

Evaluation of three different vectors on human iLRP expression in Escherichia coli

Bainan Liu (✉ bnliu@hotmail.com)

Dalian Medical University <https://orcid.org/0000-0002-5855-137X>

Dong Zhang

Zunyi Medical University

Lingli Yan

Zunyi Medical University

Ming Zhou

Cancer Research Institute, Central South University

Research article

Keywords: Immature laminin receptor protein, Prokaryotic expression vector, Recombinant expression, Protein tag

Posted Date: December 7th, 2019

DOI: <https://doi.org/10.21203/rs.2.18144/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: iLRP (immature laminin receptor protein) is a tumor associated antigen over-expressed on the surface of most human cancer cells and plays an important role in the process of tumorigenesis and development. It has strong auto-immunogenicity in patients and is a good target protein for tumor immunotherapy. The aim of this study is to find a practical and an efficient method for the production of recombinant iLRP, and to provide enough recombinant protein for further study of its potential applications.

Results: In this report, three different expression vectors based on pET-30a(+) from pET expression systems were constructed. The first one is to add a 6xHis-Tag to the N-terminal of natural iLRP, which is called as pET-His-iLRP. The second one is to include an S-Tag between the C-terminal of 6xHis-Tag and the N-terminal of natural iLRP, which is called as pET-His-S-iLRP. The third one only contains the 6xHis-Tag in front of the N-terminal of iLRP, but the natural iLRP gene was first optimized according to the bacterial genome, and the constructed vector was named as pET-His-Opt-iLRP. Then the expression of three vectors in *Escherichia coli* (*E. coli*) BL21 (DE3) was analyzed. Results demonstrated that the expression of iLRP was the highest in the vector of pET-His-Opt-iLRP, followed by the vector of pET-His-S-iLRP, and the vector of pET-His-iLRP expressed the lowest iLRP.

Conclusions: According to the results, it was concluded that the vector pET-His-Opt-iLRP combining the codon-optimized *iLRP* gene with a 6xHis-Tag had the best effect on the expression of iLRP, and the expression in vectors with the natural *iLRP* gene depends on its leader sequences in the expression frame. The longer leader sequences accommodate to the host *E. coli*, the higher expression the engineered gene would be.

Background

Immature laminin receptor protein (iLRP), the precursor of laminin receptor protein (LRP), is abundantly expressed in the early stage of embryonic development, but gradually disappeared after the middle and late stages. In normal adult tissues, its expression can hardly be detected [1]. However, studies have confirmed that iLRP is re-expressed in most of tumor tissues and plays an important role in the occurrence, development and metastasis of tumors, which is closely related to the degree of cancer deterioration [2]. Since its discovery, iLRP has been regarded as an important target in the study of cancer immunotherapy [3]. However, it is difficult to obtain sufficient pure protein for studying its functions and applications in clinics, because the natural tissues that can be used for purification are limited. In order to overcome these obstacles, large-scale production by genetic engineering is in the urgent need.

There are many factors to be considered for expressing recombinant proteins in non-native hosts. The first consideration is to choose an appropriate expression system. Usually, eukaryotic expression system is the best choice for target genes from eukaryotes, and prokaryotic expression system is the best choice for target genes from prokaryotes. However, eukaryotic expression system is more complex and difficult

to operate than prokaryotic expression system in the production of exogenous proteins. It is not suitable for large-scale production in a short time because of its low production and high cost [4]. Therefore, most of the current methods for large-scale production of exogenous proteins are carried out by prokaryotic expression system. Furthermore, prokaryotic expression system has been very mature with many advantages like as: simple operation, high yield, short production cycle, mature expression vectors easily available, and simpler purification methods [5]. Among the prokaryotic systems, the pET system developed by Novagen Company is the most successful and products by this system have been in the market, which has produced great economic value [6, 7]. In this study, the expression of iLRP was analyzed in detail by using this system, and the expression of three different vectors was compared. The difference among three vectors is: adding 6xHis-Tag to the N-terminal of natural iLRP to the construct, pET-His-iLRP; adding the S-Tag between the C-terminal of 6xHis-Tag and the N-terminal of natural iLRP to the construct, pET-His-S-iLRP; and adding 6xHis-Tag to the N-terminal of codon-optimized iLRP to the construct, pET-His-Opt-iLRP. Results showed that the most suitable expression vector for mass production of recombinant iLRP is the fusion structure of codon-optimized iLRP and 6xHis-Tag, pET-His-Opt-iLRP. By this vector, the product is not only easy to be expressed in large quantities and to be purified through Ni⁺ column, but also has little effect on the structure and function of the protein itself.

Results

Codon optimization

Codon optimization can refer to several meanings as, avoiding rare codons with low utilization, simplifying mRNA's secondary structure after gene transcription, removing the motif which is not conducive to efficient expression and add the helpful one, adjusting the GC content and other methods to re-design genes [8, 9]. In this study, GC content, codon adaptation index (CAI), mRNA structure and cis-acting elements in the human *iLRP* gene were optimized to match requirements for the expression in *Escherichia coli* (*E. coli*) by NG®Codon Optimization Technology. Through optimizing, a more smooth curve of GC content was obtained (Fig 1a), meanwhile, the stable hairpin structures were minimized to 1 from 11, which could promote ribosomes to translate more efficiently. Moreover, usage frequencies of codons and distribution of the gene became more fit to *E. coli*'s genome, and the CAI was improved from 0.60 to 0.84 after rare codons were replaced (Fig 1b). Additionally, the negative cis-acting elements and killer sequences were eliminated. Lastly, the gene optimization model evaluated the integrated scores that correlate with gene expression level negatively. For the wild *iLRP* gene, the scores are 1,632,435, and for the optimized gene, the scores are reduced to 776,954, which decreases the value by 47%. According to the results of codon optimization, the human *iLRP* gene sequence was changed to suit for the requirements of expressing in *E. coli*, but the amino acid sequence kept unchanged. The alignment of wild type sequence and optimized sequence clearly showed that both sequences encode a same protein (Fig 1c).

