Optimization of mRNA 5' End TIR Region Secondary Structure for Enhanced Recombinant sTNFα RI Expression in Escherichia coli - Postprint

Authors: Qin Jiaorong, Zhao Zhao, Luo Xinmei, Li Chunyang

Date: 2018-01-16T00:00:00+00:00

Abstract

Objective: To improve the expression level of soluble tumor necrosis factor type I receptor (sTNFαRI) in Escherichia coli (E. coli BL21(DE3)) by optimizing the secondary structure of the 5' mRNA translation initiation region (TIR) of PET11b-sTNFαRI. Methods: Through analysis of the free energy and nucleotide positional entropy of the secondary structure of the 5' TIR region of PET11b-sTNFαRI mRNA, corresponding primers were designed to mutate the codons in the 5' translation initiation region (TIR) of the mRNA, thereby exposing the ribosome binding site (RBS) and start codon (AUG) outside of the hairpin structure. Additionally, the ribosome binding site of pET11b was mutated from GAAGGAGA to GAAGAA to facilitate translation complex assembly and translation initiation. The optimized 5' TIR region sequence was cloned together with the sTNFαRI sequence into the pET11b vector through gene cloning methods and transformed into E. coli BL21(DE3). Positive transformants were induced for expression with IPTG and detected by SDS-PAGE and Western blot. Results: Through optimization of the secondary structure of the 5' TIR mRNA of PET11b-sTNFαRI, SDS-PAGE and Western blot analysis demonstrated that the expression level of recombinant sTNFαRI was increased by 50-60% compared to before optimization. Conclusion: Optimization of the secondary structure of the mRNA sequence in the translation initiation region (TIR) of recombinant vectors can effectively improve the expression level of target proteins, which holds important applied value for further industrial production.

Full Text

Preamble

Increasing the Expression Level of Recombinant sTNFαRI in *E. coli* Through Optimization of mRNA 5' Terminal TIR Secondary Structure

Jiao-rong QIN, Zhao ZHAO, Xin-mei LUO, Chun-yang LI
Chengdu Institute of Biological Products Co., Ltd., Chengdu 610023, China

Corresponding author: chunyang70@gmail.com

Abstract

Objective: To enhance the expression level of soluble tumor necrosis factor type I receptor (sTNFαRI) in *E. coli* BL21(DE3) by optimizing the secondary structure of the translation initiation region (TIR) in PET11b-sTNFαRI 5' mRNA.

Methods: We first analyzed the free energy and nucleotide position entropy of the TIR secondary structure in PET11b-sTNFαRI mRNA 5' end. Based on this analysis, we designed primers to mutate codons in the mRNA 5' translation initiation region (TIR) to expose the ribosome binding site (RBS) and start codon (AUG) outside of hairpin structures. Additionally, the pET11b ribosome binding site was mutated from GAAGGAGA to GAAGAA to facilitate translational complex assembly and initiation. The optimized 5' TIR sequence was cloned together with the sTNFαRI gene into the pET11b vector and transformed into *E. coli* BL21(DE3). Positive transformants were induced with IPTG and analyzed by SDS-PAGE and Western blot.

Results: Optimization of the PET11b-sTNFαRI 5'TIR mRNA secondary structure increased recombinant sTNFαRI expression by 50-60% compared to the original construct, as demonstrated by SDS-PAGE and Western blot analysis.

Conclusion: Optimization of the secondary structure of the translation initiation region (TIR) mRNA sequence in recombinant vectors can effectively improve target protein expression levels, offering significant value for further industrial-scale production.

Keywords: sTNFαRI; mRNA secondary structure; translation initiation region; free energy

Introduction

Tumor necrosis factor (TNF) plays a crucial role in regulating acute and chronic inflammation by binding to specific cell surface receptors [1-2]. Soluble tumor necrosis factor receptors (sTNFα-R), primarily including sTNFαRI and sTNFαRII, are naturally occurring monomeric fragments representing the extracellular portion of cell surface receptors. In vivo, they reduce TNF biological

activity and inflammatory responses by binding to transmembrane or soluble TNFα. Dalina et al. found that TNFα exerts its effects by activating TNF-R (mainly TNF-RI) and subsequently activating NF-kB transcription through a series of intracellular signaling events [3]. The soluble tumor necrosis factor receptor type II (sTNF-RII) fusion protein (Enbrel) is currently used for rheumatoid arthritis treatment [4-6], suggesting that sTNF-RI also holds potential clinical application prospects.

