Journal of Medical Microbiology and Infectious Diseases

JoMMID
ISSN: 2345-5349
eISSN: 2345-5330

The Influence of IF2 and SUMO Fusion Tags on the Expression and Solubility of the SpCas9 Protein

Farzaneh Barkhordari[1], Mahshid Hosseini Shalmaei[1, 2], Setayesh Naderi[1, 2], Mozhgan Raigani[1], Yeganeh Kayalha[3], Elham Bayat[1], Negin Sadat Hashemi Aval[1], Ahmad Adeli[1*], Fatemeh Davami[1*]

[1]Biotechnology Research Center, Pasteur Institute of Iran, Tehran, Iran; [2]Department of Biotechnology, Faculty of Advanced Science and Technology, Tehran Medical Sciences, Islamic Azad University, Tehran, Iran; [3]Faculty of Life Sciences and Biotechnology, Shahid Beheshti University, Tehran, Iran.

ARTICLE INFO

Original Article

Keywords: SpCas9 expression, CRISPR/Cas9, Initiation factor 2, Solubility-enhancing tag, Small ubiquitin-like modifier, *Escherichia coli*

Received: 24 May. 2026
Received in revised form: 23 Jul. 2026
Accepted: 09 Aug. 2026
DOI: 10.61882/JoMMID.14.2.149

***Correspondence**
Email: f.davami@gmail.com; adeliahmad@yahoo.com

ABSTRACT

Introduction: The expression of the Cas9 protein in the prokaryotic host *Escherichia coli* is often associated with the formation of inclusion bodies. This study aimed to express and purify the Cas9 protein in large quantities and soluble form in *E. coli* BL21 (DE3) for use in the CRISPR/Cas9 system. Due to the large size of Cas9, two solubility-enhancing fusion partners with relatively low molecular weights, initiation factor 2 (IF2) and small ubiquitin-like modifier (SUMO), were utilized. **Methods:** The *IF2* and *SUMO* gene sequences, 480 and 300 bp in length, respectively, were amplified by PCR and cloned into the pET-21a-3xNLS-SpCas9 vector. The expression of both recombinant proteins, with approximate molecular weights of 176 and 166 kDa, was confirmed by SDS-PAGE and Western blotting. Protein purification was performed using Ni-NTA affinity chromatography. **Results:** The presence of IF2 as a fusion partner increased the percentage of Cas9 detected in the supernatant from 5.51% to 6.99%. Moreover, under optimal cultivation conditions, the recombinant IF2-Cas9 protein was produced in greater quantities than Cas9 alone (18.50 versus 13.65 mg/L). The solubility analysis revealed that SUMO-Cas9 exhibited higher solubility (9.25%) than Cas9 alone (5.51%), representing a 67.9 ± 8.2% increase in soluble expressio. **Conclusion:** While both fusion partners enhanced Cas9 expression in *E. coli*, SUMO demonstrated superior solubility-enhancing capability compared to IF2; however, the functional activity of the purified protein remains to be confirmed.

INTRODUCTION

Since 2005, the CRISPR-Cas system has been proposed to function as a natural defense mechanism that bacteria use to protect themselves against bacteriophages [1]. This system can be adapted to create double-strand breaks in the genomes of various organisms at target sites adjacent to the protospacer adjacent motif (PAM). As a result, it has emerged as a powerful tool for genome editing. Although this technology has been widely utilized in cell-based and laboratory studies, it also holds significant potential for *in vivo* therapeutic applications and is considered an efficient method for genome editing [2, 3].

Cas9 is an essential component of type II CRISPR-Cas systems. CRISPR-associated (Cas) proteins include endonucleases identified in diverse prokaryotes; Cas9 derived from *Streptococcus pyogenes* (SpCas9) is commonly used, has a molecular weight of 158 kDa, and functions as a DNA endonuclease with two lobes: an alpha-helical lobe and a nuclease lobe [4, 5].

