



ELSEVIER

Contents lists available at ScienceDirect

Data in Brief

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

Data Article

Recombinant protein production data after expression in the bacterium *Escherichia coli*

J. Enrique Cantu-Bustos, Kevin D. Cano del Villar,
Teresa Vargas-Cortez, Jose Ruben Morones-Ramirez,
Isaias Balderas-Renteria, Xristo Zarate*

Universidad Autonoma de Nuevo Leon, Facultad de Ciencias Quimicas, Av. Universidad s/n,
Ciudad Universitaria, San Nicolas de los Garza, Nuevo Leon 66451, Mexico

ARTICLE INFO

Article history:

Received 13 January 2016

Received in revised form

12 February 2016

Accepted 26 February 2016

Available online 4 March 2016

Keywords:

Fusion protein

Affinity tag

Escherichia coli

CusF

GST

ABSTRACT

Fusion proteins have become essential for the expression and purification of recombinant proteins in *Escherichia coli*. The metal-binding protein CusF has shown several features that make it an attractive fusion protein and affinity tag: "Expression and purification of recombinant proteins in *Escherichia coli* tagged with the metal-binding protein CusF" (Cantu-Bustos et al., 2016 [1]). Here we present accompanying data from protein expression experiments; we tested different protein tags, temperatures, expression times, cellular compartments, and concentrations of inducer in order to obtain soluble protein and low formation of inclusion bodies. Additionally, we present data from the purification of the green fluorescent protein (GFP) tagged with CusF, using Ag(I) metal affinity chromatography.

© 2016 Elsevier Inc.. Published by Elsevier Inc. This is an open access article under the CC BY license

(<http://creativecommons.org/licenses/by/4.0/>).

Specifications table

Subject area	Molecular Biology
More specific subject area	Protein expression and purification

DOI of original article: <http://dx.doi.org/10.1016/j.pep.2016.01.007>

* Corresponding author.

E-mail address: xristo.zaratekl@uanl.edu.mx (X. Zarate).

<http://dx.doi.org/10.1016/j.dib.2016.02.074>

2352-3409/© 2016 Elsevier Inc.. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Type of data	<i>Pictures. Images of gel electrophoresis.</i>
How data was acquired	<i>Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE)</i>
Data format	<i>Raw</i>
Experimental factors	<i>Fusion proteins: GST, CusF, and SmbP. Inducer concentration, time and temperature during protein expression. Metal affinity chromatography with silver ions Ag(I). Periplasmic or cytoplasmic protein expression.</i>
Experimental features	<i>Production levels of soluble protein and inclusion bodies, protein purity.</i>
Data source location	<i>San Nicolas de los Garza, Nuevo Leon, Mexico</i>
Data accessibility	<i>Data is within this article</i>

Value of the data

- The data shows that the production of soluble recombinant proteins in the bacterium *Escherichia coli* can be improved by varying the fusion protein, time, temperature, and concentration of inducer during protein expression.
- The data indicates that recombinant proteins, exemplified here with the red fluorescent protein (RFP), can be expressed in the cellular periplasm when tagged with full-length CusF and SmbP.
- Protein purification data shows that CusF-tagged proteins may be purified using immobilized metal affinity chromatography with Ag(I) ions instead of the most common Cu(II) or Ni(II).

1. Data

Fig. 1 shows the SDS-PAGE analysis comparing soluble and insoluble protein content for the proteins LovR and SHY2 tagged with two different protein tags and expressed at different conditions. Additionally, the electrophoretic analysis in Fig. 2 compares the content of inclusion bodies.

Recombinant proteins tagged with the protein tags CusFp and SmbPp, containing their signal sequences, are exported to the cell's periplasm. Fig. 3 shows an image of *E. coli* BL21(DE3) cells after RFP expression and the electrophoretic analysis of the periplasmic lysates.

Fig. 4 shows pictures of the synthesized silver chromatographic media before and after incubation with the *E. coli* lysate expressing green fluorescent protein tagged with CusF. Fig. 5 shows the SDS-PAGE analysis of the purification steps, it shows the protein content in the flow-through (the lysate after incubation with the Ag(I) resin), and two elutions steps with 160 mM methionine.