Construction of expression vectors

The vectors used to express *iLRP* in this study all were derived from plasmid pET30a(+), a category of vectors in the pET System that is the most powerful system yet developed for the cloning and expression of recombinant proteins in *E. coli*. Target genes were cloned in pET plasmids under control of strong bacteriophage T7 transcription and translation signals. Native human *iLRP* gene with a 6xHis-Tag was cloned between *NdeI* and *XhoI* to form the plasmid pET-His-iLRP. However, the optimized human iLRP gene with a 6xHis-Tag was cloned between *NdeI* and *EcoRV* to form the plasmid pET-His-Opt-iLRP because the restriction site of *XhoI* is already existing in the optimized gene sequence but *EcoRV* is not within. For the plasmid pET-His-S-iLRP, the restriction sites of *BglII* and *XhoI* were used so that the original His-Tag and the thrombin protease cleavage site and S-Tag were included into the frame of target expression, which increasing the length of leader sequence significantly comparing with pET-His-iLRP. The resultant vectors were confirmed by both restriction digestion and sequencing. The structure of vectors and the difference among them are demonstrated in Fig. 2.

Analysis of protein expression by SDS-PAGE and Western Blotting

Following gene transformation by the constructed vectors, transformants were cultured in TB (Terrific Broth) with kanamycin and the expression of recombinant proteins was induced by IPTG. First, SDS-PAGE was performed to check the expression in the whole cell lysates. Results showed that the vector pET-His-Opt-iLRP expressed the highest recombinant proteins with the lowest expression from pET-His-iLRP and the vector pET-His-S-iLRP expressed high enough proteins but little lower than pET-His-Opt-iLRP (Fig. 3a). Based on published documents [10] and our preliminary data, human *iLRP* gene was expressed in *E. coli* as a form of inclusion body, therefore, inclusion bodies were made for the analysis of protein expression by SDS-PAGE and same Results were obtained (Fig. 3b). To verify the authenticity of the expressed proteins, Western Blotting with the purchased anti-iLRP monoclonal antibody from Abcam, Shanghai, China was carried out. Clearly, target bands appeared as expected on the nitrocellulose membrane, with the darkest band from the induced vector pET-His-Opt-iLRP, lighter darkness band from the induced vector pET-His-S-iLRP, and the weakest band from the induced vector pET-His-iLRP (Fig. 3c). This confirmed that the expressed products in three vectors were the recombinant iLRP with the highest expression from pET-His-Opt-iLRP.

Discussion

37 kDa iLRP is prevalently overexpressed in human cancers and highly related to poor prognosis in patients. It plays an important role in tumor invasion and metastasis and is also a key factor for tumor cell proliferation, survival and protein translation [11]. More importantly, it has been reported by researchers that iLRP is a strong tumor rejection antigen able to autologously induce immune response in cancer patient [12, 13]. These data convinced that iLRP is a suitable candidate target for tumor immunotherapy. To fast forward its application in future, producing large amount of the protein is urgently in need. In this study, the most powerful system yet developed for heterologous protein expression, pET system was used to explore a practical method of iLRP production. Accordingly, three

different expression vectors were constructed on the basis of pET30a(+) and their expression was then compared and analyzed.

In designing these three vectors, the possible influence of protein tags on the expression of the target gene and the structure of the target gene were the main factors to be taken into account. One of the frequently used protein tags, 6xHis-Tag is included for each vector in this study. The tag is not only simple in structure composing of only 6 histidines with no need to remove it from the final products, but also very convenient for the follow-up research on the engineered protein. The label has strong affinity with Ni⁺, which makes the purification of overexpressed proteins simple by using Ni⁺-affinity column [14]. In addition, the expressed products can be easily verified and detected by the monoclonal antibody against 6xHis-Tag that is very common in the current market [15]. Therefore, each vector was constructed to have the His-Tag added to the 5'-end of the fusion proteins. In certain cases, it is practical to add other kinds of protein tags for heterologous protein expression in *E. coli*, such as S-protein tag. S-Tag is an oligopeptide derived from bovine pancreatic ribonuclease A, which consists of 15 amino acids and is rich in charged and polar amino acid residues [16]. This tag has been studied extensively and can not only increase the expression yield and improve the solubility of target protein [17], but also has little effect on the structure and function of the recombinant protein with weak immunogenicity either [18]. It binds to ligand S-protein with high specificity and affinity, leading to the development of simple technique for protein purification with widely used S-protein affinity column in the market [19]. Considering of these merits, a vector containing both His-Tag and S-Tag was designed for the comparison with other expression vectors. Between pET-His-S-iLRP and pET-His-iLRP, it was apparent that the former had much higher expression level than the latter (Fig. 3). Since the only difference between both vectors is in the presence or absence of S-Tag, the result consistently confirmed that the S-Tag had critical influence on the expression of iLRP as expected. However, the overly extended sequence in front of the target protein could remain a persistent concern for its future applications and in most cases, it is required to be removed even though it does greatly enhance the expression [20]. This extra step will add uncertainties to the final production for the recombinant protein. Therefore, other methods that possibly have similar or even better effects on the expression for the target gene could be developed. Accordingly, a vector with the optimized iLRP gene was constructed. To our best knowledge, the gene structure is another crucial aspect to be considered for engineered gene expression in the host besides the addition of protein tags. This was verified by the huge difference of expression level among three vectors (Fig. 3). Between pET-His-iLRP and pET-His-Opt-iLRP, although the vectors both encode same amino acid sequences for the target protein iLRP, the expression level was totally different. The vector pET-His-iLRP expressed too few recombinant iLRPs to be detected by Coomassie Brilliant Blue staining after SDS-PAGE and even hardly be detected by Western Blotting. However, the expression increased substantially with the vector pET-His-Opt-iLRP including the optimized iLRP gene sequence according to the host *E. coli*'s genome. Obviously, the composition of the target gene sequences exerted tremendous impact on the expression level. According to the report [9], heterologous expression of rare codon-containing genes is likely to exhaust the endogenous pools of the analogous tRNAs, which leading to growth inhibition, premature termination of transcription and/or translation, decreased mRNA stability, and increased frame-shifts, deletions and

