VNU Journal of Science: Natural Sciences and Technology, Vol. 32, No. 1S (2016) 286-292

Evaluation of Recent Methods to Improve Recombinant *Helicobacter Pylori* Protein Yield and Solubility in *Escherichia Coli* Expression System

Nguyen Thi Huyen[1], Trinh Le Phuong[1],
Do Thi Thanh Trung[1], Le Hong Diep[2], Pham Bao Yen[1,*]

[1]*Key Laboratory of Enzyme and Protein Technology (KLEPT), VNU University of Science, 334 Nguyen Trai, Thanh Xuan, Hanoi, Vietnam*
[2]*Faculty of Biology, VNU University of Science, 334 Nguyen Trai, Thanh Xuan, Hanoi, Vietnam*

Received 08 August 2016
Revised 18 August 2016; Accepted 09 September 2016

Abstract: Successful expression of target genes, often indicated by high yield and solubility, is critical for studies involving recombinant proteins. Yet the most common bacterial expression system utilizing *Escherichia coli* as host cells is usually reported to produce low amounts of soluble target proteins. In this study, two *Helicobacter pylori* (*Hp*) genes, *Hp* lipase and *Hp* peptide deformylase (*Hp*-PDF), whose encoded proteins are crucial for bacterial growth and colonization, thus could be used to screen potential anti-*Hp* drugs, were designed to be expressed in such system. Genetic engineering, experimental biology, and computational biology methods were employed to enhance recombinant protein production. The result showed that *Hp*-lipase expression was most improved through construct design that used two restriction enzymes, *Nde*I and *Xho*I, including TEV sequence and 6xHis tag at the 3' end of the target gene. *Hp*-PDF production increased significantly (24%) by optimizing culture condition and IPTG concentration according to Design Expert prediction together with cobalt supplementation. Either the addition of chemicals (glycylglycine) or heat shock method enhanced the yield and solubility of the studied proteins. Conclusively, it is suggested that combination of genetic engineering and computational optimization was efficient for recombinant protein expression in *E. coli* in addition to the conventional experimental biology methods.

Keywords: Helicobacter pylori, lipase, peptide deformylase, *Escherichia coli* expression system.

1. Introduction

In recent years, recombinant DNA technology is increasingly popular due to its use in a wide range of important applications, especially in medical research. It allows controllable and efficient expression of large amounts of target proteins without having to purify from original sources. However, recombinant protein expression still faces several major issues, including low yield and solubility, particularly in the expression system in *E. coli,* possibly caused by the toxicity of the targets to the host cells [1, 2]. Despite the fact that insoluble proteins (localized in the inclusion bodies) could be used to obtain

* Corresponding author. Tel.: 84-982408770
Email: cinaus@gmail.com

renatured protein, the refolding process might be incomplete, resulting in misfolded, thus, inactive proteins [3, 4]. The optimization of expression process in *E. coli* to obtain high yields of soluble recombinant proteins still remains a matter of research.

There are various strategies to overcome obstacles in recombinant proteins expression in *E. coli* published in previous studies. A number of fusion partners, *e.g.* maltose binding protein (MBP), glutathione S-transferases (GSTs), N-utilization substance (NusA), and thioredoxin (TRX) [5, 6], were commonly included in recombinant construct design to increase protein solubility. Most notably, His-maltose binding protein (HisMBP) was reported to enhance the solubility up to 42% in 632 proteins experimented [7, 8]. Additional sequences, *e.g.* FLAG, 6xHis [9, 10, 11], are considered the most widely used tags for recombinant protein construction. Besides genetic engineering, there are chemicals reported to improve expression levels and solubility of recombinant proteins such as glycylglycine [1], glucose, ethanol [12], and cofactors of the recombinant enzymes to be expressed (*e.g.* divalent metallic ions). Physical methods like heat shock also contributed to the 5-fold increase in solubility of GlnRS protein [13]. Changing culture conditions such as temperature, induction time, and inducer concentration is another method frequently used [14]. Most recently, the incorporation of computational algorithm to the design of experiment and the prediction of optimal condition has brought success to recombinant protein production. For example, Design Expert software was reported to increase the yield of anti-keratin ScFv TS1-218 by 21 folds [15]. In this study, we examined a number of strategies to optimize expression of two *H. pylori* proteins, lipase and PDF, which are essential for the growth, colonization, and development of the bacteria. However, the discovery of anti-*H. pylori* compounds is a difficult task due to lack of mature protein targets. Therefore, increasing yields of soluble recombinant *H. pylori* proteins for screening new drugs against this bacteria is the obvious objective of our study.

