuropneumoniae* serotype 5. *Infection and*
378 *Immunity*, 63(7), 2797–2800. <https://doi.org/10.1128/iai.63.7.2797-2800.1995>
- 379 10. Seo, K.-W., Kim, D.-H., Kim, A. H., Yoo, H.-S., Lee, K.-Y., & Jang, Y.-S. (2011). Characterization of
380 antigenic determinants in ApxIIA exotoxin capable of inducing protective immunity to *Actinobacillus*
381 *pleuropneumoniae* challenge. *Immunological Investigations*, 40(5), 465–480.
382 <https://doi.org/10.3109/08820139.2011.558151>
- 383 11. Shin, M. K., Kang, M. L., Jung, M. H., Cha, S. Bin, Lee, W. J., Kim, J. M., Kim, D. H., & Yoo, H. S. (2013).
384 Induction of protective immune responses against challenge of *Actinobacillus pleuropneumoniae* by oral
385 administration with *Saccharomyces cerevisiae* expressing Apx toxins in pigs. *Veterinary Immunology and*
386 *Immunopathology*, 151(2), 132–139. <https://doi.org/10.1016/j.vetimm.2012.11.003>
- 387 12. Sun, M., Gao, A. X., Ledesma-Amaro, R., Li, A., Wang, R., Nie, J., Zheng, P., Yang, Y., Bai, Z., & Liu, X.
388 (2022). Hypersecretion of OmlA antigen in *Corynebacterium glutamicum* through high-throughput based
389 development process. *Applied Microbiology and Biotechnology*, 106(8), 2953–2967.
390 <https://doi.org/10.1007/s00253-022-11918-x>
- 391 13. Hemamalini, N., Ezhilmathi, S., & Mercy, A. A. (2020). Recombinant protein expression optimization in
392 *Escherichia coli*: A review. *Indian Journal of Animal Research*, 54(6), 653–660.
393 <https://doi.org/10.18805/ijar.B-3808>
- 394 14. Kaur, J., Kumar, A., & Kaur, J. (2018). Strategies for optimization of heterologous protein expression in *E.*

- 395 *coli*: Roadblocks and reinforcements. *International Journal of Biological Macromolecules*, 106, 803–822.
- 396 <https://doi.org/10.1016/j.ijbiomac.2017.08.080>
- 397 15. Ahmad, I., Nawaz, N., Darwesh, N. M., ur Rahman, S., Mustafa, M. Z., Khan, S. B., & Patching, S. G. (2018).
- 398 Overcoming challenges for amplified expression of recombinant proteins using *Escherichia coli*. *Protein*
- 399 *Expression and Purification*, 144, 12–18. <https://doi.org/10.1016/j.pep.2017.11.005>
- 400 16. Rosano, G. L., & Ceccarelli, E. A. (2014). Recombinant protein expression in *Escherichia coli*: advances and
- 401 challenges. *Frontiers in Microbiology*, 5, 172. <https://doi.org/10.3389/fmicb.2014.00172>
- 402 17. Sørensen, H. P., & Mortensen, K. K. (2005). Advanced genetic strategies for recombinant protein expression
- 403 in *Escherichia coli*. *Journal of Biotechnology*, 115(2), 113–128.
- 404 <https://doi.org/10.1016/j.jbiotec.2004.08.004>
- 405 18. Seo, S. W., Yang, J. S., Kim, I., Yang, J., Min, B. E., Kim, S., & Jung, G. Y. (2013). Predictive design of
- 406 mRNA translation initiation region to control prokaryotic translation efficiency. *Metabolic Engineering*,
- 407 15(1), 67–74. <https://doi.org/10.1016/j.ymben.2012.10.006>
- 408 19. Sun, M., Gao, X., Zhao, Z., Li, A., Wang, Y., Yang, Y., Liu, X., & Bai, Z. (2020). Enhanced production of
- 409 recombinant proteins in *Corynebacterium glutamicum* by constructing a bicistronic gene expression system.
- 410 *Microbial Cell Factories*, 19(1), 113. <https://doi.org/10.1186/s12934-020-01370-9>
- 411 20. Sun, M., Gao, A. X., Li, A., Liu, X., Wang, R., Yang, Y., Li, Y., Liu, C., & Bai, Z. (2021). Bicistronic design
- 412 as recombinant expression enhancer: characteristics, applications, and structural optimization. *Applied*
- 413 *Microbiology and Biotechnology*, 105(20), 7709–7720. <https://doi.org/10.1007/s00253-021-11611-5>
- 414 21. Takyar, S., Hickerson, R. P., & Noller, H. F. (2005). mRNA helicase activity of the ribosome. *Cell*, 120(1),
- 415 49–58. <https://doi.org/10.1016/j.cell.2004.11.042>
- 416 22. Liu, X., Zhao, Z., Zhang, W., Sun, Y., Yang, Y., & Bai, Z. (2017). Bicistronic expression strategy for high-

- level expression of recombinant proteins in *Corynebacterium glutamicum*. *Engineering in Life Sciences*, 17(10), 1118–1125. <https://doi.org/10.1002/elsc.201700087>
23. Mukhopadhyay, U. K., & Sahni, G. (2002). An insight into the possible mechanism of working of two-cistronic gene expression systems and rational designing of newer systems. *Journal of Biosciences*, 27(3), 219–231. <https://doi.org/10.1007/BF02704911>
24. Zhao, Z., Liu, X., Zhang, W., Yang, Y., Dai, X., & Bai, Z. (2016). Construction of genetic parts from the *Corynebacterium glutamicum* genome with high expression activities. *Biotechnology Letters*, 38(12), 2119–2126. <https://doi.org/10.1007/s10529-016-2196-y>
25. Mutalik, V. K., Guimaraes, J. C., Cambray, G., Lam, C., Christoffersen, M. J., Mai, Q. A., Tran, A. B., Paull, M., Keasling, J. D., Arkin, A. P., & Endy, D. (2013). Precise and reliable gene expression via standard transcription and translation initiation elements. *Nature Methods*, 10(4), 354–360. <https://doi.org/10.1038/nmeth.2404>
26. Baltes, N., Hennig-Pauka, I., & Gerlach, G.-F. (2002). Both transferrin binding proteins are virulence factors in *Actinobacillus pleuropneumoniae* serotype 7 infection. *FEMS Microbiology Letters*, 209(2), 283–287. <https://doi.org/10.1111/j.1574-6968.2002.tb11145.x>
27. Sohoni, S. V., Nelapati, D., Sathe, S., Javadekar-Subhedar, V., Gaikawai, R. P., & Wangikar, P. P. (2015). Optimization of high cell density fermentation process for recombinant nitrilase production in *E. coli*. *Bioresource Technology*, 188, 202–208. <https://doi.org/10.1016/j.biortech.2015.02.038>
28. Han, S. O., Inui, M., & Yukawa, H. (2008). Effect of carbon source availability and growth phase on expression of *Corynebacterium glutamicum* genes involved in the tricarboxylic acid cycle and glyoxylate bypass. *Microbiology*, 154(10), 3073–3083. <https://doi.org/10.1099/mic.0.2008/019828-0>
29. Islam, R. S., Tisi, D., Levy, M. S., & Lye, G. J. (2007). Framework for the rapid optimization of soluble protein