misincorporations. Through the technique of gene optimization, these factors could be easily modified to accommodate with the need in expression for the host. In this case, the genetic constitution of iLRP had been changed a lot by changing the GC content, improving the CAI, stabilizing the mRNA structures, and reducing the integrated scores negatively correlated with the protein expression (Fig. 1). Combining all above factors, it is very reasonable for the great enhancement of the iLRP expression in pET-His-Opt-iLRP comparing with the native gene structure. As for the vector pET-His-S-iLRP, its high level of expression might be attributed to the fusion sequence in front of iLRP. This sequence is composed of a 6xHis-Tag and a thrombin protease cleavage site and an S-protein tag, around 84 base pairs which close to 10% of the length of iLRP gene. As a common sense, constituent sequences in pET system was modified to the optimum conditions to ensure the best expression for heterologous proteins. Logically, the longer DNA sequences accommodated to bacterial components are included in the target protein, the higher expression will be reached. This logicalness meets with the results well in this study (Fig. 3). In pET-His-iLRP, the His-Tag is only 18 base pairs which merely close to 2% of the length of iLRP gene and hardly put any influence on the expression, leading to the expression level could not be improved. This is also consistent with the data from other researchers [10]. In their study, a Tobacco Etch Virus (TEV) protease site was added to the 5'-end of human iLRP gene, plus the His-Tag, around 51 extra base pairs optimized for the expression in *E. coli* were included in the expression frame. Therefore, they also had good expression even with the native iLRP gene for the production of recombinant iLRP. From this point of view, the optimized iLRP gene in pET-His-Opt-iLRP deserved the highest expression level, followed by pET-His-S-iLRP with the lowest expression in pET-His-iLRP. These results illustrated that the yield of recombinant protein in a heterologous host could be significantly improved by changing the gene constituents to accommodate to the host genome either through codon optimizing or with a proper leader sequence.

Conclusions

Through the comparison of three constructed vectors with different components for iLRP production, it was found that the combination of codon-optimized iLRP and 6xHis-Tag had the best effect on the expression, and the expression in vectors with the natural iLRP genes depends on the leader sequences in the expression frame. The longer leader sequences accommodate to the host, the higher expression the engineered gene would be. In terms of this finding, the full-length sequences of iLRP gene was optimized according to the host of *E. coli*'s genome. In this way, the whole reading frame of the fusion protein was accommodated to the host, which resulting in the highest expression among three vectors. In addition, proper protein tags can also put important influence on the expression level for the target gene. In this study, the S-Tag played the role very well, which was shown by the distinct expression between pET-His-S-iLRP and pET-His-iLRP. In summary, this study found that the heterologous expression of human iLRP in a bacterial host could be greatly increased by genetically modifying gene constituents and using proper protein-tags and laid the foundation on the future researches on its applications.

Methods

Chemicals and strains and plasmids

Terrific Broth (TB) and kanamycin were purchased from ThermoFisher Scientific, Chengdu, China for culturing bacterial strains. PCR kit, restriction enzymes, and gel extraction kit were purchased from TaKaRa Co. Ltd., Beijing, China. The primary antibody, mouse anti-iLRP and the secondary antibody, goat anti-mouse IgG conjugated with Horseradish peroxidase (HRP) were bought from Abcam, Shanghai, China. Other molecular biology reagents were from Solarbio Science & Technology Co. Ltd., Beijing, China. All chemicals in this study were with analytical grade or a higher purity.

Escherichia coli DH5 α was used as the host for recombinant DNA manipulation, and BL21(*DE3*) was used for gene expression. Competent cells of bacterial strains were purchased from ThermoFisher Scientific, Chengdu, China. Plasmids pUC57 and pET30a(+) for cloning and expression were from ChinaPeptides, Shanghai, China. The plasmid pENTER containing the natural human *iLRP* cDNA was bought from Vigene, Shandong, China. All hosts and plasmids were stored at -80 °C in our lab.