Recently, the CRISPR-Cas9 system has been successfully applied in the treatment of genetic diseases such as sickle cell disease [6]. Additionally, CRISPR-Cas-based methods have been used in highly sensitive nucleic acid detection techniques, such as SHERLOCK, and in HIV treatment [7, 8]. Typically, plasmids encoding SpCas9 and guide RNA (gRNA) are transfected into cells for expression. However, plasmid-based CRISPR-Cas delivery presents several limitations. One major drawback is unintended integration of the plasmid, or parts of it, into the host genome. Moreover, DNA transfection can be stressful for cells and may trigger the cyclic GMP-AMP synthase (cGAS)-STING pathway, causing an innate immune response. Additionally, SpCas9 expression can persist for several days after transfection, and prolonged SpCas9 activity, which

[DOI: 10.61882/JoMMID.14.2.149]

generates double-strand breaks (DSBs), increases off-target effects in genome editing [9, 10].

The Cas9-gRNA complex, known as ribonucleoprotein (RNP), has been shown to reduce off-target effects while increasing the efficiency of targeted genome editing [10, 11]. In this approach, purified Cas9 protein is complexed with synthetic gRNA *in vitro* before being delivered to the target cell. Because this RNP complex is preassembled in the laboratory, it does not require transcription or translation within the cell and can become immediately active upon delivery, thereby enhancing genome-editing efficiency [9, 12-14].

The delivery method of the CRISPR-Cas9 tool significantly affects both safety and therapeutic effectiveness. The RNP approach offers advantages such as faster activity, reduced immune responses, and reduced off-target effects [9, 11, 12, 15, 16]. Consequently, efficient production of soluble Cas9 protein is essential.

The soluble expression of recombinant proteins in *E. coli* depends on the intrinsic properties of the target protein [17]. Comparative studies indicate that as protein molecular mass increases, the likelihood of soluble expression in *E. coli* decreases. A large-scale study analyzing 95 eukaryotic recombinant proteins found that smaller proteins (40.4 kDa on average) require fusion tags to remain soluble [18]. Moreover, the N-terminal and C-terminal residues of a protein can influence expression and solubility. Klock *et al*. [19] demonstrated that the removal of just four amino acids from both the N-terminal and C-terminal regions can convert a soluble protein into an insoluble form.

The expression of the Cas9 protein in *E. coli* often leads to the formation of inclusion bodies. To overcome this challenge and enhance solubility, researchers have explored the use of solubility-enhancing fusion tags. Several fusion proteins, such as maltose-binding protein (MBP) and NusA, have been introduced to improve recombinant protein expression. These fusion partners are relatively large (40–55 kDa) and enhance both the solubility and stability of aggregation-prone proteins [17, 20]. Each solubility-enhancing tag has distinct biochemical properties, and the mechanism of its solubility enhancement is not fully understood [17, 20]. However, one hypothesis suggests that fusion proteins exhibit chaperone-like activity, preventing aggregation by interacting with partially folded proteins. Additionally, studies indicate that N-terminal fusions are generally more effective than C-terminal fusions in improving soluble protein expression [21]. Malhotra emphasized the importance of fusion tag size in selecting the appropriate solubility tag; the metabolic burden on the host cell differs depending on whether large or small tags are used [22]. Despite these insights, the selection of an appropriate fusion tag remains largely a trial-and-error process, as the effectiveness of solubility tags varies among different target proteins [22].

Solubility partners with smaller molecular masses, such as the N-terminal domain I of initiation factor 2 (IF2; ~18–20 kDa), have been explored as alternatives to larger fusion tags [17, 20]. Full-length IF2 is an initiation factor that binds to the small ribosomal subunit (30S) in *E. coli*; however, the N-terminal domain I (18 kDa, residues 1–157) is unable to bind independently to ribosomes and lacks catalytic activity [23].

Another established fusion tag is small ubiquitin-like modifier (SUMO; ~11 kDa), which has been widely used to enhance solubility and expression levels [24]. This system was initially developed based on the observation that ubiquitin fusion improves the solubility of recombinant proteins [25, 26]. Numerous studies have confirmed that SUMO fusion increases the solubility of hard-to-express proteins, such as MMP13 and eGFP, in bacterial hosts [27]. SUMO fusion has also been successfully applied to enhance the soluble expression of SARS coronavirus proteins and several membrane proteins [28, 29]. A major advantage of SUMO is its low molecular mass (~11 kDa). However, SUMO protease cannot process substrates with an N-terminal proline immediately downstream of the cleavage site, because this residue can block access to the active site [24].

Therefore, this study aimed to investigate whether N-terminal IF2 and SUMO fusions can reduce inclusion body formation and enhance the soluble expression of the high-molecular-weight SpCas9 protein in *E. coli*.

