Protein Expression and Purification 132 (2017) 44–49

Contents lists available at ScienceDirect

Protein Expression and Purification

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

Production of recombinant proteins in *Escherichia coli* tagged with the fusion protein CusF3H+

Teresa Vargas-Cortez [a], Jose Ruben Morones-Ramirez [a, b], Isaias Balderas-Renteria [a, b], Xristo Zarate [a, b, *]

[a] Universidad Autonoma de Nuevo Leon, Facultad de Ciencias Quimicas, San Nicolas de los Garza, NL 66455, Mexico
[b] Universidad Autonoma de Nuevo Leon, Centro de Investigacion en Biotecnologia y Nanotoxicologia, Facultad de Ciencias Quimicas, Parque de Investigacion e Innovacion Tecnologica, Apodaca, NL 66629, Mexico

ARTICLE INFO

Article history:
Received 14 December 2016
Received in revised form
7 January 2017
Accepted 10 January 2017
Available online 11 January 2017

Keywords:
Fusion protein
Affinity tag
CusF3H+
Recombinant proteins
Escherichia coli

ABSTRACT

Recombinant protein expression in the bacterium *Escherichia coli* still is the number one choice for large-scale protein production. Nevertheless, many complications can arise using this microorganism, such as low yields, the formation of inclusion bodies, and the requirement for difficult purification steps. Most of these problems can be solved with the use of fusion proteins. Here, the use of the metal-binding protein CusF3H+ is described as a new fusion protein for recombinant protein expression and purification in *E. coli*. We have previously shown that CusF produces large amounts of soluble protein, with low levels of formation of inclusion bodies, and that proteins can be purified using IMAC resins charged with Cu(II) ions. CusF3H+ is an enhanced variant of CusF, formed by the addition of three histidine residues at the *N*-terminus. These residues then can bind Ni(II) ions allowing improved purity after affinity chromatography. Expression and purification of Green Fluorescent Protein tagged with CusF3H+ showed that the mutation did not alter the capacity of the fusion protein to increase protein expression, and purity improved considerably after affinity chromatography with immobilized nickel ions; high yields are obtained after tag-removal since CusF3H+ is a small protein of just 10 kDa. Furthermore, the results of experiments involving expression of tagged proteins having medium to large molecular weights indicate that the presence of the CusF3H+ tag improves protein solubility, as compared to a His-tag. We therefore endorse CusF3H+ as a useful alternative fusion protein/affinity tag for production of recombinant proteins in *E. coli*.

1. Introduction

The importance of fusion proteins for recombinant protein production in *Escherichia coli* has been recognized for quite some time now [1]; nevertheless, it remains crucial to continue exploring novel fusion proteins/affinity tags for their suitability in the production of particular target proteins. The periplasmic protein CusF is part of the CusCBFA efflux complex system involved in the copper and silver resistance mechanism of *E. coli* [2]. Recently, we presented results demonstrating that CusF is a suitable fusion protein for recombinant protein production *in vitro* using *E. coli*-based cell-free expression systems and *in vivo*, because it improves protein solubility, to a degree that is similar to the performance of maltose-binding protein (MBP) and glutathione S-transferase (GST). Use of CusF is particularly convenient because of its molecular weight, just 10 kDa, is considerably smaller than that of MBP (42 kDa) and GST (26 kDa), and this feature leads to higher protein yields after tag removal [3]. In our work, Green Fluorescent Protein (GFP) tagged with CusF (CusF_GFP), purified by immobilized metal affinity chromatography (IMAC) functionalized with Cu(II) ions, showed only 50% purity after chromatography, possibly due to the presence of other histidine-rich proteins in the *E. coli* lysate that bind to the

Abbreviations: CusF, Cation efflux system protein; CusF3H+, mutant of CusF; GFP, Green Fluorescent Protein; MBP, maltose-binding protein; GST, glutathione S-transferase; IMAC, immobilized-metal affinity chromatography; dNTP, deoxynucleotide triphosphate; BSA, bovine serum albumin; PIF3, phytochrome-interacting factor 3; PixE, PixD response regulator-like protein; AXR3, auxin-responsive protein IAA17; LB, Luria-Bertani broth; IPTG, Isopropyl β-D-1-thiogalactopyranoside; OD_{600}, optical density at 600 nm; SDS-PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis; Tris, Tris(hydroxymethyl)-aminomethane.