- expression in *Escherichia coli* combining microscale experiments and statistical experimental design. *Biotechnology Progress*, 23(4), 785–793. <https://doi.org/10.1021/bp070059a>
30. Marx, C. K., Hertel, T. C., & Pietzsch, M. (2007). Soluble expression of a pro-transglutaminase from *Streptomyces mobaraensis* in *Escherichia coli*. *Enzyme and Microbial Technology*, 40(6), 1543–1550. <https://doi.org/10.1016/j.enzmictec.2006.10.036>
31. Semba, H., Ichige, E., Imanaka, T., Atomi, H., & Aoyagi, H. (2008). Efficient production of active form of recombinant cassava hydroxynitrile lyase using *Escherichia coli* in low-temperature culture. *Applied Microbiology and Biotechnology*, 79(4), 563–569. <https://doi.org/10.1007/s00253-008-1464-8>
32. Zhang, W., Yang, Y., Liu, X., Liu, C., & Bai, Z. (2019). Development of a secretory expression system with high compatibility between expression elements and an optimized host for endoxylanase production in *Corynebacterium glutamicum*. *Microbial Cell Factories*, 18(1), 72. <https://doi.org/10.1186/s12934-019-1116-y>
33. Hunt, I. (2005). From gene to protein: a review of new and enabling technologies for multi-parallel protein expression. *Protein Expr. Purif.*, 40(1), 1–22. <https://doi.org/10.1016/j.pep.2004.10.018>
34. Freudl, R. (2017). Beyond amino acids: Use of the *Corynebacterium glutamicum* cell factory for the secretion of heterologous proteins. *Journal of Biotechnology*, 258, 101–109. <https://doi.org/10.1016/j.jbiotec.2017.02.023>
35. Sharifi-Rad, J., Soufi, L., Hoseini-Alfatemi, S. M., Sharifi-Rad, M., Iriti, M., Sharifi-Rad, M., & Hoseini, M. (2015). Recombinant Proteins in *Escherichia coli*: Optimizing Production Strategies. *American-Eurasian Journal of Agricultural & Environment*, 15(11), 2149–2159. <https://doi.org/10.5829/idosi.aejaes.2015.15.11.96100>
36. Sivashanmugam, A., Murray, V., Cui, C., Zhang, Y., Wang, J., & Li, Q. (2009). Practical protocols for

- production of very high yields of recombinant proteins using *Escherichia coli*. *Protein Science*, 18(5), 936–948. <https://doi.org/10.1002/pro.102>
37. Kudla, G., Murray, A. W., Tollervey, D., & Plotkin, J. B. (2009). Coding-sequence determinants of gene expression in *Escherichia coli*. *Science*, 324(5924), 255–258. <https://doi.org/10.1126/science.1170160>
38. Smit, M. H., & van Duin, J. (1990). Secondary structure of the ribosome binding site determines translational efficiency: a quantitative analysis. *Proc Natl Acad Sci USA*, 87(19), 7668–7672. <https://doi.org/10.1073/pnas.87.19.7668>
39. Kruth, S., Schibajew, L., & Nett, M. (2022). Biocatalytic production of the antibiotic aurachin D in *Escherichia coli*. *AMB Express*, 12(1). <https://doi.org/10.1186/s13568-022-01478-8>
40. Claassens, N. J., Finger-Bou, M., Scholten, B., Muis, F., De Groot, J. J., De Gier, J. W., De Vos, W. M., & Van Der Oost, J. (2019). Bicistronic design-based continuous and high-level membrane protein production in *Escherichia coli*. *ACS Synthetic Biology*, 8(7), 1685–1690. <https://doi.org/10.1021/acssynbio.9b00101>
41. Marino, J., Hohl, M., Seeger, M. A., Zerbe, O., & Geertsma, E. R. (2015). Bicistronic mRNAs to enhance membrane protein overexpression. *Journal of Molecular Biology*, 427(4), 943–954. <https://doi.org/10.1016/j.jmb.2014.11.002>
42. Liu, X., Sun, M., Gao, A. X., Ledesma-Amaro, R., Fang, Q., Yang, Y., & Bai, Z. (2023). Leaderless bicistronic design for precise and reliable control of gene expression in *Corynebacterium glutamicum*. *ACS Synthetic Biology*. <https://doi.org/10.1021/acssynbio.3c00246>
43. Kelley, D. R., Snoek, J., & Rinn, J. L. (2016). Basset: learning the regulatory code of the accessible genome with deep convolutional neural networks. *Genome Research*, 26(7), 990–999. <https://doi.org/10.1101/gr.200535.115>
44. Duan, Y., Zhai, W., Liu, W., Zhang, X., Shi, J. S., Zhang, X., & Xu, Z. (2021). Fine-tuning multi-gene clusters

via well-characterized gene expression regulatory elements: case study of the arginine synthesis pathway in

C. glutamicum. *ACS Synthetic Biology*, 10(1), 38–48. <https://doi.org/10.1021/acssynbio.0c00405>

45. Jia, L., Mao, Y., Ji, Q., Dersh, D., Yewdell, J. W., & Qian, S. B. (2020). Decoding mRNA translatability and stability from the 5' UTR. *Nature Structural and Molecular Biology*, 27(9), 814–821. <https://doi.org/10.1038/s41594-020-0465-x>

Figure captions

Fig. 1 Construction of bicistronic T7 expression structure

(A) Monocistronic and bicistronic T7 expression structure. The bicistronic T7 expression system contains T7 promoter, 5'UTR and its original SD motif (SD1), fore-cistron sequence, the second SD sequence (SD2), and downstream target gene EGFP. (B) Construction process of bicistronic T7 expression plasmid.

