Protein Expression and Purification 138 (2017) 46–55

Contents lists available at ScienceDirect

Protein Expression and Purification

journal homepage: www.elsevier.com/locate/yprep

Insulin chains as efficient fusion tags for prokaryotic expression of short peptides

Ligang Deng, Xiaoying Xue, Cangjie Shen, Xiaohan Song, Chunyang Wang, Nan Wang*

Beijing Key Laboratory of New Drug Mechanisms and Pharmacological Evaluation Study, Institute of Materia Medica, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, 100050, China

ARTICLE INFO

Article history:
Received 23 February 2017
Received in revised form
22 May 2017
Accepted 30 June 2017
Available online 13 July 2017

Keywords:
Fusion tag
Inclusion bodies
Insulin chains
Mini-proinsulin
Short peptides

ABSTRACT

Insulin chains are usually expressed in *Escherichia coli* as fusion proteins with different tags, including various low molecular weight peptide tags. The objective of this study was to determine if insulin chains could facilitate the recombinant expression of other target proteins, with an emphasis on low molecular weight peptides. A series of short peptides were fused to mini-proinsulin, chain B or chain A, and induced for expression in *Escherichia coli*. All the tested peptides including glucagon-like peptide 1 (GLP-1), a *C*-terminal extended GLP-1, oxyntomodulin, enfuvirtide, linaclotide, and an unstructured artificial peptide were expressed with reasonable yields, identified by Tricine-SDS-PAGE and immunoblotting. All recombinant products were expressed in inclusion bodies. The effective accumulation of products was largely attributed to the insoluble expression induced by fusion with insulin chains, and was confirmed by the fusion expression of transthyretin. Insulin chains thus show promise as efficient fusion tags for mass production of heterologous peptides in prokaryotes.

1. Introduction

Recombinant production of proteins by prokaryotic expression has been a well-established method in industry as well as in laboratory practice. However, rapid degradation by proteases makes direct expression of short peptides in *Escherichia coli* difficult. Fusion technology has been used for the expression of short peptides, as it has been shown to decrease degradation and improve production of proteins [1]. Commonly used fusion tags such as glutathione S-transferase (GST) [2], small ubiquitin-like modifier (SUMO) [3], thioredoxin (TRX) [4,5], maltose binding protein (MBP) [6], ThiS [7], ketosteroid isomerase [8] and purF fragment [9] have high molecular weights, and although they improve the total expression of fusion proteins, they reduce the effective production efficacy of low molecular weight short peptide, which constitutes only a small portion in the fusion molecule.

The matured peptide hormone insulin consists of two separate chains joined by disulfide bridges. Chain A and chain B are short peptides comprising 21 and 30 amino acid residues respectively. They are originally produced in the pancreatic beta cells in the form of proinsulin as a single peptide B-C-A with the C peptide between them. The fusion strategy was initially employed in recombinant insulin production, wherein chain B and chain A were fused to β-galactosidase [10] and Trp-LE [11], respectively, and separately expressed followed by *in vitro* processing to form the insulin molecule. This two-chain approach suffered from a poor yield of single chains and low efficiency of processing, and has been replaced by an efficient single-chain approach, wherein chains B and A are expressed as a single peptide in the form of proinsulin [12–15] or mini-proinsulin with a shortened C-peptide [16–18], followed by efficient renaturation and processing. Various fusion tags have been used for single-chain expression, including high molecular weight β-galactosidase with either the complete sequence (110 kDa) [19,20] or a partial sequence (58 kDa) [19]. Tags with mid-molecular weight such as Trp-LE (22 kDa) [21,22], IgG-binding domains (ZZ) (15 kDa) [12] and B-domain of protein A (7.3 kDa) [14] have also been used. Low-molecular weight tags such as *N*-terminus of interleukin-2 (IL-2) (2.4 kDa) [23] and oligopeptides with or without His-tag [14,18,24–27] have shown previous success, forming the motivation for this study. With this background, the objective of this work thus was to determine the efficacy of low molecular weight short peptide insulin chains as fusion tags for the expression of heterologous peptides.

* Corresponding author. Tel: +86 10 83172984.
E-mail address: wangnan@imm.ac.cn (N. Wang).

http://dx.doi.org/10.1016/j.pep.2017.06.019
1046-5928/

2. Materials and methods

2.1. Materials

The pQE80L vector plasmid was purchased from Qiagen (Hilden, Germany). All restriction enzymes and the T4 DNA ligase were obtained from TaKaRa (Dalian, China). Pfu DNA polymerase and LA Taq DNA polymerases were obtained from Vigorous Biotechnology (Beijing, China). Oligonucleotides were obtained from Invitrogen (Shanghai, China). Anti-His-tag mouse monoclonal antibody and goat anti-mouse IgG (HRP conjugated) were acquired from CoWin Biotech (Beijing, China). Tricine was purchased from Amresco (Cleveland, USA). Isopropyl β-D-1-thiogalactopyranoside (IPTG) was from Solarbio (Beijing, China). Polyvinylidene difluoride (PVDF) membrane was from Millipore (Boston, USA). α-Cyano-4-hydroxycinnamic acid (CHCA) was from Sigma (St. Louis, USA). Dithiothreitol (DTT) was from Merck (Darmstadt, Germany). Trypsin and carboxypeptidase B were from Yaxin Biotechnology (Shanghai, China). *E. coli* TG1 cells were used for cloning, maintenance, and propagation of plasmids, and also for expression. *E. coli* BL21 (DE3) pLysS cells were used as hosts for protein expression studies. Both strains were purchased from ComWin Biotech (Beijing, China) and cultivated in LB broth under appropriate selective conditions.