2. Materials and methods

2.1. Construction of lipase and PDF expression vectors

H. pylori lipase and PDF genes were introduced into pET22b(+) vector between restriction sites of either *Nco*I or *Nde*I and *Xho*I, respectively. In case of *Nco*I, the recombinant proteins contained extra 22 amino acids of pET22b(+) vector at the N terminus (constructs Lip-1, Lip-2). All four constructs were designed with 6xHis tag before stop codon for protein purification purposes. Two of four constructs carried Tobacco etch virus (TEV) cleavage sequence in front of the 6xHis tag sequence (constructs Lip-2, Lip-3).

2.2. Expression of Hp-PDF and Hp-lipase recombinant proteins in BL21(DE3)RIL strain

Hp-PDF and *Hp*-lipase were transformed into competent BL21(DE3)RIL cells and the transformants were grown overnight on TSA plates containing 100 μg/ml ampicillin and 34 μg/ml chloramphenicol. Colonies from the plates were first inoculated in 3 ml LB medium containing appropriate antibiotics and grown overnight at 37°C with shaking. An aliquot of the start cultures was grown in 50 ml LB medium with antibiotics at different temperatures, ranging from 25°C to 42°C. IPTG was added when the cell density reached an OD_{600} between 0.6-0.8, then the cells were grown for additional 3 hours and harvested by centrifuging at 7000 rpm at 4°C for 20 minutes.

The cell pellet was resuspended in lysis buffer and sonicated on ice to obtain the crude lysate. The lysate was then centrifuged at 13000 rpm at 4°C for 20 minutes to separate cellular proteins into soluble and insoluble fractions. Collected fractions were analyzed by SDS-PAGE to confirm the expression level and the presence of target proteins.

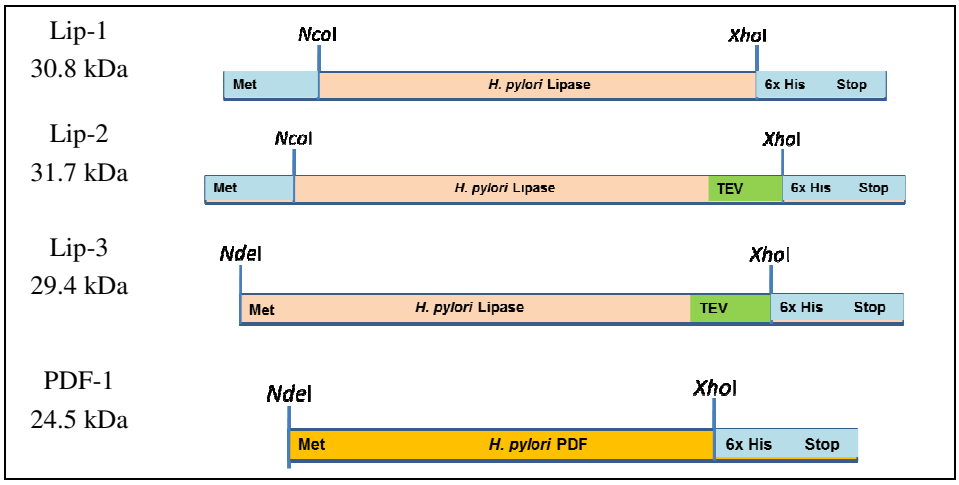


Figure 1. Schematic diagram of *Hp*-PDF and *Hp*-lipase vector components and the predicted molecular weights of the recombinant proteins.