2. Experimental design, materials and methods

2.1. DNA constructs

Full-length CusF (CusFp, for periplasmic expression) was amplified with primers 5'-AGTCAGTCA-CATATGAAAAAGCACTGCAAGTCG-3' (NdeI, forward) and 5'-ATGCATGCAGGTACCCCTGGCTGACTT-TAATATCCTGTAA-3' (KpnI, reverse). The 50 μ L reaction comprised 10 ng of template DNA, 60 pmol of each primer, 1.5 μ L of 10 mM dNTPs mix, and 2 units of Vent DNA polymerase (New England Biolabs) in 1 \times ThermoPol reaction buffer. The thermocycler conditions were 95 $^{\circ}$ C for 2 min; 30 cycles of 95 $^{\circ}$ C–1 min, 59 $^{\circ}$ C–1 min, 72 $^{\circ}$ C–1 min; and a final extension at 72 $^{\circ}$ C for 10 min. Amplification of CusF lacking the signal sequence (for cytoplasmic expression) was done with forward primer 5'-AGTCAGTCA-CATATGGCTAACGAACATCATGAAAC-3' (NdeI) and the same reverse primer and thermocycler conditions as before. pET30a vector was linearized with NdeI and KpnI, and the CusF

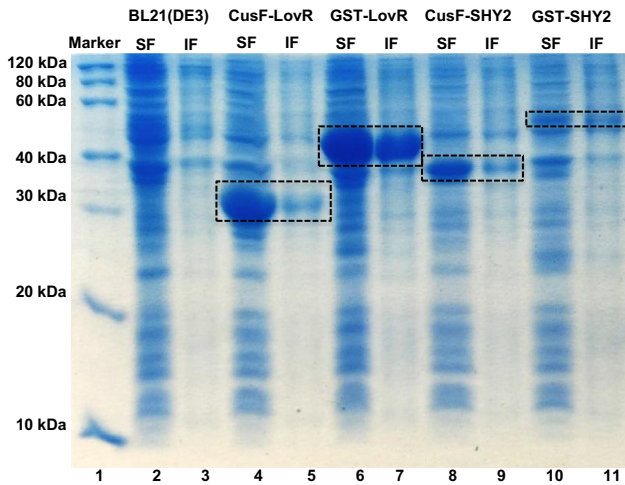


Fig. 1. 15% SDS-PAGE analysis of small-scale expression for the LovR and SHY2 constructs tagged with CusF and GST at 37 °C, 1 mM IPTG, 4 h. Lane 1: protein marker; lanes 2 and 3: soluble and insoluble fractions of *E. coli* BL21(DE3) lysate control; lanes 4 and 5: soluble and insoluble fractions of CusF-LovR; lanes 6 and 7: soluble and insoluble fractions of GST-LovR; lanes 8 and 9: soluble and insoluble fractions of CusF-SHY2; lanes 10 and 11: soluble and insoluble fractions of GST-SHY2. Calculated molecular weights for CusF-LovR: 27.4 kDa, GST-LovR: 43.6 kDa, CusF-SHY2: 33.1 kDa, and GST-SHY2: 49.1 kDa.