* Corresponding author. Universidad Autonoma de Nuevo Leon, Facultad de Ciencias Quimicas, San Nicolas de los Garza, NL 66455, Mexico.
E-mail address: xristo.zaratekl@uanl.edu.mx (X. Zarate).

http://dx.doi.org/10.1016/j.pep.2017.01.006
1046-5928/

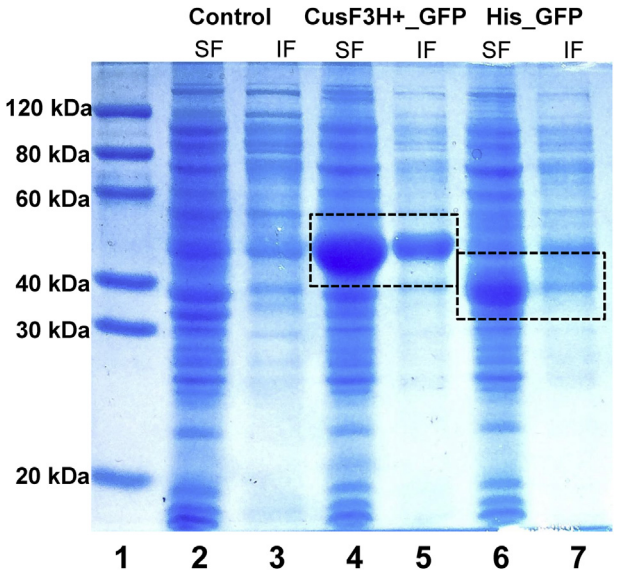

Fig. 1. 12% SDS-PAGE analysis of small-scale expression of GFP constructs. Lane 1: protein marker; lanes 2 and 3: soluble and insoluble fractions of *E. coli* BL21(DE3) lysate control; lanes 4 and 5: soluble and insoluble fractions of CusF3H+_GFP expression; lanes 6 and 7: soluble and insoluble fractions of His-tag_GFP. Calculated molecular weights for CusF3H+_GFP: 36.7 kDa and His-tag_GFP: 26.0 kDa. Estimated soluble protein yields for CusF3H+_GFP: 36% and His-tag_GFP: 31%.

column as well. Another difficulty that can be encountered by the use of Cu(II) chromatography is that the target protein may show intrinsic affinity for copper, complicating the removal of the CusF tag from the protein of interest, which could affect the final protein yield [3].

In order to improve purity and tag removal, in this work we altered the primary sequence of CusF by specifically adding three histidine residues to the *N*-terminus, hence the name CusF3H+. We tested its ability to bind IMAC resins charged with Ni(II) ions for recombinant protein purification. GFP tagged with wild-type CusF bound poorly to the immobilized phase, protein being lost during the wash steps, even in the absence of imidazole in the buffer. On the contrary, CusF3H+_GFP bound tightly to the Ni(II) resin and no protein was lost during washing with buffers containing 10 mM imidazole, thus achieving higher purity after protein elution compared to Cu(II) IMAC. Finally, expression levels of CusF3H+ tagged proteins were compared to those of the corresponding His-tag constructs. The results indicate that CusF3H+ improves solubility even for proteins as large as 57 kDa.