2.3. Data analysis

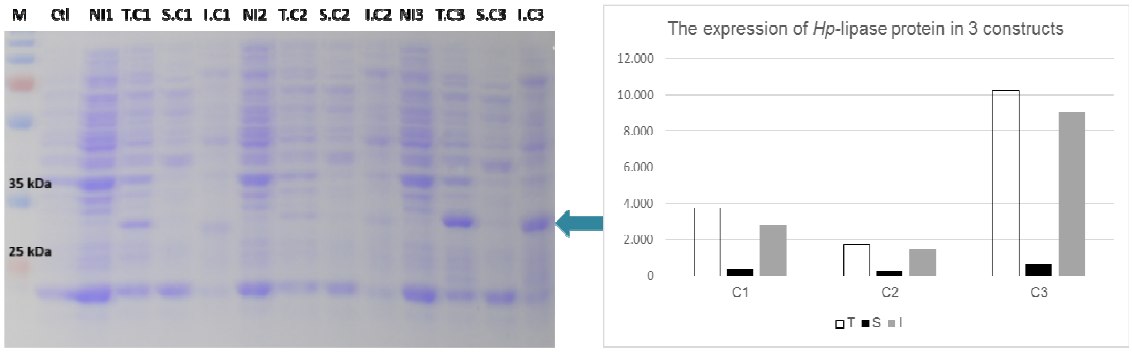
ImageJ software (https://imagej.nih.gov/ij/) was used to analyze the expression level of recombinant proteins through quantification of protein band intensities on SDS-PAGE gels. The exported data was used to calculate the yield and solubility of the proteins obtained. From these input values, Design Expert software (http://www.statease.com/dx10.html) established the matrix of factors affecting recombinant protein expression. The predicted optimal conditions were verified by experiment.

3. Results and discussion

3.1. Genetic engineering improved expression level of H. pylori lipase

The construct design step significantly increased lipase expression. Three lipase constructs had different expression levels of the recombinant enzyme when induced, however, the majority of the target protein remained in the insoluble fraction (Fig. 2A, lanes NI, T, and I). This solubility problem was described in a number of studies with the possible explanation of non-native proteins' toxicity to host cells, resulting in immediate transport to the inclusion body after synthesis [16, 17]. As shown as Fig. 2A, there was a band of approximately 30 kDa in each of the three total extracts (*i.e* crude lysates, labeled T.C1-T.C3).

Compared to the other constructs, Lip-3 had highest expression, about 3-fold higher than Lip-1 and 10-fold higher than Lip-2 (Fig. 2B). It is important to note that both Lip-3 and PDF-1 constructs had similar N termini which did not have 22 extra amino acids as a result of *Nco*I restriction digest, thus, started with wild-type methionine. This result provided the evidence for genetic engineering, in others words, constructs design, could be a useful strategy to improve recombinant protein expression level.



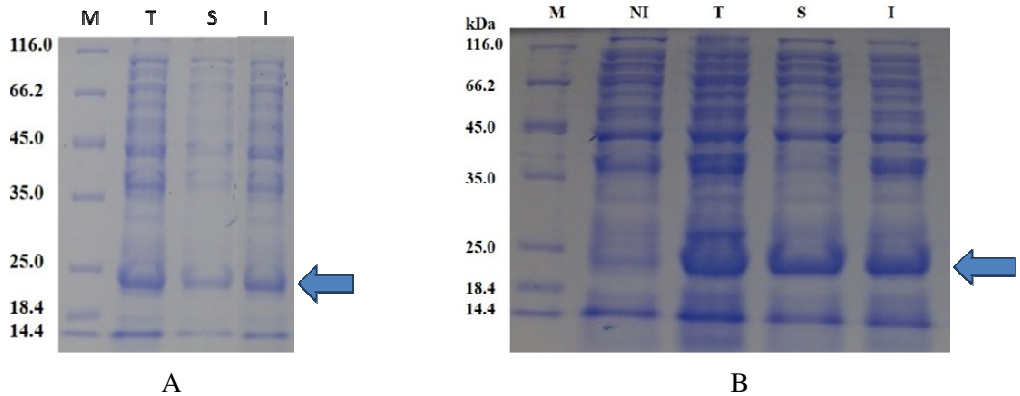
A B

Figure 2. The result of *Hp*-lipase expression with 3 constructs
M: molecular weight standards (kDa), Ctl: pET22b(+) empty, NI: Total protein noninduced,
T: Total protein, S: soluble fractions, I: Insoluble protein fractions, C1, 2, 3: Constructs 1, 2, 3.