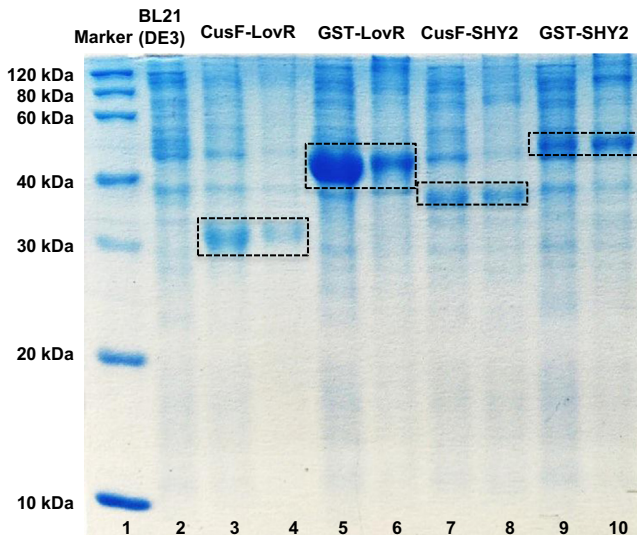


Fig. 2. 15% SDS-PAGE analysis of the insoluble fractions (inclusion bodies) produced from small-scale expressions at 37 °C and 25 °C for the LovR and SHY2 constructs tagged with CusF and GST. Lane 1: molecular weight marker; lane 2 *E. coli* BL21(DE3) background control; lanes 3 and 4: insoluble fractions of CusF-LovR at 37 °C and 25 °C respectively; lanes 5 and 6: insoluble fractions of GST-LovR at 37 °C and 25 °C respectively; lanes 7 and 8: insoluble fractions of CusF-SHY2 at 37 °C and 25 °C respectively; lanes 9 and 10: insoluble fractions of GST-SHY2 at 37 °C and 25 °C respectively.

insert was ligated into it following digestion with the same enzymes. The gene for GST was cloned with the same restriction sites for comparative studies. Target proteins were cloned using NcoI and XhoI restriction sites, in order to place the enterokinase recognition sequence between CusF and the target protein for tag removal. The target proteins expressed in the cellular cytoplasm (using the CusF construct) were LovR from *Caulobacter crescentus* [2], the *Arabidopsis thaliana* protein Short Hypocotyl

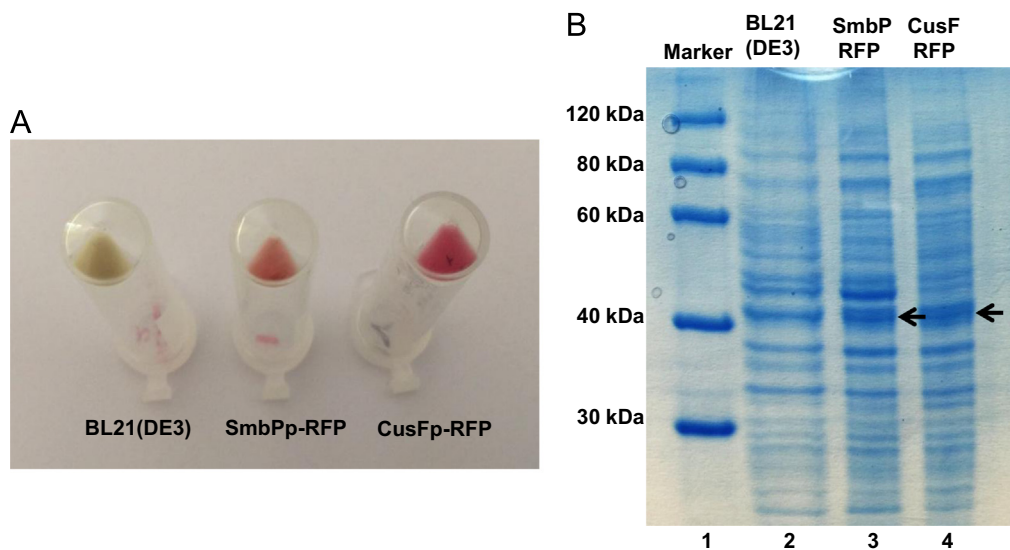


Fig. 3. Expression of red fluorescent protein in the periplasm of *E. coli* tagged with SmbPp and CusFp. (A) Picture of *E. coli* BL21 (DE3) cell pellets after expression of the different RFP constructs. The protein tag used for each construct is indicated below. A control, uninduced *E. coli* cells on the left, is included to show regular cell color. (B) 10% SDS-PAGE analysis of RFP periplasmic expression. Lane 1: protein marker; lane 2: *E. coli* BL21(DE3) periplasmic background control; lane 3: SmbPp-RFP periplasmic fraction; lane 4: CusFp-RFP periplasmic fraction. Arrows indicate the band for SmbPp-RFP and CusFp-RFP. Calculated molecular weights for SmbPp-RFP: 37.2 kDa, CusFp-RFP: 38.1 kDa.