mRNA non-coding regions and secondary structures play vital roles in protein translation. The sequence and structure of the mRNA 5' translation initiation region (TIR) determine protein translation initiation efficiency [4]. The TIR primarily consists of four components: (1) the Shine-Dalgarno (SD) sequence, (2) the start codon, (3) the sequence between the SD sequence and start codon, and (4) enhancer elements. Numerous studies have reported that optimizing the secondary structure of the mRNA TIR region can control and enhance recombinant protein expression in *E. coli* [7-11].

In cellular environments, RNA often folds into secondary or tertiary structures that affect RNA degradation and translation initiation [6]. RNA secondary structure analysis methods can be broadly categorized into thermodynamic models, homology comparison models, and statistical models. In thermodynamic models, literature reports indicate that lower free energy in the mRNA TIR region facilitates translation initiation. Moreover, when the relative free energy is higher and nucleotide position entropy is lower at ribosomal binding sites and start codons, the probability of hairpin structure formation is reduced, indicating optimal sequences [12-15].

This study employs bioinformatics analysis based on thermodynamic models to analyze free energy and position entropy changes in nucleotide sequences, thereby determining the probability of hairpin structure formation at ribosome binding sites and start codons. By optimizing the TIR region secondary structure and screening for optimal sequences, we designed primers to mutate the mRNA translation initiation region (TIR) sequence to enhance target protein expression levels in *E. coli*.

Materials and Methods

1.1 Strains and Plasmids

The prokaryotic expression vector pET11b and host strain BL21(DE3) were purchased from Novagen.

1.2 Reagents

DNA polymerase, T4 DNA ligase, and Wizard SV Gel and PCR Clean-up System (A9282) were purchased from Promega. Plasmid mini-prep kits were obtained from Tiangen. XbaI and BamHI were from NEB. Yeast extract and tryptone were from OXOID. IPTG was from MERCK.

1.3 mRNA Sequence Design and Secondary Structure Analysis

Based on literature reports, we designed four mRNA sequences (rbs1, rbs2, rbs3, rbs4). We used RNAstructure (http://rna.urmc.rochester.edu/RNAstructureWeb/Servers/Predict1/Predict1.h to simulate the secondary structure of the 70 bp sequence upstream of the start codon (AUG) and calculate total free energy. The RNAfold Web Server (http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi) was used to analyze nucleotide position entropy and relative free energy changes to determine the probability of hairpin structure formation at ribosome binding sites and start codons, enabling selection of optimal sequences for TIR region mutation.

1.4 Construction of Recombinant Strains

1.4.1 sTNFαRI Gene Synthesis and TIR Region Mutant Sequence Amplification The sTNFαRI gene sequence was synthesized by Invitrogen. TIR region mutant sequences were introduced at the 5' end of sTNFαRI via PCR. The PCR reaction mixture contained: template 1 L, forward and reverse primers 0.5 L each, 10× PCR buffer 5 L, dNTP 1 L, Taq DNA polymerase 1 L, sterile ddH_2O 41 L, totaling 50 L. PCR conditions were: 94°C pre-denaturation for 4 min; 30 cycles of 94°C denaturation for 30 s, 58°C annealing for 30 s, 72°C extension for 30 s; final extension at 72°C for 5 min. PCR products were analyzed by 1% agarose gel electrophoresis.

1.4.2 Recombinant Engineering Strain Construction PCR products were recovered using the Wizard SV Gel and PCR Clean-up System. Both PCR products and pET11b plasmid were double-digested with XbaI and BamHI at 37°C for 3 h, followed by addition of 1 L CIAP for an additional 0.5 h. Digested products were recovered and identified by 1% agarose gel electrophoresis. Target gene fragments were ligated to vector fragments with T4 DNA ligase overnight at room temperature, then transformed into *E. coli* DH5α competent cells. Transformants were plated on LB agar containing 100 g/mL carbenicillin and incubated at 37°C overnight. Single colonies were screened by colony PCR, and positive clones were selected for plasmid extraction and double-digestion identification. Confirmed plasmids were sequenced by Genewiz. Correct plasmids were transformed into *E. coli* BL21(DE3) competent cells, plated on LB agar with 100 g/mL carbenicillin, and positive clones were preserved in glycerol at -80°C.