3.2. Culture condition optimization increased Hp-PDF production in the soluble fraction

For optimizing growth condition, we examined different chemicals, physical and biological methods. General chemicals like glycylglycine did not improve expression and seemed to be toxic to the cells at high concentrations (0.5 and 1 M). However, the addition of cofactor (cobalt ion for PDF) and different concentrations of the inducer (IPTG) produced more soluble recombinant proteins, from approximately 8% to 20%, equal to a 2.5-fold increase (Fig. 3B, lane S).

Induction point (at different OD_{600} that reflect early-, mid-, or late- exponential growth phases) and culture temperature after induction (ranging from 16-42°C) are considered biological factors that affect the cells' ability to grow and produce recombinant proteins [18, 19]. Previous studies pointed out that IPTG addition at mid-exponential phase (OD_{600}=0.4-0.6) and lower growth temperature (25-30°C) could improve solubility. The culture conddition variation worked well for *Hp*-PDF, which was induced at OD_{600}= 0.6 using 0.8 mM IPTG and cultured at 37°C [20]. As indicated in Fig. 4, the final optimized culture conditions for *Hp*-PDF included 1 mM IPTG induction at OD_{600}= 0.6-0.8 and growing at 25°C, resulting in a 14% increase of the target protein in the soluble fraction.



A B

Figure 3. *Hp*-PDF expression in the absence (A) or presence (B) of Co^{2+}, M: molecular weight standards (kDa), NI: Total protein noninduced, T: Total protein, S: soluble fractions, I: Insoluble fractions.

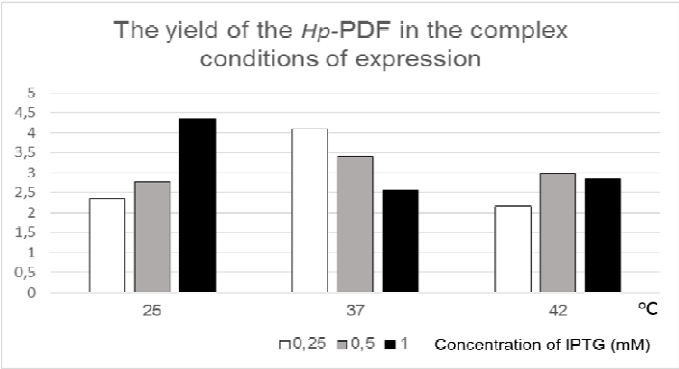


Figure 4. The yield of the *Hp*-PDF in the complex conditions of expression.

Heat shock method, presumably generating heat shock proteins that act as chaperones for proper protein folding, only slightly increased expression (data not shown). Similarly, effects of the culture condition optimization on *Hp*-lipase expression were insignificant.

3.3. Optimization through design expert software

In the previous culture condition optimization, only three different values for each factor (IPTG concentration and growth temperature) were experimented. The major disadvantage of experimental method is that it is almost impossible to conduct full factorial set of design that covers wide ranges of the variables. To solve this problem, we used the obtained data as inputs for further analysis by the Design Expert software. The algorithms were allowed to calculate and predict optimal IPTG concentration and temperature towards maximal yield (Fig. 5A), maximal solubility (Fig. 5B), and both (Fig. 5C). The dark red color indicated optimal ranges, clearly showing temperature of approximately 31°C and 1 mM IPTG were most productive regarding yield and solubility, respectively. When optimizing towards both factors, the software predicted maximal yield and solubility for expression at 28°C using 1 mM IPTG with 91% desirability.

The experiment carried at these conditions produced more soluble *Hp*-PDF, approximately a 15% increase compared to Do *et al.,* (2015), proving the validity of this strategy. The combination of experimental methods and analysis software was also applied successfully to some other recombinant proteins as reported in Jafari *et al.,* (2011) [15] regarding the optimization of culture conditions for anti-keratin ScFv TS1-218 production in the yeast expression system of *Pichia pastoris*. Furthermore, preliminary assay of *Hp*-PDF activity was conducted, showing that the recombinant enzyme was active, thus, could be used for future experiments.