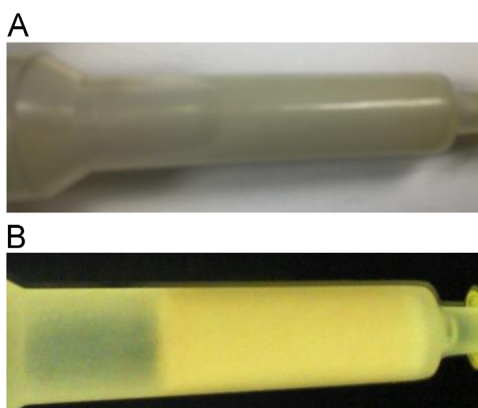


Fig. 4. Binding of CusF-GFP to Ag(I) chromatography media. (A) Image of the functionalized Bio-Gel P2 polyacrylamide media with glutaraldehyde and thiourea for immobilization of Ag(I) ions. (B) the same Ag(I) media after incubation with the lysate from *E. coli* expressing the CusF-GFP construct.

2 (SHY2) [3], and the S65T mutant of Green Fluorescent Protein (GFP) [4]. Red Fluorescent Protein [5], was expressed in the periplasm with the CusFp construct. A second RFP construct was tested for expression in the periplasm but tagged with SmbPp as previously described [6].

2.2. Protein expression

Recombinant proteins produced in *E. coli* are typically expressed at 37 °C for 3–4 h and induced with 1.0 mM isopropyl β -D-1-thiogalactopyranoside (IPTG) to deactivate the LacI repressor; these

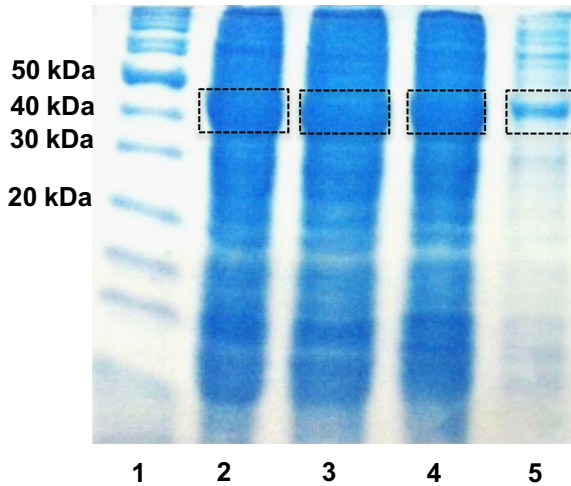


Fig. 5. 14% SDS-PAGE analysis of the purification of CusF-GFP with Ag(1) affinity chromatography. Lane 1: protein marker; lane 2: lysate; lane 3: flow-through; lane 4: first elution fraction with 160 mM methionine; lane 5: second elution fraction. Calculated molecular weight for CusF-GFP: 40.8 kDa.

conditions sometimes form insoluble inclusion bodies depending on the target protein [7]. The proteins LovR and SHY2, tagged with CusF and GST at the N-terminal, were expressed at the conditions mentioned above and alternatively at 25 °C, 16 h with 0.1 mM IPTG to avoid aggregation. DNA constructs were transformed into *E. coli* BL21(DE3). 2 mL of Luria–Bertani broth (with 30 µg/mL kanamycin) was inoculated with a single colony and incubated at 37 °C and 220 rpm until an OD_{600} of 0.5 was attained. IPTG was added to a final concentration of 0.1 mM to induce expression; cells were incubated overnight at room temperature (25 °C) and 220 rpm. If expression was induced with 1.0 mM IPTG, the cells were incubated at 37 °C for 4 h and 220 rpm. After protein expression cells were harvested by centrifugation and resuspended in 100 µL 1 × SDS-PAGE sample buffer, incubated in boiling water for 10 min, and then centrifuged for 10 min at 13,000 rpm; the supernatant was used for soluble protein content analysis with SDS-PAGE. The residual pellet was resuspended in 100 µL 8 M urea buffer and boiled for 10 min to solubilize the inclusion bodies, then centrifuged at the same speed; the supernatant was used for the analysis of insoluble protein content with SDS-PAGE. For CusF-GFP large-scale expression, cells were grown in baffled flasks until OD_{600} reached 0.5, expression was induced by adding IPTG up to 0.1 mM, cells were incubated overnight at room temperature (25 °C) and 220 rpm. For large-scale expression of CusFp-RFP and SmbPp-RFP expression, cells were grown in the same manner.