2. Materials and methods

2.1. DNA constructs

CusF is a periplasmic protein, the first five amino acids of the mature polypeptide after signal sequence proteolysis being NEHHH. The forward primer (full sequence below) used to amplify CusF and produce CusF3H+ DNA included an NdeI restriction site (CATATG) which contains the start codon, followed by three histidine codons, and finally the same DNA sequence from wild type CusF from its first histidine. It was decided to use the CAC codon in the primer to add the histidines in order to simplify annealing to the CusF template (which uses CAT for the three histidines). CAC and CAT codons are almost equally used in *E. coli* for the amino acid histidine, so translation rates should not be affected [4]. CusF3H+ was amplified using the primers 5' - AGTCAGTCA**CATATG**CACCACCACCATCATCATGAAACCATGAGCGAAGTAC - 3' (NdeI, forward) and 5' – ATGCATGCA**GGTACC**CTGGCTGACTTTAATATCCTGTAA – 3' (KpnI, reverse). For the amplification, the 50 μL reaction comprised 10 ng of CusF DNA template, 60 pmol of each primer, 1 μL of dNTP mix (10 mM), and 2 units of *Vent* DNA polymerase (New England Biolabs) in 1× ThermoPol reaction buffer. pET30a vector (EMD Millipore) was linearized with NdeI and KpnI (New England Biolabs), and CusF3H+ was digested with the same enzymes and inserted into the vector. Target proteins were cloned using NcoI and XhoI restriction sites in order to place an enterokinase recognition sequence between CusF3H+ and the target protein for tag removal. The reporter protein for expression and purification was the S65T mutant of GFP [5]; other target proteins were PixE, the response regulator of PixD from *Synechocystis* sp [6], and the *Arabidopsis thaliana* proteins PIF3, phytochrome-interacting factor 3 [7], and AXR3, auxin-responsive protein IAA17 [8]. Constructs for GFP, PixE, and PIF3 tagged with an *N*-terminal His-tag instead of CusF3H+ were made for protein expression comparisons.

2.2. Protein expression

DNA constructs were transformed into the *E. coli* strain BL21(DE3). For small-scale expression experiments, 2 mL of Luria-Bertani (LB) broth with kanamycin at a final concentration of 30 μg/mL were inoculated with a single colony and incubated at 37 °C and 200 rpm until an OD_{600} between 0.4 and 0.6 was reached. Protein expression was induced with IPTG to a final concentration of 0.1 mM, and the cells were incubated at 25 °C and 220 rpm for 16 h. The cells were harvested and resuspended in 100 μL 1× SDS-PAGE sample buffer, incubated in boiling water for 10 min, and then centrifuged for 10 min at 13000 rpm. The supernatant was used to analyze soluble protein content. The pellet was resuspended in an 8 M urea buffer and boiled for 10 min, and then centrifuged at the same speed; the supernatant was used for the analysis of insoluble protein content. Both protein fractions were analyzed by SDS-PAGE electrophoresis and soluble protein was semi-quantitated by densitometry using the ImageJ software (version 1.49) from the National Institute of Health [9]. The software takes into consideration the intensity of all protein bands present in a selected lane by plotting a peak for each one; the area of the peak is proportional to the intensity of the band. Based on the area of each peak, a percentage from the total area can be assigned for each protein peak/band, thus, soluble protein yields were calculated for each of the target proteins. For large-scale protein expression, CusF_GFP and CusF3H+_GFP transformed cells were grown in baffled flasks until OD_{600} reached 0.5, expression was induced by adding IPTG up to 0.1 mM, and cells were incubated for 16 h at 25 °C and 220 rpm.

2.3. Protein purification

CusF_GFP cells were harvested by centrifugation at 4 °C. The cells were resuspended in ice-cold lysis buffer (50 mM Tris, 500 mM NaCl, pH 8.0), and lysed using a bead-beater and 0.1 mm glass beads (BioSpec Products). The lysate was clarified by centrifugation at 4 °C, 15000 rpm for 20 min. Purification was done using the ÄKTA Primer Plus FPLC (GE Healthcare). The supernatant was loaded into a HisTrap FF 1 mL-column charged with Ni(II) previously equilibrated with lysis buffer. After loading with the CusF_GFP lysate, the column was washed with 10 column-volumes of lysis buffer, then eluted with elution buffer (50 mM Tris-HCl, 500 mM NaCl, 50 mM imidazole, pH 8.0). Purification of CusF3H+_GFP was treated differently, the lysis buffer including 5 mM imidazole, the washing buffer 10 mM imidazole, and the elution buffer containing 250 mM imidazole. Elution fractions (0.5 mL) were collected and analyzed by SDS-PAGE. All protein