Fig. 2 Characterization of effects of different fore-cistron sequences on the performance of bicistronic T7 expression system

C (control) represents the monocistronic expression vector pET28-EGFP. The abscissa represents the BCD vector containing the corresponding fore-cistron. (A) EGFP expression level of the BCD expression vectors. (B) The relative transcriptional level and relative translation efficiency of EGFP in bicistronic T7 expression system.

Fig. 3 Expression level of PCP vaccine antigen in four bicistronic T7 expression vectors.

Line M: Protein marker, Line 1: BL21 WT strain, Line 2: pET28 without target gene, Line 3-7 represented the target protein expression of vector pET28a, pET28-HT5, pET28-HT8, pET28-HP11Y, and pET28-HP12Y,

respectively. (A), (B), and (C) represented the SDS-PAGE analysis of Oml1, Oml7, and Apx II , respectively. Approximately 5 μ L sample was loaded onto each lane. (D), (E), and (F) represented the relative protein yield of Oml1, Oml7 and Apx II in BCD vectors per liter of cells, respectively. *represents $P<0.05$, ** represents $P<0.01$.

Fig. 4 Effects of cultivation temperature and induction conditions on the production of PCP vaccine proteins.

For the optimization experiment of cultivation temperature, the OD₆₀₀ and protein yield at 37 °C were used as the control, defined as 1. For the optimization experiment of induction conditions, the OD₆₀₀ and protein yield with 1.0 mM IPTG addition and 0 h pre-induction period was used as the control, defined as 1. *represents $P<0.05$, ** represents $P<0.01$. (A) Effect of cultivation temperature on Oml1 production. (B) Effect of cultivation temperature on Oml7 production. (C) Effect of cultivation temperature on Apx II production. (D) Effect of IPTG concentration on Oml1 production. (E) Effect of IPTG concentration on Oml7 production. (F) Effect of IPTG concentration on Apx II production. (G) Effect of pre-induction period on Oml1 production. (H) Effect of pre-induction period on Oml7 production. (I) Effect of pre-induction period on Apx II production.

Fig. 5 Effects of medium components on the production of PCP vaccine proteins.

For the optimization of carbon source, the OD₆₀₀ and protein yield under glucose carbon source were used as the control, defined as 1. For the optimization of nitrogen source, the OD₆₀₀ and protein yield in the original TBSB medium were used as the control, defined as 1. *represents $P<0.05$, ** represents $P<0.01$. (A) Effect of carbon source on Oml1 production. (B) Effect of carbon source on Oml7 production. (C) Effect of carbon source on Apx II production. (D) Effect of nitrogen source on Oml1 production. (E) Effect of nitrogen source on Oml7

527 production. (F) Effect of nitrogen source on Apx II production.

528

529

530 **Fig. 6 High-level production of Oml1, Oml7, and Apx II in 5 L bioreactor.**

531 (A), (B), and (C) represented the time profile of cell growth, DO level, and protein concentration during

532 fermentation. Arrows represented IPTG addition. (D), (E), and (F) represented SDS-PAGE analysis of purified