MATERIAL AND METHODS

Materials. Phusion High-Fidelity DNA polymerase and the SalI, BglII, XhoI, HindIII, and NheI restriction enzymes were obtained from Thermo Fisher Scientific, USA. A Ni-NTA affinity chromatography column was purchased from Biobasic Corporation (Biobasic, Canada). Additional required chemicals were sourced from Sigma-Aldrich, USA. The bacterial strains used as the host organisms for the expression of the cloned genes were *E. coli* BL21(DE3), Rosetta(DE3), and BL21(DE3)pLysS.

Cloning and expression of IF2 fusion in the pET-21a-3xNLS-SpCas9 vector. Gene-specific primers carrying a SalI restriction site at their *5′* termini were designed to facilitate cloning of the *IF2* gene (forward primer: CGAGTCGACATGACAGATGTAACG, reverse primer: TTGGTCGACCACTTTGTCTTTTTC). PCR was performed using the following cycling conditions: an initial denaturation at 95°C for 4 min, followed by 30 cycles of 95°C for 20 s, 65°C for 20 s, and 72°C for 20 s, and a final extension at 72°C for 10 min. PCR amplification of the *IF2* gene generated a 480-bp band. The amplified fragment was subsequently inserted upstream of the *cas9* gene in the pET-21a-3xNLS-SpCas9 vector. The recombinant IF2-pET-21a-3xNLS-SpCas9 plasmid was subsequently transformed into the *E. coli* expression strain BL21(DE3). After overnight incubation at 37°C, single colonies were then cultured in 5 mL of ampicillin-containing medium in tubes to examine the expression of the IF2-Cas9 protein at 37°C and 30°C with shaking at 200 rpm. Once the optical density (OD_{600}) reached 0.5, the culture was induced with 0.2 mM IPTG for 4 and 8 h. The culture conditions, including IPTG

[DOI: 10.61882/JoMMID.14.2.149]

concentration, post-induction incubation time, glucose supplementation, culture volume, and temperature, were also assessed for protein expression. Pre-induction and 4- and 8-h post-induction samples were analyzed on an 8% polyacrylamide gel to compare IF2-Cas9 protein expression yield. The expression of the recombinant IF2-Cas9 fusion protein was then verified by Western blot analysis using a horseradish peroxidase (HRP)-conjugated anti-His tag antibody.

Purification of recombinant IF2-Cas9 protein. The recombinant IF2-Cas9 protein was purified using affinity chromatography under native conditions with a nickel-Sepharose column, according to the manufacturer's protocol (QIAexpress® Ni-NTA Fast Start Handbook, Qiagen). The purity of the isolated target protein was assessed using SDS-PAGE followed by Coomassie Brilliant Blue staining.

Cloning and expression of SUMO fusion in the pET-21a-3xNLS-SpCas9 vector. To clone the *SUMO* gene, suitable primers (forward primer: AAAGTCGACATGGGGTCCCTGCAG, reverse primer: TTGTCGACACCTCCAATCTGTTCG) were designed with the inclusion of the restriction enzyme site for *Sal*I at the start of the sequence. A PCR reaction was carried out to amplify the *SUMO* gene using the following cycling conditions: an initial denaturation at 95°C for 4 min, followed by 30 cycles of 95°C for 15 s, 65°C for 15 s, and 72°C for 15 s, and a final extension at 72°C for 10 min. This yielded a product of approximately 300 bp in length. The resulting PCR product was cloned upstream of the *cas9* gene within the pET-21a-3xNLS-SpCas9 vector.

After confirmation of cloning, the verified plasmid, SUMO-pET-21a-3xNLS-SpCas9, was transformed into the BL21 (DE3) expression strain. The transformed cells were plated and incubated overnight at 30°C and 37°C. The resulting colonies were cultured in 5 mL of ampicillin-containing medium to assess the expression of the SUMO-Cas9 protein. Expression was evaluated at 30°C and 37°C with shaking at 200 rpm. Once the OD_{600} reached 0.5, induction was performed using 0.5 mM IPTG for 4 h. Samples were collected before and after induction to evaluate protein expression levels, which were subsequently analyzed using an 8% polyacrylamide gel.

Following the test of recombinant SUMO-Cas9 protein expression, its expression was confirmed by performing Western blotting using an antibody specifically against the His-tag fused with HRP.