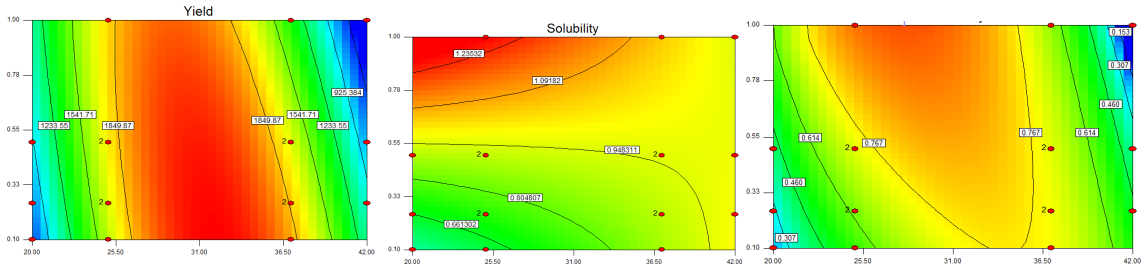


Figure 5. Analyzing results by Design Expert software, A: Yield, B: Solubility, C: Both.

4. Conclusion

Recombinant protein expression faces common obstacles of low yield, insolubility, and weak activity, especially when *E. coli* system is used. In this study, we examined solutions to these problems using three strategies including genetic engineering, experimental and computational biology with two targets, *Hp*-PDF and *Hp*-lipase. The results indicated that there was a 3- to 10-fold increase in *Hp*-lipase yield using genetic engineering, while *Hp*-PDF was obtained with more than 65% in the soluble fraction by culture condition and software optimization. Experiments showed that the efficiency of these strategies was different, depending on the target protein; therefore, we recommend the combination of methods for effective expression of soluble recombinant proteins in *E. coli.*

References

[1] Ghosh S., Sheeba R, Sheikh S. R., Sharmistha B., Rakesh K.C., Prachee C., Nasreen Z. E., Sangita M., Seyed E.H., Method for enhancing solubility of the expressed recombinant proteins in Escherichia coli, BioTechniques, 37 (2004) 418.

[2] Duong-Ly K.C., Gabelli S.B., Explanatory chapter: troubleshooting recombinant protein expression: general, Methods Enzymol (2014) 209.

[3] Saïda F., Uzan M., Odaert B., Bontems F., Expression of highly toxic genes in E. coli: special strategies and genetic tools, Curr Protein Pept Sci, 7 (1) (2006) 47.

[4] Dumon-Seignovert L., Cariot G., Vuillard L., The toxicity of recombinant proteins in Escherichia coli: a comparison of overexpression in BL21(DE3), C41(DE3), and C43 (DE3), Protein Expr Purif, 37 (1) (2004) 203.

[5] Stevens R.C., Design of high-throughput methods of protein production forstructural biology, Ways & Means, 8 (2000) 177.

[6] Cabrita L.D., Gilis D., Robertson A.L., Dehouck Y., Rooman M., Bottomley S.P., Enhancing the stability and solubility of TEV protease using in silico design, Protein Sci, 16 (11) (2007) 2360.

[7] Nallamsetty S., Waugh D.S., A generic protocol for the expression and purification of recombinant proteins in Escherichia coli using a combinatorial His6-maltose binding protein fusion tag, Nat Protoc, 2 (2) (2007) 383.

[8] Lebendiker M., Danieli T., Purification of proteins fused to maltose-binding protein, Methods Mol Biol, (2011) 281.

[9] Bornhorst J.A. and Falke J.J., Purification of Proteins Using Polyhistidine Affinity Tags, Methods Enzymol (2000) 245.

[10] Einhauer A., Jungbauer A., The FLAG peptide, a versatile fusion tag for the purification of recombinant proteins, J Biochem Biophys Methods, 49 (1-3) (2001) 455.

[11] Shih Y.P., Wu H.C., Hu S.M., Wang T.F., and Wang A.H., Self-cleavage of fusion protein in vivo using TEV protease to yield native protein, Protein Sci, 14 (4) (2005) 936.

[12] Chhetri G., Kalita P., and Tripathi T., An efficient protocol to enhance recombinant protein expression using ethanol in Escherichia coli, MethodsX, 2 (2015) 385.

[13] Chen J., Acton T.B., Basu S.K., Montelione G.T., and Inouye M., Enhancement of the Solubility of Proteins Overexpressed in Escherichia coli by Heat Shock, J. Mol. Microbiol. Biotechnol. 4 (6) (2002) 519.