533 Oml1, Oml7 and Apx II , respectively. (G) Induction of protective immunity against challenge infection with APP7.

534

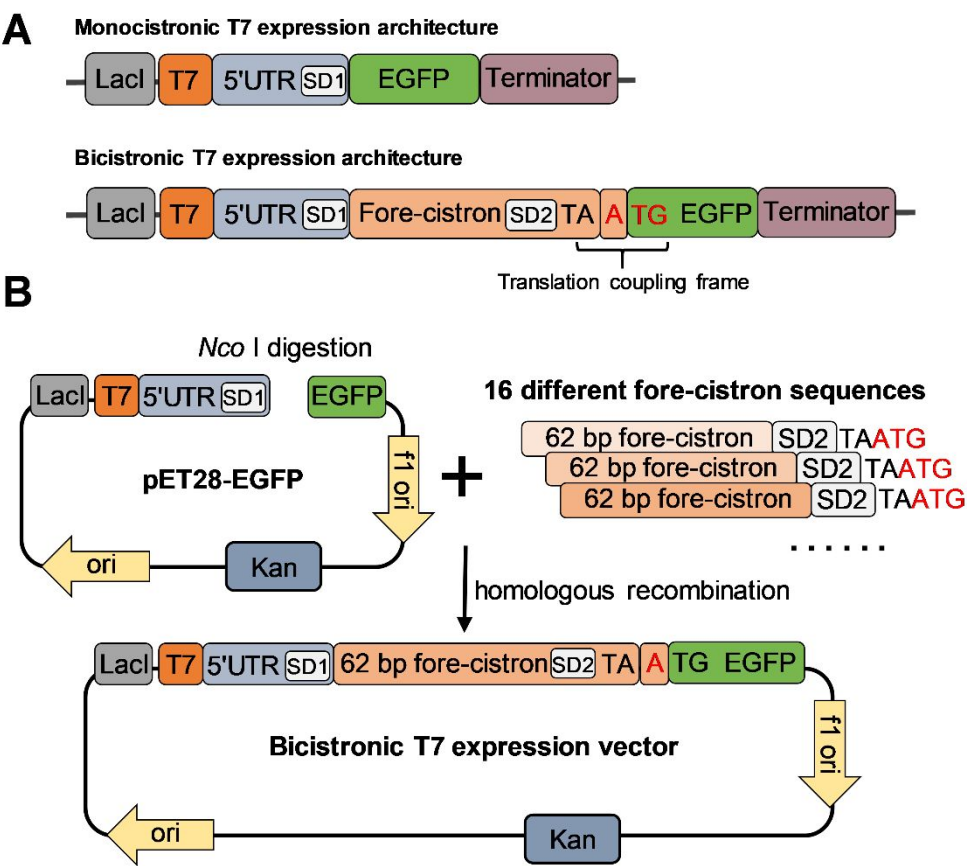


Fig. 1

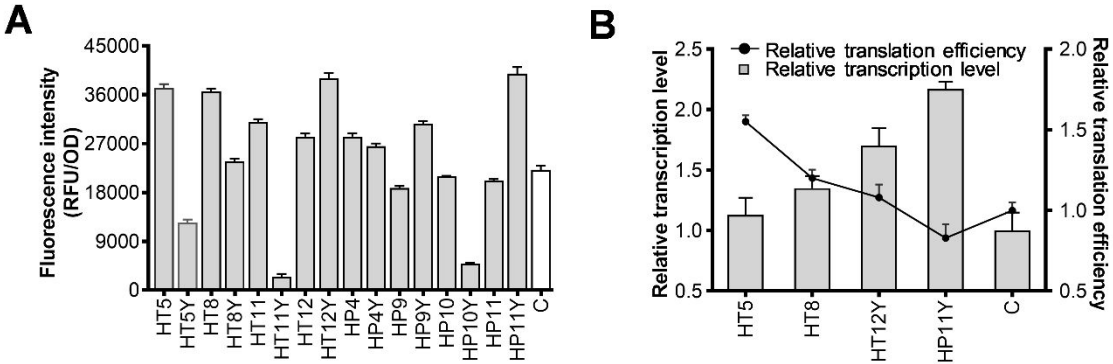


Fig. 2

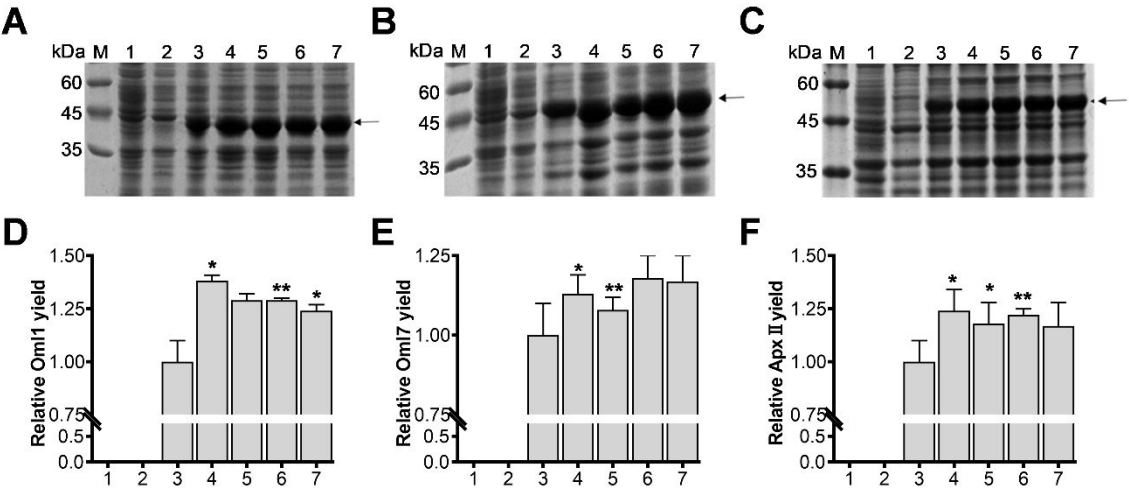


Fig. 3

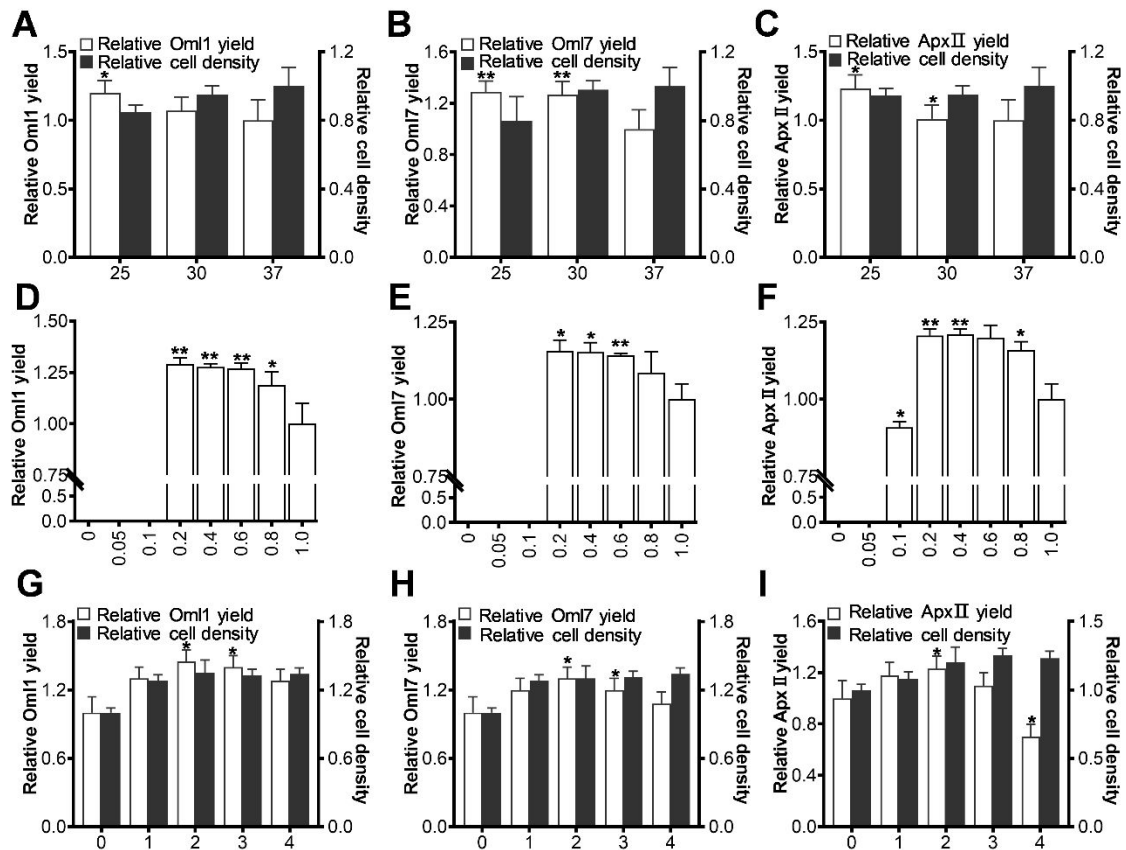


Fig. 4

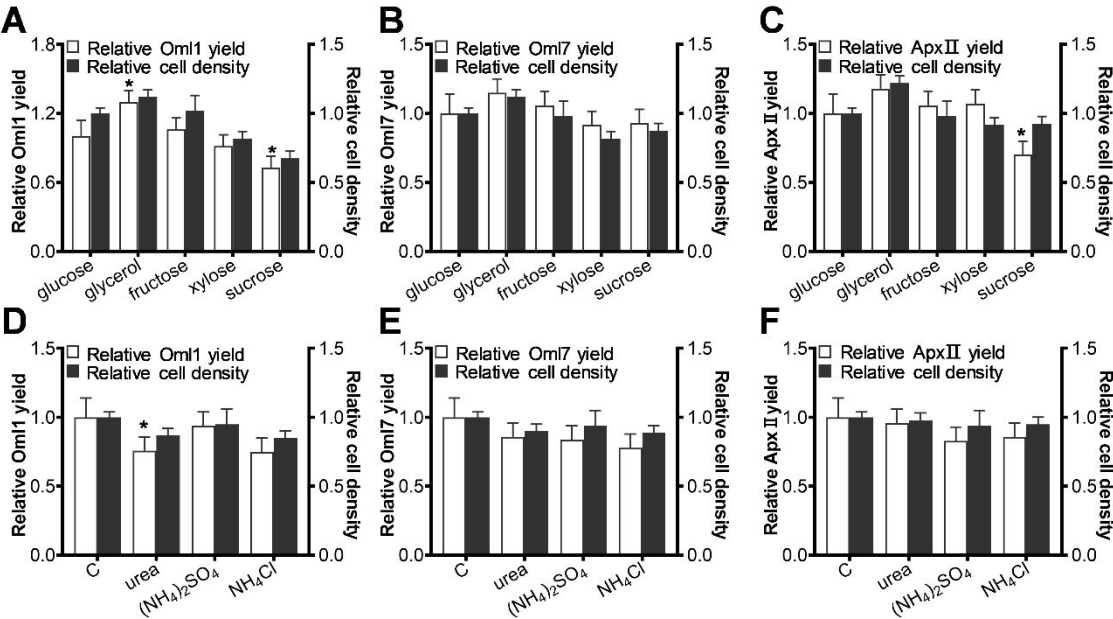


Fig. 5

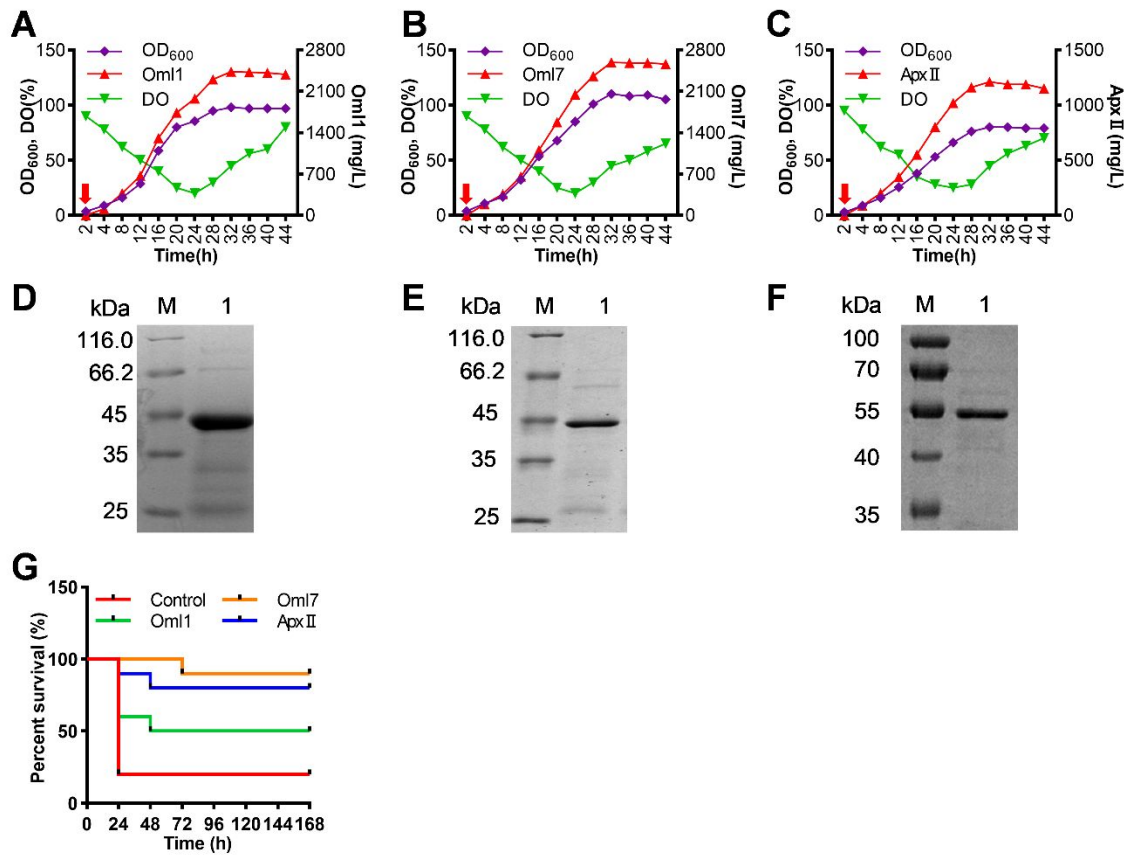


Fig. 6

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

553 **Table 1** The sequence information of fore-cistron

Number	The fore-cistron sequence (5'-3')	Sources
HT5	gtgggagatgtgtaaaaggcaacgacgcgcacacggagacggtgatacgcgccgaaaaat	NCgl1504
HT5Y	gtgggtgatgtggtgaaaggcaatgatgccatactggtgatggtgatactagaagaaaaat	HT5 codon optimization
HT8	atgtacgcgcatcgtcaagaccggcggaagcagtagaaggtgccgaaggtgacctcgtaa	NCgl2280
HT8Y	atgtatgccattgtgaaaactggtggcaaacagtataaagtggcagaaggtgatctggtgaa	HT8 codon optimization
HT11	atgactcgacgtctacatggtggtgagcaggatggccaggaacacgttaaaggacagctaaa	NCgl0576
HT11Y	atgactagaagactgcatggtggtgaacaagatggccaagaacatgtgaaaggtcagctgaa	HT11 codon optimization
HT12	gtgtcgctttcagagcaggagcaacgcgcactgcgcgaaattgagcaagcgctcatggcaga	NCgl2088
HT12Y	gtgagcctgagtgaacaagaacagagagccctgagagaaattgaacaagccctgatggcaga	HT12 codon optimization
HP4	atggggtccatgggctaaaacacatttcaaggcaacgaaactgctacctccggcgaaactgcc	NCgl1041
HP4Y	atgggcagcatggccaaaaccatttcaaggcaatgaaactgccacctctggtgaactgcc	HP4 codon optimization
HP9	atgagcgatattcgatggcagcccagggtggcgctggttcggaaatgacgtctttgatcg	NCgl2328
HP9Y	atgagtgatattagaatggcagcccaagggtggccctggctttggcaatgatgtgtttgatcg	HP9 codon optimization
HP10	atgcgactcgactcctcggacctcccggtgctggtgaaggcaccaggtgcaattctctc	NCgl0533
HP10Y	atgagactggtgctgctggggccacctggtgctggcgaaggcaccgaagcagccattctgtc	HP10 codon optimization