Purification of recombinant SUMO-Cas9 protein. The recombinant SUMO-Cas9 protein was purified using affinity chromatography, following the standard Qiagen protocol under native conditions. The purity of the isolated protein was assessed through SDS-PAGE, followed by Coomassie Brilliant Blue staining.

Comparison of Cas9, SUMO-Cas9, and IF2-Cas9 expression in different *E. coli* hosts. After confirming expression, the SUMO-pET-21a-3xNLS-SpCas9 and IF2-pET-21a-3xNLS-SpCas9 vectors were transformed into different *E. coli* expression strains, including BL21(DE3), Rosetta(DE3) and BL21(DE3) pLysS. Different strains were compared to determine whether expression trends were consistent across hosts. The transformed cells were plated and incubated overnight at 37°C with shaking at 200 rpm. At an OD_{600} of 0.5, induction was carried out using 0.2 mM IPTG for 4 h. Pre-and post-induction samples were analyzed on an 8% polyacrylamide gel to assess the expression levels.

Comparison of expression and solubility of Cas9, SUMO-Cas9, and IF2-Cas9 proteins. To determine whether IF2 and SUMO fusion proteins improve the solubility of Cas9, each of these proteins was individually overexpressed in 10 mL of LB broth containing ampicillin under optimized expression conditions, including an OD_{600} of 0.5 and induction with 0.2 mM IPTG for 4 h in BL21(DE3).

The solubility of Cas9, IF2-Cas9, and SUMO-Cas9 proteins was then assessed following a standardized procedure developed in this study. After induction, bacterial pellets were resuspended in 0.5 mL phosphate-buffered saline (PBS), pH 7.2, and incubated on ice for 20 minutes. The resuspended cells were lysed by ultrasonication for 10 cycles of 20s pulses and 30s rest intervals between cycles. An aliquot of the sonicated sample was reserved as the total fraction, and the supernatant (Sup) was separated from the cell pellet (Pellet) by centrifugation at 17,000 × *g* for 30 minutes. The protein concentrations of the supernatant (Sup), pellet (Pellet), and total (Total) samples were measured using a NanoDrop spectrophotometer at a wavelength of 280 nm. Then, 100 µg of each sample was loaded onto an SDS-PAGE gel and analyzed using ImageJ software.

RESULTS

Analysis of cloning and expression of IF2 fusion. PCR amplification of the *IF2* gene yielded the expected 480 bp product. The presence of the *IF2* gene was confirmed by restriction enzymes *Bgl*II and *Xho*I in the pET-21a-3xNLS-SpCas9 vector. The observation of fragments of 490 and 525 bp confirmed the cloning. To assess the expression of the IF2-Cas9 protein, samples were collected before and 4 and 8 h after induction with 0.2 mM IPTG at 37°C and 30°C and were analyzed on an 8% polyacrylamide gel (Figure 1). A distinct band corresponding to 176 kDa was observed, confirming the expression of the IF2-Cas9 protein. In addition, various parameters, including IPTG concentration, post-induction incubation time, glucose supplementation, culture volume, and temperature, were optimized. The final optimized expression conditions were an OD_{600} of 0.5 and induction with 0.2 mM IPTG for 4 h at 37°C or 30°C. We selected a temperature of 37°C for identification and purification.

[DOI: 10.61882/JoMMID.14.2.149]

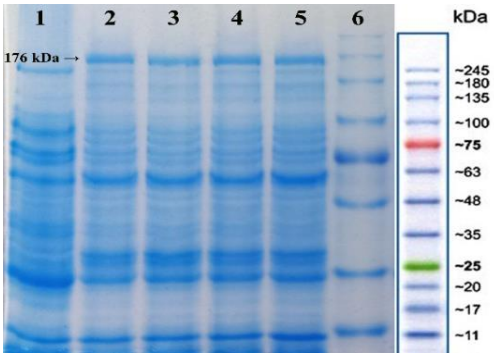

Fig.1. SDS-PAGE gel analysis of the expression of IF2-Cas9 at 30°C and 37°C. The recombinant IF2-pET-21a-3xNLS-SpCas9 vector was transformed into *E. coli* BL21(DE3). After overnight incubation at 37°C, single colonies were cultured in 5 mL of ampicillin-containing medium in tubes. Once the OD_{600} reached 0.5, the culture was induced with 0.2 mM IPTG for 4 and 8 h. Lane 1: before induction; Lane 2: after induction for 4 h at 30°C; Lane 3: after induction for 8 h at 30°C; Lane 4: after induction for 4 h at 37°C; Lane 5: after induction for 8 h at 37°C; Lane 6: protein marker. A distinct band corresponding to 176 kDa was observed, confirming the expression of the IF2-Cas9 protein.