Codon optimization

The technology of codon optimization has been a valuable tool for producing proteins as therapeutic agents or research reagents in heterologous host. With the availability of large amount of genomics data, and with the increased knowledge of protein expression, function and structure relationships, gene expression levels can be significantly improved by this technique [8, 21]. During the process of applying the technique, a variety of key elements involved in different stages of protein expression, like as GC content, codon adaptability, mRNA structure, and various cis-elements in transcription and translation should be carefully taken into consideration [9, 22]. In this study, human *iLRP* gene was optimized in terms of above points using a commercial proprietary algorithm, NG® Codon Optimization Technology (Synbio Technologies, NJ, USA) in order to analyze its expression in *Escherichia coli*.

Construction of expression vectors

Plasmid pET-30a(+) (from Novagen) was used as the backbone for the construction of vectors in this study. The vector pET-His-iLRP was constructed by inserting the natural *iLRP* gene between restriction sites of *Nde*I and *Xho*I. The natural *iLRP* gene was amplified by PCR using the plasmid pENTER (Vigene, Shandong, China) containing the human iLRP gene [23] as template. In order to add a restriction site of *Nde*I and a 6xHis-Tag to the N-terminal of iLRP, the 5'-primer sequence was synthesized by Sangon Biotech, Shanghai, China for this PCR with the sequence as: 5'-ATACATATGCACCATCATCATCATCATTCCGGAGCCCTTGATG-3'. Similarly, a restriction site of *Xho*I was added to the 3'-primer, and the sequence is: 5'-CCGCTCGAGTTAAGACCAGTCAGTGGTTGC-3'. To construct the vector pET-His-S-iLRP, the restriction sites of *Bgl*II and *Xho*I were used for the process of cloning so that the original 6xHis-Tag and the S-Tag were kept as it was. Therefore, the 5'-primer for amplifying the *iLRP* gene from pENTER was designed as: 5'-ACGAGATCTTCCGGAGCCCTTGATG-3', and the 3'-primer is same as above. PCR was both performed at the following conditions: initial denaturation at 98°C for 2 min, 30 cycles of denaturation at 98°C for 40 s, annealing at 58°C for 40 s, and extension at

72°C for 1 min, and the final extension at 72°C for 4 min and then maintenance at 4°C forever. In the construction of pET-His-Opt-iLRP, the codon-optimized iLRP gene with the additional elements of restriction site of *NdeI* and the 6xHis-Tag at 5'-end of *iLRP* and *EcoRV* at the 3'-end of *iLRP* was synthesized by ChinaPeptides, Shanghai, China and then cloned into pET-30a(+) between *NdeI* and *EcoRV*. All constructed vectors were characterized by sequencing and digestion of restriction enzymes.

Protein expression

Above vectors were transformed into *Escherichia coli* BL21 (*DE3*) (from Novagen) for the analysis of protein expression at a small scale. Briefly, on day 1, a fresh TB agar plate containing 50 µg/ml kanamycin was inoculated with bacterial stocks for each vector. On day 2, well isolated single clones were picked from the plates and cultured overnight in test tubes with 3 ml TB containing 50 µg/ml kanamycin, shaking at the speed of 200 rpm and at 37°C. In the morning of day 3, overnight cultures were transferred to 500 ml flasks with 100 ml fresh TB and 50 µg/ml kanamycin, continuously shaking at the speed of 200 rpm and at 37°C until the optical density at 600 nm (OD₆₀₀) reached to 1.0. At this point, a small aliquot of cultures for each sample was collected as negative control. Then, the rest of cultures was induced by adding isopropyl-thio-β-D-galactoside (IPTG) to a final concentration of 1.0 mM and shaking at the conditions above until the OD₆₀₀ reached to the same point of 6.0. Cell cultures were harvested by centrifuging at 4°C, 5000 rpm for 10 min and stored at -80°C.

SDS-PAGE and Western Blotting

Cell pellets harvested above were divided into two parts. One of them was directly lysed to prepare samples of whole cell lysates for SDS-PAGE (sodium dodecyl sulfate-polyacrylamide gel electrophoresis) and another one was used to prepare inclusion bodies. Cell pellets were washed with ice-cold phosphate-buffered saline (PBS, PH 7.2) twice, and resuspended with 5 ml of PBS. The cells in suspension were disrupted by sonication with 20 s pulse-on and pulse-off time each using an ultrasonicator (Ningbo Scientz Biotechnology Co. LTD, China). The inclusion bodies were then isolated by centrifugation at 12,000 rpm for 20 min at 4°C. After washing twice with 20 ml ice-cold PBS, the inclusion bodies were dissolved in 5 ml of solubilization buffer (10 mM Tris-HCl, 100 mM sodium phosphate, 6 M urea, 2 mM 2-mercaptoethanol, PH 8.0). After centrifugation at 12,000 rpm for 20 min, the total protein in the supernatant was determined with a BCA protein assay kit (ThermoFisher Scientific, Chengdu, China), and then, the supernatant was used to prepare the samples for SDS-PAGE with equal protein amount of each sample being loaded onto the gel. All samples for SDS-PAGE were boiled in 1x reducing loading buffer at 98°C for 10 min and subjected to 12% denaturing SDS-PAGE. Gels were stained by Coomassie Brilliant R-250 (ThermoFisher Scientific, Chengdu, China) to identify the expected protein bands. For Western Blotting, gels were blotted to nitrocellulose membrane, and then the membrane was blocked with 2.5% (w/v) skimmed milk powder in PBS, pH 7.2 for 2 h at room temperature (RT), and successively incubated with mouse anti-iLRP monoclonal antibody (Abcam, Shanghai, China) at a dilution of 1:5,000 at 4°C overnight. Next day, the membrane was washed 3 times, and then incubated with a 1:10,000 dilution of horseradish peroxidase (HRP) conjugated anti-mouse IgG secondary antibody at RT for 2 h. After the

incubation, wash the membrane 5 times, then incubate the membrane with electrochemiluminescence (ECL) reagent (ThermoFisher Scientific, Chengdu, China) for 5 min in dark. Protein bands were detected by ChemiDoc™ MP imaging System (Bio-Rad Laboratories Co. LTD, Shanghai, China).