1.5 Small-Scale Fermentation and Induction

A small amount of recombinant glycerol stock was inoculated into LB liquid medium containing 100 g/mL carbenicillin and activated overnight at 37°C with shaking at 220 rpm. The activated culture was inoculated at 1% into fresh LB medium containing 100 g/mL carbenicillin and grown at 37°C, 260 rpm to OD_{600} 0.8. Expression was induced with 0.25 mM IPTG for 3.5 h.

1.6 Detection and Analysis of sTNFαRI Expression

Cells were harvested by centrifugation, resuspended in an equal volume of ultrapure water, and mixed with 5× protein loading buffer. Samples were vortexed and heated at 94°C for 5 min. sTNFαRI protein expression levels were detected and analyzed by 15% SDS-PAGE and Western blot. Gel imaging system was used to analyze sTNFαRI band percentages and compare with pre-optimization strains.

Results

2.1 Sequence Design

Four different mutant sequences were designed and numbered rbs1, rbs2, rbs3, and rbs4. The pre-mutation sequence was designated Ori-sTNFαRI (Table 1).

2.2 mRNA Secondary Structure Simulation Analysis

Using the minimum free energy (MFE) method, RNAstructure was employed to simulate the secondary structure of the 70 bp nucleotide sequence centered on the start codon (AUG) in the sTNFαRI mRNA 5' TIR region and calculate total free energy. Different colors marked bases to indicate the probability of stem-loop structure formation at each position. Secondary structure predictions for the four mutated sequences (Rbs1-4) and original sequence (ori-sTNFαRI) are shown in Figure 1 [Figure 1: see original paper]A-1E.

As shown in Figure 1, the original sequence (Figure 1E) exhibited a 70-80% probability (green, see reference 1F) of paired structures at both the ribosome binding site (···GAAGGAGA···) and start codon (AUG). In mutated sequences (1A-1D), the probability of base pairing at the ribosome binding site (···AGGAGG···) was lower than the original sequence for Rbs1-3, except Rbs4 which showed significantly higher pairing probability (red). Rbs1 and Rbs2 had the AUG exposed outside the hairpin structure, with Rbs1 showing lower free energy than Rbs2, making it more favorable for translation initiation. Based on this analysis, theoretically, mutating the sTNFαRI upstream TIR region to the Rbs1 sequence would be optimal.

2.3 Nucleotide Position Entropy and Relative Free Energy Analysis of Different Sequences

Literature reports indicate that predicting RNA secondary structure using minimum free energy alone has limited accuracy [16]. We therefore analyzed free energy and nucleotide position entropy changes in the 70 bp sequence centered on the start codon (AUG) using the RNAfold Webserver (Figure 2 [Figure 2: see original paper]).

Figure 2 presents relative free energy predictions for different nucleotide sequences at various positions using comprehensive methods including minimum

free energy (mfe), thermodynamic ensemble matching (pf), and centroid approaches [17], along with nucleotide positional entropy values.

Theoretically, when nucleotide sequences are located in non-pairing regions of RNA secondary structure, positional entropy is higher; when located in pairing-prone regions, entropy is lower [18-21]. The start codon (AUG) and upstream ribosome binding site (RBS) being in an unpaired state is most conducive to translational complex assembly and translation initiation [13].

As shown in Figure 2, Rbs1 sequence (Figure 2A) exhibited the highest entropy at the ribosome binding site (···AGGAGG···) with relatively low free energy, indicating low probability of hairpin structure formation (optimal). In Figures 2B-2E, positional entropy at the RBS was lower, with Figure 2D showing the lowest, indicating the highest pairing probability for Rbs4 sequence (worst). Regarding the start codon (AUG) position, Rbs1 and Rbs4 showed the lowest relative free energy. Although nucleotide positional entropy was also low near the stem-loop structure, both sequences remained unpaired (Figure 1), favoring translation initiation. However, Rbs4 showed high pairing probability at the upstream ribosome binding site, which is unfavorable for translational complex assembly.

Comprehensive analysis of relative free energy and nucleotide positional entropy for all five sequences predicts that the Rbs1 sequence would yield the best expression results.