Recombinant IF2-Cas9 protein expression was further confirmed by Western blot analysis using an HRP-conjugated anti-His tag antibody, which detected a distinct band at the expected molecular weight of 176 kDa (Figure 2).

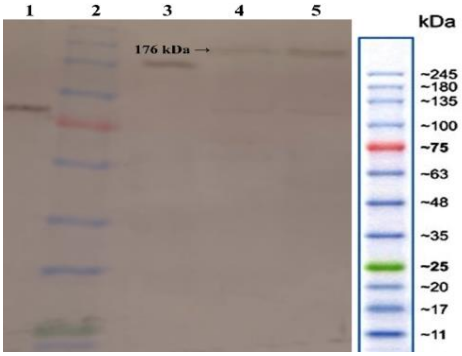

Fig. 2. Western blotting of IF2-Cas9 protein expression. The expression of the recombinant IF2-Cas9 fusion protein was verified by Western blot analysis using an HRP-conjugated anti-His tag antibody. Lane 1: positive control sample containing a His-tagged protein; Lane 2: protein marker; Lane 3: Cas9 protein (~158 kDa); Lanes 4 and 5: IF2-Cas9 protein samples (~176 kDa).

Purification of recombinant IF2-Cas9 protein. The recombinant IF2-Cas9 protein was purified using affinity chromatography with a Ni-NTA column, following the Qiagen protocol under native conditions. As shown in Figure 3, after SDS-PAGE, a distinct band corresponding to the IF2-Cas9 protein with an apparent molecular mass of 176 kDa was observed, confirming successful purification.

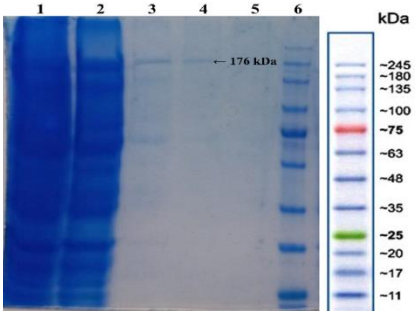

Fig. 3. SDS-PAGE gel showing purification of His-tagged recombinant IF2-Cas9 protein via native affinity chromatography. To obtain the purified target protein, native affinity chromatography was performed, and the fractions were assessed by SDS-PAGE, followed by Coomassie Brilliant Blue staining. Lane 1: initial sample; Lane 2: flow-through sample; Lanes 3 – 5: protein samples eluted from the column after washing with elution buffer; Lane 6: protein marker. The purified IF2-Cas9 protein appears as a ~176 kDa band in the eluted fractions.

Analysis of cloning and expression of SUMO fusion. A PCR reaction was carried out to amplify the *SUMO* gene. This yielded a product of approximately 300 bp. The presence of the *SUMO* gene was confirmed by HindIII and

[DOI: 10.61882/JoMMID.14.2.149]

NheI restriction enzymes in the pET-21a-3xNLS-SpCas9 vector by observation of fragments of 2345 and 7638 bp and fragments of 1485 and 8498 bp, respectively. After confirmation of cloning, the expression of the SUMO-Cas9 protein was assessed in the resulting colonies. Following induction with 0.5 mM IPTG, samples were collected before and after induction to evaluate protein expression levels, which were subsequently analyzed using an 8% polyacrylamide gel (Figure 4). A distinct band at 166 kDa was detected, confirming the expression of the SUMO-Cas9 protein. The final optimized expression conditions were an OD_{600} of 0.5 and induction with 0.5 mM IPTG for 4 h at 37°C or 30°C. We selected **a** temperature of 37°C for identification and purification.