HP11	atgattggagcaccacccgacatgggcaatgtgtacaacaacatcaccgaaaccatcggccca	NCgl2473
HP11Y	atgattggtgccccacctgatatgggcaatgtgtataacaacattactgaaaccattggtca	HP11 codon optimization

554

Hyper-production of porcine contagious pleuropneumonia subunit vaccine proteins in *Escherichia coli* by developing a bicistronic T7 expression system

Manman Sun^{1,2,3,#}, Alex Xiong Gao^{4,#}, An Li^{1,#}, Rodrigo Ledesma-Amaro², Peng Wang³, Wenchao Chen⁵,
Zhonghu Bai¹, Xiuxia Liu^{1*}

[#] Co-first author

* Correspondence: liuxiuxia@jiangnan.edu.cn

¹ National Engineering Research Center of Cereal Fermentation and Food Biomanufacturing, Jiangnan University, Wuxi 214112, China

² Department of Bioengineering and Imperial College Centre for Synthetic Biology, Imperial College London, London SW7 2AZ, UK

³ Key laboratory of high magnetic field and Ion beam physical biology, Hefei Institutes of Physical Science, Chinese Academy of Sciences

⁴ Division of Life Science, The Hong Kong University of Science and Technology, Hong Kong, China

⁵ Oil Crops Research Institute of the Chinese Academy of Agricultural Sciences, Wuhan, 430062, China

Supplementary material 1 Gene sequence of Oml1, Oml7, and ApxII.

Sequence 1: the codon-optimized Oml1 sequence, the green sequence represents the 6×his tag.

atgaccccgaagtgatagcgccgaaagcggaacagccgaaaaagaagtgccgaagcggataacagcaaggcggaa
gaaccgaaagagatggcgccacaagtgatagcccgaagcagaagagccgaagaacatggccccgcagatgggcaaccgaaactg
aacgatccgcaagtgtgcccgaatggataacccgcagaagatgcccgaagggcgaagaactgagcaagataaaagcaacg
cggaattctgaaagaactggcggtgaaagataaataacagcggcattattaacaacgcggatgtggtgctgaacctgaaattgatgaaaaag
atcatattaccgtggtgctggataaagataaaataacgcacatctgaaagtgaacacaccattagcgcgcaagatataaaacctgaa