2.4 Primer Synthesis

PCR mutagenesis primers were synthesized by Genewiz (Table 2).

2.5 Identification of sTNFαRI Gene Amplification Products

Using rbs1, rbs2, rbs3, and rbs4 as forward primers and PP6 as reverse primer, PCR amplification products were analyzed by 1% agarose gel electrophoresis. As shown in Figure 3 [Figure 3: see original paper], bands corresponding to the positive control (approximately 330 bp) were visible in all four mutated PCR products.

2.6 Double Enzyme Digestion and Sequencing Identification of Recombinant pET11b-Rbs1/2/3/4-sTNFαRI Plasmids

Recombinant pET11b-Rbs1/2/3/4-sTNFαRI plasmids were double-digested with XbaI and BamHI, and fragments of approximately 330 bp were observed by 1% agarose gel electrophoresis (Figure 4 [Figure 4: see original paper]). Sequencing results confirmed consistency with the expected sequences, indicating successful construction of all recombinant pET11b-rbs1/2/3/4-sTNFαRI plasmids.

2.7 Identification and Expression Comparison of Recombinant sTNFαRI Protein in BL21(DE3) Cells

Harvested cells from small-scale fermentation were analyzed by 4-20% gradient SDS-PAGE and Western blot (Figure 5 [Figure 5: see original paper]). The pET11b-rbs1-sTNFαRI/BL21(DE3) strain showed the highest sTNFαRI expression (molecular weight 12.9 kDa) after TIR region mutation, consistent with predictive results.

Image Lab 4.0 software analysis of sTNFαRI band percentages after induction of different mutant recombinant strains (Figure 6 [Figure 6: see original paper]) revealed significantly improved expression with the optimized rbs1 sequence (Figure 7 [Figure 7: see original paper]). Calculations showed a 58.4% increase compared to pre-optimization expression levels.

Discussion

Similar to the TNFαRII recombinant protein (Enbrel), TNFαRI recombinant protein may have clinical applications in treating autoimmune diseases such as psoriasis and rheumatoid arthritis [22]. Currently, approximately 30% of therapeutic recombinant proteins are produced through *E. coli* expression [7], where expression levels directly impact downstream processing and production costs, making it a critical factor in product development.

Multiple factors influence heterologous protein expression in *E. coli*, including codon preference, promoter strength, mRNA stability and secondary structure, and translation termination signals [23-24]. This study focused on how mRNA translation initiation region sequence and secondary structure affect TNFαRI expression levels.

We designed different primers to alter the mRNA TIR region sequence, thereby changing hairpin structure position and length in the mRNA secondary structure and modifying mRNA free energy and entropy values to enhance TNFαRI expression. Results demonstrated that when the total free energy of the TNFαRI mRNA 5' TIR region was lowest and the probability of hairpin structure formation at the ribosome binding site and start codon was minimized, translation initiation was facilitated, thereby improving TNFαRI inclusion body expression. After TIR region optimization, sTNFαRI inclusion body expression increased by 58.4% compared to pre-optimization levels.

This study optimized the mRNA 5' TIR secondary structure of the target gene to enhance sTNFαRI expression in *E. coli*, providing a solid foundation for future drug development targeting TNF. Additionally, it offers novel strategies for improving recombinant protein expression levels.

References

[1] Arend WPD, Dayer JM. Inhibition of the production and effects of

interleukin-1 and tumor necrosis factor alpha in rheumatoid arthritis. *Arthritis Rheum*, 1995, 38:151-160.

[2] Alsalameh S, Winter K, Al-Ward R, et al. Distribution of TNF-alpha, TNF-R55 and TNF-R75 in the rheumatoid synovial membrane: TNF receptors are localized preferentially in the lining layer; TNF-alpha is distributed mainly in the vicinity of TNF receptors in the deeper layers. *Scand J Immunol* 1999, 49:278-285.

[3] de Oliveira D C, Hastreiter A A, Mello A S, et al. The effects of protein malnutrition on the TNF-RI and NF-kB expression via the TNF-αsignaling pathway. *Cytokine*. 2014, 69(2): 218-225.

[4] Moreland L W, Baumgartner S W, Schiff M H, et al. Treatment of rheumatoid arthritis with a recombinant human tumor necrosis factor receptor (p75)-Fc fusion protein. *N Engl J Med*, 1997, 337:141-148.