2.2. Construction of expression vectors

Standard molecular biology techniques were applied for cloning the vectors [28]. Genes for insulin chain A and chain B were synthesized as described previously [7]. All other genes for target proteins/peptides barring the unstructured artificial peptide were designed with a codon bias for expression in *E. coli* host cells. Gene fusions were generated by overlap PCR or restricted fragment ligation. All clones used were confirmed by DNA sequencing (Invitrogen). Expression vectors were based on the pQE80L backbone with 6-His at 5′ fusion site. Target peptides reported in this study were expressed with insulin chains fused at the *N*-terminus. The amino acid sequences of recombinant fusion peptides with predicted molecular weights were listed in Table 1.

2.3. Expression of recombinant proteins

For mid-scale expression, overnight *E. coli* cultures were subcultured in 20–30 ml fresh LB broth containing 100 μg/ml ampicillin at 37 °C in shaking flasks until the mid-log phase was reached with appropriate optical density (OD) values at 600 nm. Expression of fusion proteins was induced by IPTG. Optimized expression conditions specific to recombinant peptides were used to evaluate the expression efficacy, as listed in Table 2. Then 14 ml of cells were harvested (yielding about 0.11 g at 4 h, or 0.18 g at 20 h of wet cell pellet), resuspended in PBS containing 1% Triton X-100, subjected to freezing and thawing cycles three times, and then lysed by sonication. Inclusion bodies were separated from soluble fractions by centrifugation (14,000×*g* for 15 min at 4 °C), and washed further with PBS containing 1% Triton X-100 and 5 mM EDTA.

For large-scale expression of B′-LN2, 1.5 L of LB were used. The five flasks were shaked to grow for 20 h at 37 °C with IPTG of 0.1 mM. Harvested cells of about 5.3 g of wet pellet were resuspended in 30 ml PBS containing 1% Triton X-100 and then lysed as described above. Inclusion bodies were washed with PBS buffer of 10 ml containing 1% Triton X-100 and 5 mM EDTA.

2.4. Electrophoresis and Western blot

Total cell lysates induced with or without IPTG as well as the separated fractions of induced cells at the same dilution were mixed with Laemmli buffer, heated in a boiling water bath for 20 min in order to break down the DNA and decrease the viscosity of samples. Samples were analyzed in duplicate gels by 13% (or 16.5% as indicated) Tricine-SDS-PAGE as described previously [29] or in 15% SDS-PAGE. One of the gels was stained with Coomassie brilliant blue R-250, the other was electroblotted onto PVDF membranes. His-tagged fusion proteins were identified by immunoblotting using anti-His-tag mouse monoclonal antibody and goat anti-mouse IgG (HRP conjugated) as primary and secondary antibodies, respectively. Chemiluminescence was recorded per the supplier's protocol (Vigorous Biotechnology, Beijing, China).

2.5. Protein quantification

Target monomer bands of the expressed fusion proteins were quantified by densitometric analysis using the QuantiScan Software (Biosoft, Cambridge, UK), using a predefined amount of bovine serum albumin (BSA) as the standard. Recombinant productivity was estimated from the mid-scale expression. Results from batches of independent production of the same fusion protein were averaged for the estimation, and represented as the mean ± SD (n = 3).

2.6. Mass spectrometry

Acidified samples were mixed with CHCA solution (30 g/L in 70% acetonitrile and 30% methanol, with 0.1% TFA) at a ratio of 1:1 (v/v), spotted onto the sample plate and air-dried. Mass spectra of the samples were acquired using a MALDI-TOF/TOF Analyzer 4800 Plus (Applied Biosystem, Foster City, CA, USA) in the linear or reflection mode under a laser intensity of ~4800 W/cm^2.

2.7. Proteolytic processing

Inclusion bodies of B′-LN2 were dissolved in 20 mM Tris-HCl (pH 8.0) containing 8 M urea and 5 mM DTT, incubated for 20 min at room temperature. After centrifuged (14,000×*g* for 10 min at 4 °C), supernatants were diluted 20 times with 20 mM Tris-HCl (pH 8.0) and subjected to protease treatment. B′-LN2 was first digested by trypsin (4 mg/L) at 37 °C overnight to generate linaclotide with an additional 'R' at the *C*-terminus. Then the 'R' residue was removed by carboxypeptidase B (0.5 mg/L) for 30 min.

2.8. Reverse phase chromatography

The solution of B′-LN2 treated with trypsin only was oxidized completely by air for 24 h, followed by adding acetic acid to a final concentration of 40 mM. The supernatant obtained by centrifugation was loaded onto C18 column. The column was washed with H_2O containing 40 mM acetic acid and eluted stepwise by 25% and 50% acetonitrile all containing 40 mM acetic acid. The fractions monitored at OD_{280} were pooled and analyzed by mass spectrometry. The fraction containing target peptide was evaporated and lyophilized. Further purification was performed on Resource RPC column (6.4 × 100 mm, GE Healthcare, Uppsala, Sweden) using AKTA Purifier 10 FPLC system (GE Healthcare, Uppsala, Sweden). The lyophilized powder was dissolved in 2% acetonitrile containing 0.065% TFA at a concentration of 2.5 mg/ml. After the sample was loaded, the column was washed with 2% acetonitrile containing 0.065% TFA and eluted with a linear gradient of 2%–50% acetonitrile containing 0.065% TFA. The effluents were monitored at OD_{214}. The fractions corresponding to the elution peaks were pooled and analyzed by mass spectrometry.