2.3. *E. coli* periplasm lysis

Expression of recombinant proteins in the cellular periplasm has certain advantages since it contains less proteases [8], and provides the proper conditions for the formation of disulfide bonds [9]. If the protein of interest is tagged with the CusFp or SmbPp constructs that include the signal sequences, then the target protein is exported to the periplasm. An osmotic shock procedure was performed to obtain the periplasmic fraction. Cells from a 1 L culture were resuspended in 20 mM Tris–HCl, 30% sucrose, 2.5 mM EDTA, pH 8.0 (5 mL/g of cell pellet) and incubated on ice for 1 h. Cells were centrifuged at 10,000 rpm for 15 min at 4 °C, the supernatant was recovered and saved, and the sediment was resuspended in 20 mL ice-cold 5 mM $MgSO_4$ buffer for 1 h. Cells were centrifuged at 10,000 rpm for 15 min at 4 °C, and the supernatant was also considered as periplasmic fraction, the two supernatants were pooled, and analyzed by SDS-PAGE.

2.4. Purification of CusF-GFP with immobilized Ag(I) affinity chromatography

CusF is a metal-binding protein with high binding affinities to copper and silver ions, especially Cu (I) and Ag(I) [10]; although it has been demonstrated that binds Cu(II) as well [11]. This last feature has allowed the purification of CusF-tagged proteins with immobilized Cu(II) affinity chromatography [1]. Nevertheless, the first metal ion tested to purify CusF-tagged proteins was Ag(I); this metal ion was immobilized on a functionalized acrylamide resin with glutaraldehyde and thiourea. Polyacrylamide beads (Bio-Gel P2, from BIO-RAD) were functionalized with silver ions as previously described [12]. First, a 25% glutaraldehyde solution was polymerized for 24 h at 50 °C. After hydration of the Bio-Gel P2 beads with a 50 mM sodium phosphate buffer (pH 8.0) for 24 h, the polymerized glutaraldehyde solution was added to the wet resin (gel) in a ratio of 20 mL/1 g of dry gel. They were incubated at 37 °C/120 rpm for 24 h, afterwards the resin was filtered and rinsed with deionized water. For every 1 g of initial dry gel, 20 mL of 1.0 M thiourea solution was added to the glutaraldehyde-activated gel resin and the slurry was incubated at 37 °C/120 rpm. After 24 h of reaction, the thiourea solution was removed by decantation and the gel was filtered and washed with deionized water again. The final step involved the addition of silver nitrate in a ratio of 0.9 mL of 1.0 N AgNO₃ for 1 g of dry gel and the solution was incubated at 37 °C. After 24 h the resin was washed with deionized water to remove all the unbounded silver. Cells expressing CusF-GFP were harvested by centrifugation at 4 °C and then resuspended in ice-cold lysis buffer (50 mM sodium phosphate, 300 mM NaCl, pH 8.0). Cells were lysed using a bead-beater and 0.1 mm glass beads (both from BioSpec Products); the lysate was clarified by centrifugation at 4 °C, 15,000 rpm for 20 min. 1 mL of the Ag(I) resin was incubated with the 5 mL of CusF-GFP lysate 30 min at room temperature. The resin was washed thoroughly with lysis buffer. Since two methionines and one histidine are used by CusF to bind Ag(I) [13], elution steps were performed using a buffer with a high methionine concentration (50 mM sodium phosphate, 300 mM NaCl, 160 mM methionine, pH 8.0). Elution fractions were analyzed by SDS-PAGE.

Acknowledgments

We thank Mexico's Consejo Nacional de Ciencia y Tecnologia (CONACYT) for financial support provided to JECB and TVC during their graduate studies. This work was funded by CONACYT Grant CB 2012-179774-B awarded to XZ.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at <http://dx.doi.org/10.1016/j.dib.2016.02.074>.