Abbreviations

CA: Bicinchoninic acid; CAI: Codon adaptation index; *E. coli*: *Escherichia coli*; ECL: Electrochemiluminescence; HRP: Horseradish peroxidase; iLRP: Immature laminin receptor protein; IPTG: Isopropyl-thio-β-D-galactoside; LRP: Laminin receptor protein; OD: Optical density; PBS (phosphate-buffered saline); PCR: Polymerase chain reaction; RT: Room temperature; SDS-PAGE: Sodium dodecyl sulfate-polyacrylamide gel electrophoresis; TB: Terrific Broth; TEV: Tobacco Etch Virus.

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

Not applicable

Competing interests

The authors declare that they have no competing interests.

Funding

The study was funded by the National Natural Science Foundation of China (Grant No. 81560276), and the Foundation for Returnees from the Department of Human Resources and Social Security of Guizhou Province of China (Grant No. [2014]08). These funding bodies did not play any role in the design of the study, the collection, analysis, and interpretation of data and in writing the manuscript.

Authors' contributions

BL conceived and designed the study and performed the data analysis and wrote the manuscript; DZ designed and performed most of experiments and data analysis; LY performed some experiments and data analysis; MZ reviewed the manuscript and provided constructive suggestions during the research. All authors read and approved the final manuscript.

Acknowledgements

Not applicable

References

1. Barsoum AL, Schwarzenberger PO. Oncofetal antigen/immature laminin receptor protein in pregnancy and cancer. *Cell Mol Biol Lett*. 2014;19(3):393-406.
2. Coggin JH, Jr., Rohrer JW, Barsoum AL. True immunogenicity of oncofetal antigen/immature laminin receptor protein. *Cancer Res*. 2004;64(13):4685; author reply
3. Pesapane A, Ragno P, Selleri C, Montuori N. Recent Advances in the Function of the 67 kDa Laminin Receptor and its Targeting for Personalized Therapy in Cancer. *Curr Pharm Des*. 2017;23(32):4745-57.
4. Yousefi-Rad N, Shokrgozar MA, Behdani M, Moradi-Kalbolandi S, Motamedi-Rad M, Habibi-Anbouhi M. Antigenic assessment of a recombinant human CD90 protein expressed in prokaryotic expression system. *Protein Expr Purif*. 2015;116:139-43.
5. Rosano GL, Ceccarelli EA. Recombinant protein expression in *Escherichia coli*: advances and challenges. *Front Microbiol*. 2014;5:172.
6. Hou WH, Fang T, Chai YR, Wang TY, Wang JM, Xue LX. Expression of recombinant kringle 1-5 domains of human plasminogen by a prokaryote expression system. *Protein Expr Purif*. 2006;47(1):93-8.
7. Kalim M, Chen J, Wang S, Lin C, Ullah S, Liang K, et al. Construction of high level prokaryotic expression and purification system of PD-L1 extracellular domain by using *Escherichia coli* host cell machinery. *Immunol Lett*. 2017;190:34-41.
8. Gao CY, Xu TT, Zhao QJ, Li CL. Codon optimization enhances the expression of porcine beta-defensin-2 in *Escherichia coli*. *Genet Mol Res*. 2015;14(2):4978-88.
9. Hanson G, Collier J. Codon optimality, bias and usage in translation and mRNA decay. *Nat Rev Mol Cell Biol*. 2018;19(1):20-30.
10. Barsoum AL, Liu B, Rohrer JW, Coggin JH, Jr., Tucker JA, Pannell LK, et al. Production, safety and antitumor efficacy of recombinant Oncofetal Antigen/immature laminin receptor protein. *Biomaterials*. 2009;30(17):3091-9.
11. Poon SL, Klausen C, Hammond GL, Leung PC. 37-kDa laminin receptor precursor mediates GnRH-II-induced MMP-2 expression and invasiveness in ovarian cancer cells. *Mol Endocrinol*. 2011;25(2):327-38.
12. Rohrer JW, Barsoum AL, Coggin JH, Jr. Identification of oncofetal antigen/immature laminin receptor protein epitopes that activate BALB/c mouse OFA/iLRP-specific effector and regulatory T cell clones. *J Immunol*. 2006;176(5):2844-56.