agatagcagcggcaaacgtggtgctattatggtatagcagctgaaccaagtgcgccaagatgaaactatagcagatgaaaaagttagcc
tgaacgaatattatctgctgagcatgaacgatgcggataaaatcgccgacaaaaagcattagctataaaggcgatattttatagctataaa
gatgtgggcaatcagaactgaaagcgagcgtggaagcgagctatgatgtgacaaaaagtgaacatgaaagtgtttggcgaaaaca
acgattattgaaactggcggaattggccgaccaacctgtgaaaaccaagtaccggcggaagtgggcggaagatggcaccattatt
aacggcacctgtatagcaaaattgataactttccgtgaaactgacccggatgcgaacttttagcggcgcatTTTTGGCGAAAAACGGCGAA
GTGCTGGCGGGCAGCGCGATTAGCGAAAAATGGCAAGGCGTGATTGGCGCGACCGCGACCACCAAGAAGATAAACACCACCACCACCA
CCACTAA

Sequence 2: the codon-optimized Oml7 sequence, the green sequence represents the 6×his tag.

atgaccctaagtagatagcgccggaaggtgaacagccgaaaaagaagaagcgccgaagcggatagtcctaaagctgagaa
accgaaaagcattgcgcgctgatgtgaaaacccgaaggtggaacacagaagaaacacacctgcaagaaaaatctccgaaagcgg
atgaaccgaagtgtgacccgaaattaggcgccgcgagaaagatgacgaaactggaagaaccgaaaaacaaagcaacgcggaa
attctgaaagaactgggcattaaagataataaaacccgcattattaccgcagcgtgtggtgctgaacctgacctggtgaacaagaaac
attcagattgcctgagcgaagcgatattgtgcgaacgatctgaaaattaccaacaccattccgaaccaagatattgcacctgaaagata
gcaccggccgctgctgggtattatggctatagcagctgaaccaagtgcgcgaagggcaacgctatggcattaacaacgtggtatctgtg
ggccattatctgctgagcatggatgaaagcaccaaaaccgcgcgaacaaagcattgaatatcgcggcaaaatgctgtatggctataaaaa
cgtggataaccgcaacctgtggcggtgtgcaagcgagctataaccatagcgataaaaaactgagcatggaaattttggcgatcatggcg
attattgaaactggcgcgattggcaacaaccgctgccgaaagatatgtgaccggcggtggtggtgataaagatggcaccattagcaac
ggcgccctgtatagcaaaattgataacccccgggcaaacctgacccggatgcgaccttttagcggcgcatTTTTGGCGAAAAACGGCGATGT
GCTGGCGGGCAGCGCGGATGGCAAAAACGGCAAGGCGTGATTGGCGCGACCGCGACCACCAAGAAGATAAACACCACCACCACCAC
CACTAA

Sequence 3: the codon-optimized truncated ApxII sequence, the green sequence represents the 6×his tag.