References

- [1] J.E. Cantu-Bustos, T. Vargas-Cortez, J.R. Morones-Ramirez, I. Balderas-Renteria, D.W. Galbraith, M.M. McEvoy, et al., Expression and purification of recombinant proteins in *Escherichia coli* tagged with the metal-binding protein CusF, *Protein Expr. Purif.* 121 (2016) 61–65. <http://dx.doi.org/10.1016/j.pep.2016.01.007>.
- [2] R. Foreman, A. Fiebig, S. Crosson, The *lovK-lovR* two-component system is a regulator of the general stress pathway in *Caulobacter crescentus*, *J. Bacteriol.* 194 (2012) 3038–3049. <http://dx.doi.org/10.1128/JB.00182-12>.
- [3] B.C. Kim, M.S. Soh, S.H. Hong, M. Furuya, H.G. Nam, Photomorphogenic development of the *Arabidopsis* shy2-1D mutation and its interaction with phytochromes in darkness, *Plant J.* 15 (1998) 61–68. <http://dx.doi.org/10.1046/j.1365-313X.1998.00179.x>.
- [4] R. Heim, R.Y. Tsien, Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy transfer, *Curr. Biol.* 6 (1996) 178–182. [http://dx.doi.org/10.1016/S0960-9822\(02\)00450-5](http://dx.doi.org/10.1016/S0960-9822(02)00450-5).
- [5] R.E. Campbell, O. Tour, A.E. Palmer, P.A. Steinbach, G.S. Baird, D.A. Zacharias, et al., A monomeric red fluorescent protein, *Proc. Natl. Acad. Sci. U.S.A.* 99 (2002) 7877–7882. <http://dx.doi.org/10.1073/pnas.082243699>.

- [6] T. Vargas-Cortez, J.R. Morones-Ramirez, I. Balderas-Renteria, X. Zarate, Expression and purification of recombinant proteins in *Escherichia coli* tagged with a small metal-binding protein from *Nitrosomonas europaea*, *Protein Expr. Purif.* 118 (2016) 49–54. <http://dx.doi.org/10.1016/j.pep.2015.10.009>.
- [7] G.L. Rosano, E.A. Ceccarelli, Recombinant protein expression in *Escherichia coli*: advances and challenges., *Front. Microbiol.* 5 (2014) 172. <http://dx.doi.org/10.3389/fmicb.2014.00172>.
- [8] B.A. Dow, S.A. Tatulian, V.L. Davidson, Use of the amicyanin signal sequence for efficient periplasmic expression in *E. coli* of a human antibody light chain variable domain, *Protein Expr. Purif.* 108 (2015) 9–12. <http://dx.doi.org/10.1016/j.pep.2014.12.017>.
- [9] F. Rastgar Jazii, A.A. Karkhane, B. Yakhchali, S.S. Fatemi, A. Deezagi, A simplified purification procedure for recombinant human granulocyte macrophage-colony stimulating factor from periplasmic space of *Escherichia coli*, *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* 856 (2007) 214–221. <http://dx.doi.org/10.1016/j.jchromb.2007.06.013>.
- [10] J.T. Kittleson, I.R. Loftin, A.C. Hausrath, K.P. Engelhardt, C. Rensing, M.M. McEvoy, Periplasmic metal-resistance protein CusF exhibits high affinity and specificity for both CuI and AgI, *Biochemistry* 45 (2006) 11096–11102. <http://dx.doi.org/10.1021/bi0612622>.
- [11] A.V. Astashkin, A.M. Raitsimring, F.A. Walker, C. Rensing, M.M. McEvoy, Characterization of the copper(II) binding site in the pink copper binding protein CusF by electron paramagnetic resonance spectroscopy, *J. Biol. Inorg. Chem.* 10 (2005) 221–230. <http://dx.doi.org/10.1007/s00775-005-0631-y>.
- [12] A.A. Garcia, D.H. Kim, D.R. Miles, Immobilization of silver and platinum ions for metal affinity chromatography, *React. Polym.* 23 (1994) 249–259. [http://dx.doi.org/10.1016/0923-1137\(94\)90027-2](http://dx.doi.org/10.1016/0923-1137(94)90027-2).
- [13] I.R. Loftin, S. Franke, N.J. Blackburn, M.M. McEvoy, Unusual Cu(I)/Ag(I) coordination of *Escherichia coli* CusF as revealed by atomic resolution crystallography and X-ray absorption spectroscopy, *Protein Sci.* 16 (2007) 2287–2293. <http://dx.doi.org/10.1110/ps.073021307>.