13. Friedrichs B, Siegel S, Kloess M, Barsoum A, Coggin J, Jr., Rohrer J, et al. Humoral immune responses against the immature laminin receptor protein show prognostic significance in patients with chronic lymphocytic leukemia. *J Immunol.* 2008;180(9):6374-84.
14. Zhao D, Huang Z. Effect of His-Tag on Expression, Purification, and Structure of Zinc Finger Protein, ZNF191(243-368). *Bioinorg Chem Appl.* 2016;2016:8206854.
15. Yin H, Sun B, Zhou Y, Wang M, Xu Z, Fu Z, et al. A new strategy for methylated DNA detection based on photoelectrochemical immunosensor using Bi₂S₃ nanorods, methyl bonding domain protein and anti-his tag antibody. *Biosens Bioelectron.* 2014;51:103-8.
16. Raines RT, McCormick M, Van Oosbree TR, Mierendorf RC. The S.Tag fusion system for protein purification. *Methods Enzymol.* 2000;326:362-76.
17. Kobayashi H, Swapna GV, Wu KP, Afinogenova Y, Conover K, Mao B, et al. Segmental isotope labeling of proteins for NMR structural study using a protein S tag for higher expression and solubility. *J Biomol NMR.* 2012;52(4):303-13.
18. Zhang Y, Wang D, Sadee W. Calmodulin interaction with peptides from G-protein coupled receptors measured with S-Tag labeling. *Biochem Biophys Res Commun.* 2005;333(2):390-5.
19. Zhao X, Li G, Liang S. Several affinity tags commonly used in chromatographic purification. *J Anal Methods Chem.* 2013;2013:581093.
20. Zhang J, Sun A, Dong Y, Wei D. Recombinant Production and Characterization of SAC, the Core Domain of Par-4, by SUMO Fusion System. *Appl Biochem Biotechnol.* 2018;184(4):1155-67.
21. Song H, Li G, Mai W, Huang G, Chen K, Zhou Y, et al. Codon optimization enhances protein expression of *Bombyx mori* nucleopolyhedrovirus DNA polymerase in *E. coli*. *Curr Microbiol.* 2014;68(3):293-300.
22. Gvritshvili AG, Leung KW, Tombran-Tink J. Codon preference optimization increases heterologous PEDF expression. *PLoS One.* 2010;5(11):e15056.
23. Vania L, Rebelo TM, Ferreira E, Weiss SFT. Knock-down of LRP/LR promotes apoptosis in early and late stage colorectal carcinoma cells via caspase activation. *BMC Cancer.* 2018;18(1):602.