[5] Berard R A, Laxer R M. Etanercept(Enbrel) in the treatment of Juvenile idiopathic arthritis. *Expert Opin Biol Ther.*, 2013, 13(11): 1623-1630.

[6] Belmellat N, Semerano L, Segueni N, et al. Tumor Necrosis Factor-Alpha Targeting can protect against arthritis with low sensitization to infection. *Front Immuno*, 2017, 14:8:1533.

[7] Huang C J, Lin H, Yang X. Industrial production of recombinant therapeutics in *Escherichia coli* and its recent advancements. *Journral of Inductrial Microbiology & Biothechnology*, 2012, 39(3): 383.

[8] Alibolandi M, Mirzahoseini H, Abad MAK, Azami movahed M. High level expression of human basic fibroblast growth factor in *Escherichia coli*: evaluating the eVect of the GC content and rare codons within the Wrst 13 codons. *Afr J Biotechnol*, 2010, 9(16):2456–2462.

[9] Simmons L C, Yansura D G. Translational level is a critical factor for the secretion of heterologous proteins in *Escherichia coli*. *Nat Biotechnol*, 1996, 14(5):629–634.

[10] Vimberg V, Tats A, Remm M. Translation initiation region sequence preferences in *Escherichia coli*. *BMC Mol Biol*, 2007, 8:100.

[11] Behloul N, Wei W, Baha S, et al. Effects of mRNA secondary structure on the expression of HEV ORF2 proteins in *Escherichia coli*. *Microb Cell Fact*, 2017, 16(1):200.

[12] Zhang H W, Yang Y C, Lu Z. From sequence to structure: RNA secondary structure prediction methods and the applications. *Chinese Bulletin of Life Sciences*, 2014, 26(3):219-218.

[13] Garcia-Martin J A, Clote P. RNA Thermodynamic Structural Entropy. *PloS ONE*, 2015, 10(11): e037859.

[14] Zhang W C, Xiao W H, Wei H M, et al. mRNA secondary structure at start AUG codon is a key limit factor for human protein expression in *Escherchia coli*. *Biochemical and Biophysical Research Communications*, 2006, 349:69-78.

[15] Seetin M G, Mathews D H. RNA structure prediction: an overview of methods. *Methods in Molecular Biology*, 2012, 905:99-122.

[16] Li X. The simulation Analysis of Secondary Structure Prediction Optimization Model. *Computer Simulation*, 2016, 7: 323-326.

[17] Hamada M, Kiryu H, Sato K, et al. Prediction of RNA secondary structure using generalized centroid estimators. *Bioinformatics*, 2009, 25(4): 465-473.

[18] Mathews D H, Disney M D, Childs J L, et al. Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA Secondary Structure. *PNAS*, 2004, 101(19):7287-7292.

[19] Eren A M, Morrison H G, Lescault P J, et al. Minimum entropy decomposition: Unsupervised oligotyping for sensitive partitioning of high-throughput marker gene sequences. *The ISME Journal*, 2015, 9:968-979.

[20] Dotu I, Garcia-Martin J A, Slinger B L, et al. Complete RNA inverse folding: computational design of functional hammerhead ribozymes. *Nucleic Acids Research*, 2014, 42(18): 11752-11762.

[21] Zhang Y P, Wang P, Yan M D. An Entropy-Based Position Projection Algorithm for Motif Discovery. *BioMed Research International*, 2016, ID9127474.

[22] Bao C H, Wu L Y, Wu, H G et al. Moxibustion Inhibits Apoptosis and Tumor Necrosis Factor-Alpha/Tumor Necrosis Factor Receptor 1 in the Colonic Epithelium of Crohn' s Disease Model Rats. *Digestive Diseases & Scences*, 2012, 57(9):2286-2295.

[23] Farshadpour F, Taherkhani R, Makvandi M, et al. Condon-optimized expression and purification of truncated ORF2 protein of Hepatitis E Virus in *Escherichia coli*. *JUndishapur J Microbiol*. 2014, 7(7): e11261.

[24] Molina-Garcia L, Ciralda R. Enabling stop codon read-through translation in bacteria as a probe for amyloid aggregation. *Sci Rep*. 2017, 7(1):1908.

Note: Figure translations are in progress. See original paper for figures.