Atcaaggctatgatagccgccatctggcgatctgaagataacatgaaatttctgattaacctgaacaaagaactgaagcggaaacgct
ggtggcgattacgcagcagcgtgggataatcagattggcgatctggcagctattagccgccaccgataaaattagcagcggcaaacgc
tatgtggatgcgtttgaagaaggtcagcatcagagctatgatagcagcgtgacgtggataacaaaaacggcattattaacattagcaacacca
accgcaaacgcagagcgtgtgttcgacccccgtgtgacccggcggaagaaaccggaacgcattcaagaaggcaaaaacagc
tatattaccaaacgcatattcagcgcgtgtagctggaccgtgaccgatggcgacgcgagcagcagcgtgattttaccaacgtggtgcag
cgattgcggtgaaatttgatgatggggcaacattattgaaagcaaaataacaaaattattgcgaacctggcgcgggcaacgataacgtg
tttgtggcgagcagcaccaccgttattgacggtggcgacggcgatgacgctgacattatagccggcggaatattggcgcgctggtgattgat
gcgaccgcggaaccgaaaaaggcagctatagcgtgaaacgctatgtggcgatagcaaaagcgtgcatgaaccattgacacccatcag
accaacgtgggcaaccgcgaagaaaaaattgaatatcgccggaagatgatcgctttcataccggctataccgtgaccgacagcctgaaaa
gctgtgaggaaattattggagccaattcaacgatattttaaaggcagtcagttcgacgatgtgtttcatggcggaatggcggtgataacatt
gatggcaacgacggcgatgatcatctgtttggcgcgggcgatgatgtgattgacggtgtaacggcaacaacttttgggtggcggcac
cggcaacgatattattagtgccggaagataacgatattatgtcataaaaccggcgacggcaacgacagcattaccACCACCACCACC
ACCCTAA

Table S1 Bacteria strains and plasmids used in this study.

Strains or plasmids	Description	Sources
Strains		
<i>E. coli</i> DH5 α	\	Lab stock
<i>E. coli</i> BL21	\	Lab stock
APP7	\	Tecon Biology Co. , Ltd
Plasmids		
pET28a (+)	<i>E. coli</i> expression vector, kanamycin resistance	Lab stock
pET28-EGFP	pET28a, carrying monocistronic T7 promoter and EGFP gene	This study
pET28-Oml1	pET28a, carrying monocistronic T7 promoter and Oml1 gene	This study
pET28-Oml7	pET28a, carrying monocistronic T7 promoter and Oml7 gene	This study
pET28-ApxII	pET28a, carrying monocistronic T7 promoter and ApxII gene	This study
pET28-HT5-EGFP	Bicistronic T7 system, carrying HT5 fore-cistron and EGFP gene	This study
pET28-HT5Y-EGFP	Bicistronic T7 system, carrying HT5Y fore-cistron and EGFP gene	This study
pET28-HT8-EGFP	Bicistronic T7 system, carrying HT8 fore-cistron and EGFP gene	This study
pET28-HT8Y-EGFP	Bicistronic T7 system, carrying HT8Y fore-cistron and EGFP gene	This study
pET28-HT11-EGFP	Bicistronic T7 system, carrying HT11 fore-cistron and EGFP gene	This study
pET28-HT11Y-EGFP	Bicistronic T7 system, carrying HT11Y fore-cistron and EGFP gene	This study
pET28-HT12-EGFP	Bicistronic T7 system, carrying HT12 fore-cistron and EGFP gene	This study
pET28-HT12Y-EGFP	Bicistronic T7 system, carrying HT12Y fore-cistron and EGFP gene	This study
pET28-HP4-EGFP	Bicistronic T7 system, carrying HP4 fore-cistron and EGFP gene	This study
pET28-HP4Y-EGFP	Bicistronic T7 system, carrying HP4Y fore-cistron and EGFP gene	This study
pET28-HP9-EGFP	Bicistronic T7 system, carrying HP9 fore-cistron and EGFP gene	This study
pET28-HP9Y-EGFP	Bicistronic T7 system, carrying HP9Y fore-cistron and EGFP gene	This study
pET28-HP10-EGFP	Bicistronic T7 system, carrying HP10 fore-cistron and EGFP gene	This study
pET28-HP10Y-EGFP	Bicistronic T7 system, carrying HP10Y fore-cistron and EGFP gene	This study
pET28-HP11-EGFP	Bicistronic T7 system, carrying HP11 fore-cistron and EGFP gene	This study
pET28-HP11Y-EGFP	Bicistronic T7 system, carrying HP11Y fore-cistron and EGFP gene	This study
pET28-HT5-Oml1	Bicistronic T7 system, carrying HT5 fore-cistron and Oml1 gene	This study
pET28-HT8-Oml1	Bicistronic T7 system, carrying HT8 fore-cistron and Oml1 gene	This study
pET28-HT12Y-Oml1	Bicistronic T7 system, carrying HT12Y fore-cistron and Oml1 gene	This study
pET28-HP11Y-Oml1	Bicistronic T7 system, carrying HP11Y fore-cistron and Oml1 gene	This study
pET28-HT5-Oml7	Bicistronic T7 system, carrying HT5 fore-cistron and Oml7 gene	This study
pET28-HT8-Oml7	Bicistronic T7 system, carrying HT8 fore-cistron and Oml7 gene	This study
pET28-HT12Y-Oml7	Bicistronic T7 system, carrying HT12Y fore-cistron and Oml7 gene	This study
pET28-HP11Y-Oml7	Bicistronic T7 system, carrying HP11Y fore-cistron and Oml7 gene	This study
pET28-HT5-ApxII	Bicistronic T7 system, carrying HT5 fore-cistron and ApxII gene	This study
pET28-HT8-ApxII	Bicistronic T7 system, carrying HT8 fore-cistron and ApxII gene	This study
pET28-HT12Y-ApxII	Bicistronic T7 system, carrying HT12Y fore-cistron and ApxII gene	This study
pET28-HP11Y-ApxII	Bicistronic T7 system, carrying HP11Y fore-cistron and ApxII gene	This study

Table S2 Primers used in this study.