Figures

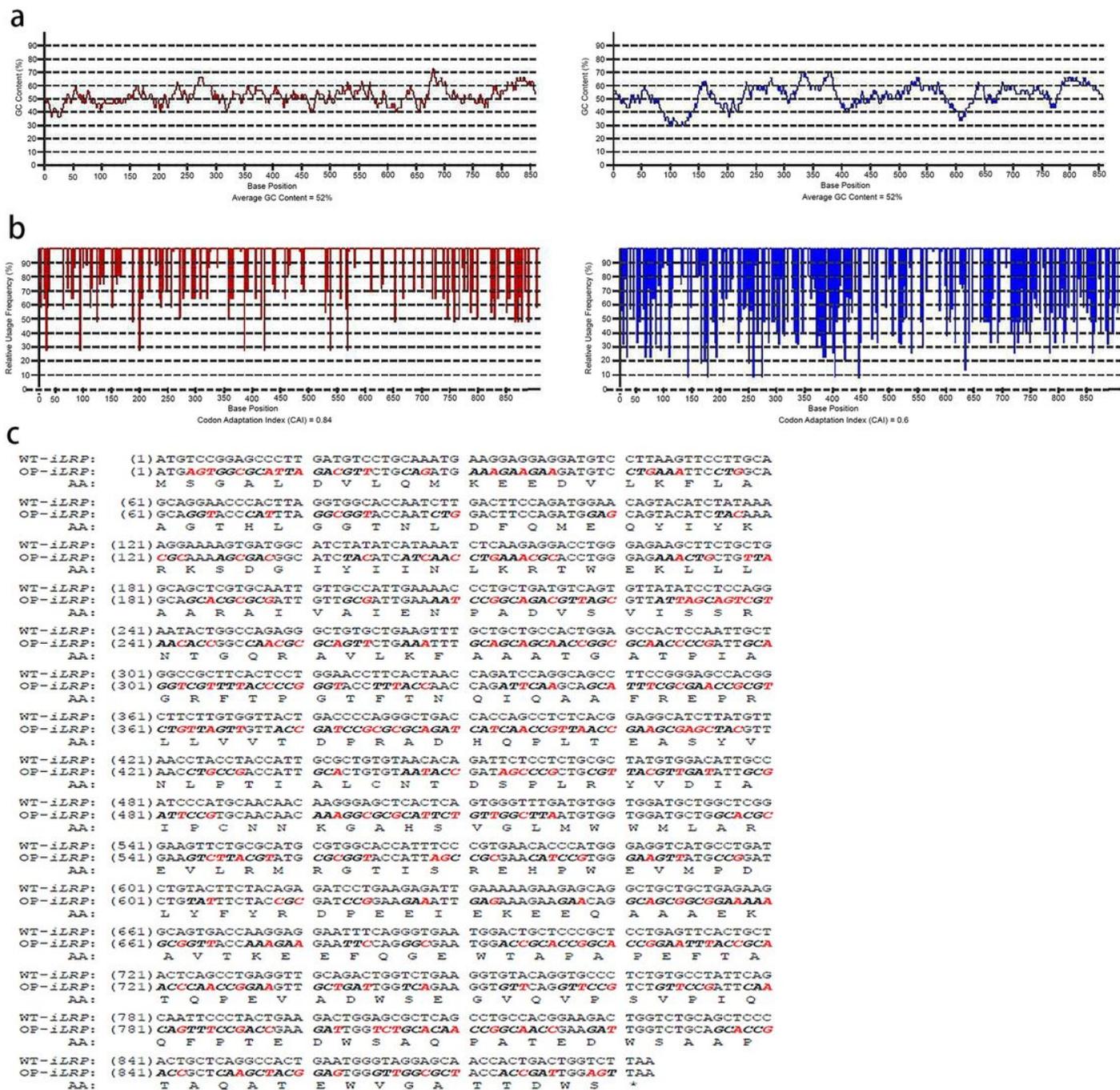


Figure 1

Comparison on feature parameters between the native iLRP gene and the optimized one. a The GC content change. Y axis denotes local GC content calculated by the window of 30 base pairs, and X axis represents base position. In spite of the mean value of 52% in GC content for both genes, a smoother variation curve was obtained after optimizing, with significant changes at base positions 50 to 100 and 300 to 400. The left panel denotes the variation curve for the optimized gene, and the right panel for the wild gene. b Analysis of codon usage bias by Codon Adaptation Index (CAI). CAI measures the deviation

of a given protein coding sequence according to a reference set of genes. It provides an indication of gene expression level with a value range from 0 to 1. A higher value predicts a higher expression level. The left panel represents the plot of optimized iLRP gene with CAI value of 0.84, and the right panel represents the plot of wild iLRP gene with CAI value of 0.6. c Sequence comparison between the optimized and wild iLRP gene. OP: codon optimized; WT: wild type; AA: amino acid. Differences in synonymous codons between the wild and optimized gene were indicated by bold and italic letters respectively. Individual bases in red represent the base replacement by bacterial preferred ones. Translated amino acids are shown in standard single-letter. There is no difference in amino acid sequences between two DNA sequences after translation.

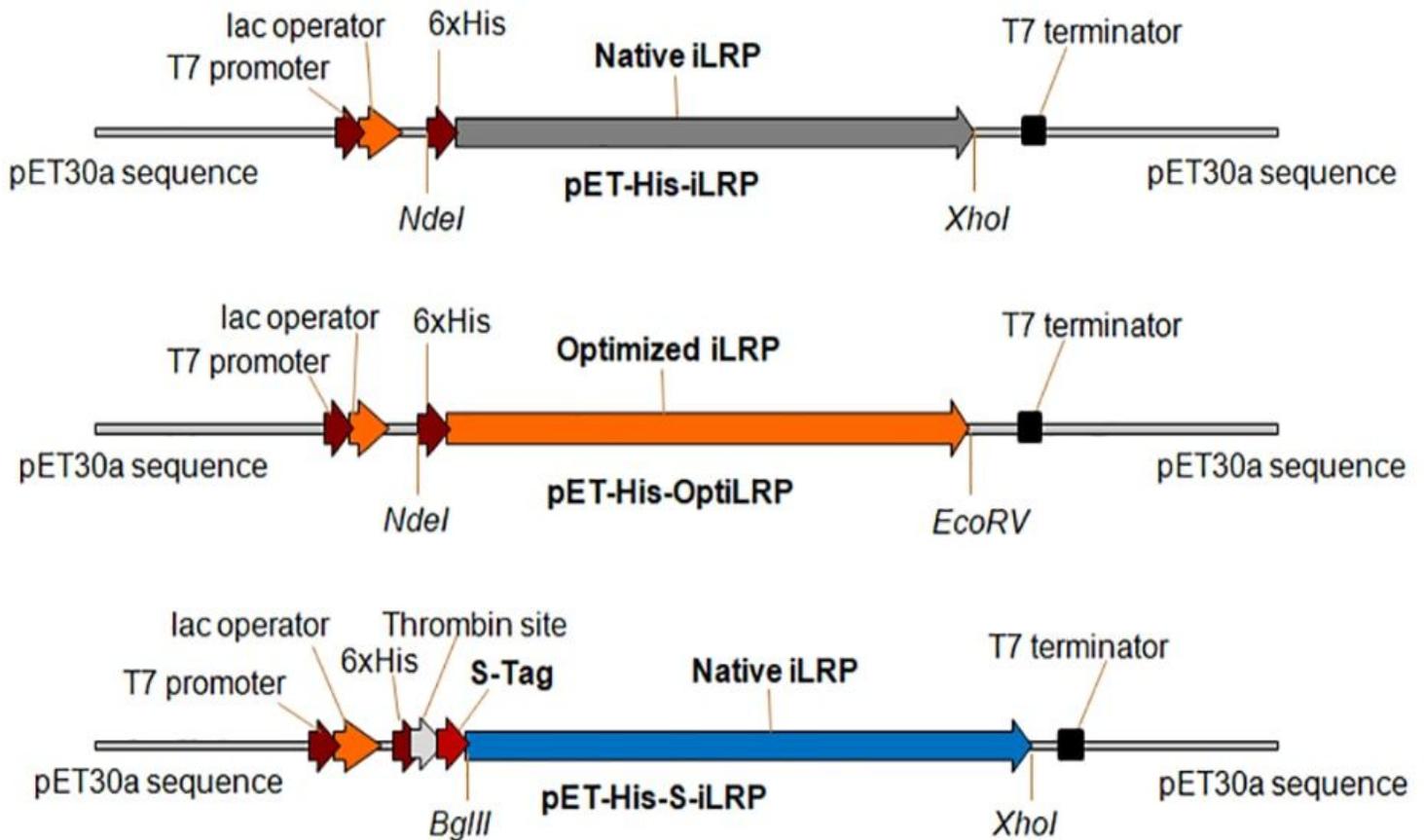


Figure 2

Structural comparison among three vectors. The vector pET-His-iLRP contains a His-Tag and the native human iLRP gene sequence that was inserted into pET30a(+) between NdeI and XhoI. The His-Tag uses the same sequence as in pET30a(+), therefore, the percentage of inherent composites from the host is close to 2% in the iLRP expression frame. The vector pET-His-Opt-iLRP harbors the optimized iLRP gene sequence that was inserted into the same plasmid between NdeI and EcoRV. In this vector, the full length of the expression structure matches the codon usage bias for the host. The vector pET-His-S-iLRP was constructed by inserting the native iLRP gene sequence into the plasmid between BglII and XhoI in order to make use of intrinsic components of His-Tag and S-Tag in it. Accordingly, the percentage of inherent