[14] Francis D.M., Page R., Strategies to optimize protein expression in E. coli, Curr Protoc Protein Sci. 2010 5 (5) (2010) 24.

[15] Jafari R., Sundström B.E., Holm P., Optimization of production of the anti-keratin 8 single-chain Fv TS1-218 in Pichia pastoris using design of experiments, Microb Cell Fact, 16 (2011) 10.

[16] Gopal G. J., Kumar A., Strategies for the production of recombinant protein in Escherichia coli, Protein J,32 (6) (2013) 419.

[17] Kim T. I., Lee H., Hong H. K., Kim K. S., Choi S. I., Maeng Y. S., Kim E. K., Inhibitory Effect of Tranilast on Transforming Growth Factor-Beta-Induced Protein in Granular Corneal Dystrophy Type 2 Corneal Fibroblasts, Cornea, 34 (8) (2015) 950.

[18] Ariane L L, Nicolau J. F. M. Q, Esteves G. D. S, Vareschini D. T, Strategies for the production of recombinant protein in Escherichia coli, Protein J, 32 (6) (2013) 419.

[19] HunterT., BibbyE., BautistaD., A Time Course Induction Study of Recombinant Protein Expression in E. coli, Protein Sample Preparation, 2006.

[20] Cai ., Han C., Hu ., Zhang J., Wu D., Wang F., Liu Y., Ding J., Chen K., Yue J., Shen X., and Jian H., Peptide deformylase is a potential target for anti-Helicobacter pylori drugs: Reverse docking enzymatic assay, and X-ray crystallography validation, Protein Science 15 (2006) 2071.

Đánh giá một số phương pháp tăng mức độ biểu hiện của protein tái tổ hợp từ *Helicobacter pylori* trong hệ thống biểu hiện *Escherichia coli*

Nguyễn Thị Huyền[1], Trịnh Lê Phương[1], Đỗ Thị Thanh Trung[1], Lê Hồng Điệp[2], Phạm Bảo Yên[1]

[1]*Phòng Thí nghiệm trọng điểm Công nghệ Enzyme và Protein, Trường Đại học Khoa học Tự Nhiên, ĐHQGHN, 334 Nguyễn Trãi, Thanh Xuân, Hà Nội, Việt Nam*
[2]*Bộ môn Sinh lý Thực vật và Hóa sinh, Khoa Sinh học, Trường Đại học Khoa học Tự Nhiên, ĐHQGHN, 334 Nguyễn Trãi, Thanh Xuân, Hà Nội, Việt Nam*

Tóm tắt: Các nghiên cứu đã chỉ ra hai yếu tố quan trọng đánh giá sự thành công trong việc biểu hiện gen đích là hàm lượng và độ tan của protein tái tổ hợp. Trong các hệ thống biểu hiện ở vi khuẩn mà phổ biến nhất là *E. coli*, hàm lượng và độ tan của protein tái tổ hợp thu được còn thấp. Để giải quyết vấn đề này, bằng các phương pháp liên quan đến kĩ thuật di truyền, thực nghiệm và phần mềm tin sinh học, nghiên cứu của chúng tôi đã biểu hiện hai gen mã hóa cho hai protein quan trọng đối với sự sinh trưởng và xâm nhập của vi khuẩn *Helicobacter pylori* trong hệ thống *E. coli*. Kết quả thu được cho thấy, sự biểu hiện của *Hp*-lipase được cải thiện rõ rệt thông qua việc thiết kế cấu trúc sử dụng hai enzyme giới hạn *Xho*I và *Nde*I kèm theo các trình tự TEV và 6xHis ở đầu 3' của gen đích. Trong khi đó, việc bổ sungion Co^{2+}, chất cảm ứng IPTG và nhiệt độ biểu hiện thích hợp đã làm tăng thêm 24% độ tan của *Hp*-PDF. Ngoài ra, các phương pháp vật lý như sốc nhiệt hoặc phương pháp hóa học như bổ sung glycylglycine cũng đã được nghiên cứu ảnh hưởng đến hàm lượng và độ tan của protein tái tổ hợp.

Từ khóa: Helicobacter pylori, lipase, peptide deformylase, hệ thống biểu hiện ở vi khuẩn *Escherichia coli.*