Primer	Sequence (5'-3') ^a	Intension
HT5-F	aactttaagaaggagatatagtgggagatgttgtaaaaggcaacgacgcgcacaccgga gacggt	Amplifying HT5
HT5-R	tcctcgcccttgctcaccatcattagttgtcctctttatttttcggcgctatcaccgtctccggtg tgcgcgt	Amplifying HT5
HT5Y-F	aactttaagaaggagatatagtgggtgatgtggtgaaaggcaatgatgccatactggtga tggt	Amplifying HT5Y
HT5Y-R	tcctcgcccttgctcaccatcattagttgtcctctttatttttcttagtatcaccatcaccagtatg ggcat	Amplifying HT5Y
HT8-F	aactttaagaaggagatataatgtacgcgatcgtaagaccggcgaaagcagtacaag gttgcc	Amplifying HT8
HT8-R	tcctcgcccttgctcaccatcattagttgtcctcttttaacgaggtcaccttcggcaaccttgta ctgctttc	Amplifying HT8
HT8Y-F	aactttaagaaggagatataatgtatgccattgtgaaaactggtggcaaacagtataaagtg gca	Amplifying HT8Y
HT8Y-R	tcctcgcccttgctcaccatcattagttgtcctcttttcaccagatcaccttctgccactttatact gtttgc	Amplifying HT8Y
HT11-F	aactttaagaaggagatataatgactcgacgtctacatggtggtgagcaggatggccagg aacac	Amplifying HT11
HT11-R	tcctcgcccttgctcaccatcattagttgtcctcttttttagctgtcctttaacgtgttcctggccat cctgct	Amplifying HT11
HT11Y-F	aactttaagaaggagatataatgactagaagactgcatggtggtgaacaagatggccaag aacat	Amplifying HT11Y
HT11Y-R	tcctcgcccttgctcaccatcattagttgtcctcttttcagctgacctttcacatgttcttgccat ctgtt	Amplifying HT11Y
HT12-F	aactttaagaaggagatatagttcgctttcagagcaggagcaacgcgcactgcgcgaaa agtgcg	Amplifying HT12
HT12-R	tcctcgcccttgctcaccatcattagttgtcctcttttctgcatgagcgcttgctcaatttcgcgc agtgcg	Amplifying HT12
HT12Y-F	aactttaagaaggagatatagtgagcctgagtgaacaagaacagagagccctgagagaa attgaa	Amplifying HT12Y
HT12Y-R	tcctcgcccttgctcaccatcattagttgtcctcttttctgcatcagggtgttcaatttctctca gggctc	Amplifying HT12Y
HP4-F	aactttaagaaggagatataatgggtccatggctaaaacacattttcaaggcaacgaaac tgct	Amplifying HP4
HP4-R	tcctcgcccttgctcaccatcattagttgtcctctttggcagttcgcggaggtagcagtttcgtt gccttgaa	Amplifying HP4
HP4Y-F	aactttaagaaggagatataatgggcagcatggccaaaaccattttcaaggcaatgaaac tgcc	Amplifying HP4Y
HP4Y-R	tcctcgcccttgctcaccatcattagttgtcctctttggcagttcaccagaggtggcagtttcatt gccttgaa	Amplifying HP4Y
HP9-F	aactttaagaaggagatataatgagcgatattcgatggcagcccagggtgggcctggttt cgga	Amplifying HP9

HP9-R	tcctcgcccttgctcaccatcattagttgcctcctt	Amplifying HP9
	cgatcaaagacgtcatttcgaaccaggcccaccc	
HP9Y-F	aactttaagaaggagatataatgagtgatattagaatggcagcccaaggtggccctggcctt	Amplifying HP9Y
	ggc	
HP9Y-R	tcctcgcccttgctcaccatcattagttgcctccttcgatcaaacacatcattgccaagccag	Amplifying HP9Y
	ggccacctt	
HP10-F	aactttaagaaggagatataatgcgactcgctactcctcgacccctcccggtgctgtaagg	Amplifying HP10
	cacc	
HP10-R	tcctcgcccttgctcaccatcattagttgcctccttgagagaattgcagcctgggtgcccttac	Amplifying HP10
	cagcaccgg	
HP10Y-F	aactttaagaaggagatataatgagactgggtgctgctgggcccacctggtgctggcaaag	Amplifying HP10Y
	gcacc	
HP10Y-R	tcctcgcccttgctcaccatcattagttgcctccttgacagaatggctgctgggtgccttggc	Amplifying HP10Y
	agcaccag	
HP11-F	aactttaagaaggagatataatgattggagcaccaccgcacatgggcaatgtgtacaaca	Amplifying HP11
	acatc	
HP11-R	tcctcgcccttgctcaccatcattagttgcctcctttggccgatggttcggtgatgtgtgtac	Amplifying HP11
	acattgc	
HP11Y-F	aactttaagaaggagatataatgattgggtccccacctgatatgggcaatgtgtataacaac	Amplifying HP11Y
	att	
HP11Y-R	tcctcgcccttgctcaccatcattagttgcctcctttgaccaatggttcagtaattgtgtataca	Amplifying HP11Y
	cattgc	
EGFP-F	cgcccatgggcatggtgagcaaggcgaggagctg	Amplifying EGFP for pET28
EGFP-R	cgcaagcttttactgtacagctcgtcca	Amplifying EGFP for pET28
EGFP-qPCR-F	acaacatcgaggacggcag	qPCR for EGFP
EGFP-qPCR-F	gtccatgccgagagtgtacc	qPCR for EGFP
16s-qPCR-F	attgacgttaccgcagaagaa	qPCR for 16s
16s-qPCR-F	gcttgcaccctccgtattacc	qPCR for 16s

^a The red sequence represents homology arm, and the green sequence represents SD2 sequence.