composites from the host reached up to 10% of the recombinant gene. All inserts are under the control of T7 promoter and lac operator.

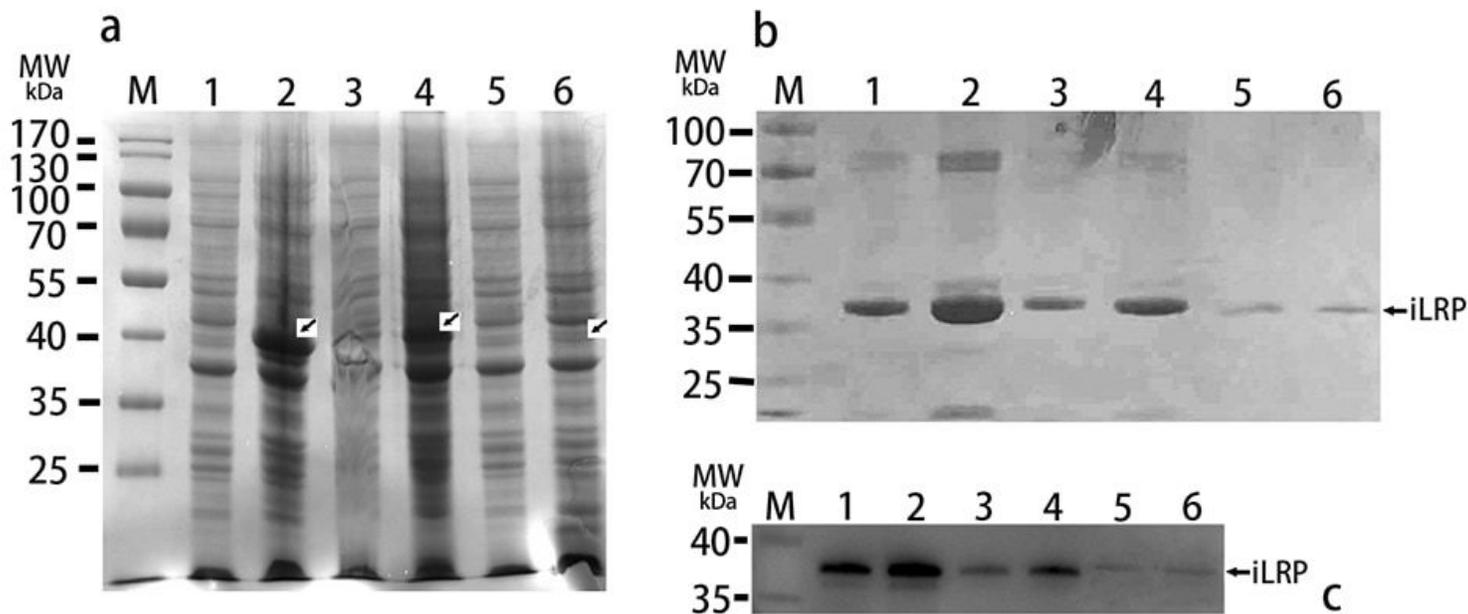


Figure 3

Evaluation of three vectors on their expressions. a Analysis of iLRP expression in whole cell lysates by SDS-PAGE. Bands representing the target protein expression were pointed at with arrows. The darker the band is, the higher expression it means. The result showed that the highest expression of iLRP was from lane 2, the second from lane 4 and the least from lane 6. Lane 1: cell lysates from pET-His-Opt-iLRP before IPTG induction; Lane 2: cell lysates from pET-His-Opt-iLRP induced by 1.0 mM IPTG; Lane 3: cell lysates from pET-His-S-iLRP before IPTG induction; Lane 4: cell lysates from pET-His-S-iLRP induced by 1.0 mM IPTG; Lane 5: cell lysates from pET-His-iLRP before IPTG induction; Lane 6: cell lysates from pET-His-iLRP induced by 1.0 mM IPTG. b Analysis of iLRP expression in inclusion bodies. Equal amount of lysates of inclusion bodies from each vector was subjected to SDS-PAGE. The result indicated that the highest iLRP expression was from pET-His-Opt-iLRP, the second from pET-His-S-iLRP and the least from pET-His-iLRP, which was consistent with the results in part a. The order of sample loading from lane 1 to lane 6 was same as in part a but loading with lysates of inclusion bodies. c Western Blotting analysis. The result not only showed that expression levels among three vectors were consistent with the above two experiments, but also confirmed the protein identity for the expressed products. The order of sample loading from lane 1 to lane 6 was same as in part a and b, with each lane loading 10 μ g of total protein from inclusion bodies determined by BCA. Lane M: molecular weight marker. MW: molecular weight (kDa).