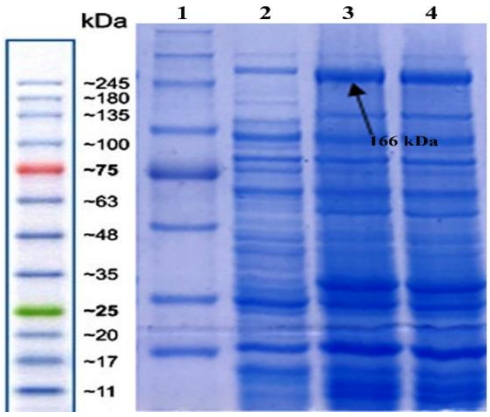

Fig. 4. SDS-PAGE gel analysis of SUMO-Cas9 protein expression at 30°C and 37°C with 0.5 mM IPTG. Cells of *E. coli* BL21(DE3) transformed with the SUMO-pET-21a-3xNLS-SpCas9 vector were plated and incubated overnight at 30°C and 37°C. The resulting colonies were cultured in 5 mL of ampicillin-containing medium. Expression of the SUMO-Cas9 protein was carried out at 30°C and 37°C with shaking at 200 rpm. Once the OD_{600} reached 0.5, induction was performed using 0.5 mM IPTG for 4 h. Lane 1: protein marker; Lane 2: before induction with 0.5 mM IPTG; Lane 3: after induction with 0.5 mM IPTG for 4 h at 37°C; Lane 4: after induction with 0.5 mM IPTG for 4 h at 30°C. A distinct band at 166 kDa was detected, confirming the expression of the SUMO-Cas9 protein under the optimized expression conditions (OD_{600} of 0.5 and induction with 0.5 mM IPTG for 4 h at 37°C or 30°C).

Following the expression of the recombinant SUMO-Cas9 protein, further confirmation was successfully performed by Western blot analysis using an HRP-conjugated anti-His-tag antibody (Figure 5).

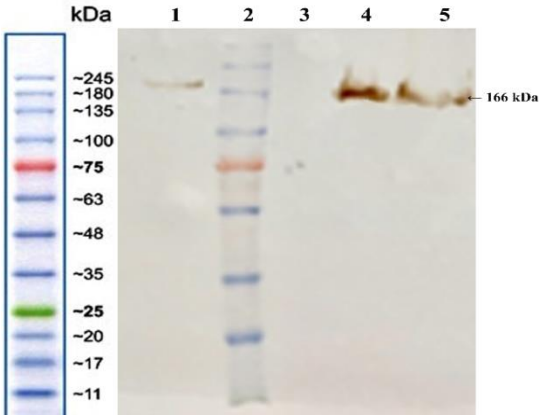

Fig. 5. Western blot confirmation of recombinant SUMO-Cas9 protein expression using an HRP-conjugated anti-His-tag antibody. Lane 1: positive control Cas9 (~158 kDa); Lane 2: protein marker; Lane 3: sample before induction; Lanes 4 and 5: SUMO-Cas9 protein samples (~166 kDa).

Purification of recombinant SUMO-Cas9 protein. The recombinant SUMO-Cas9 protein was purified using Ni-NTA affinity chromatography under native conditions, and the purity of the isolated protein was assessed by SDS-PAGE, followed by Coomassie Brilliant Blue staining. A distinct band corresponding to the SUMO-Cas9 protein was observed at 166 kDa, confirming successful purification (Figure 6).

[DOI: 10.61882/JoMMID.14.2.149]

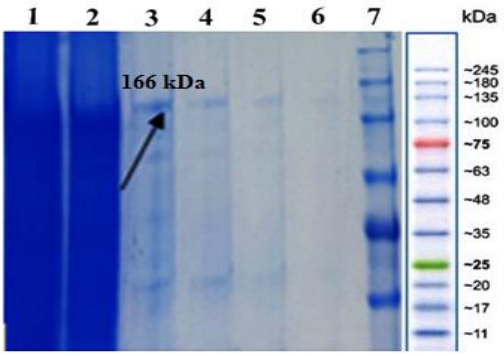

Fig. 6. SDS-PAGE gel showing purification of His-tagged recombinant SUMO-Cas9 protein by native affinity chromatography. Lane 1: initial sample; Lane 2: flow-through sample; Lanes 3 – 6: samples eluted from the column after washing with elution buffer; Lane 7: protein marker. The purified SUMO-Cas9 protein appears as a ~166 kDa band in the eluted fractions.

Comparison of Cas9, SUMO-Cas9, and IF2-Cas9 expression in different *E. coli* hosts. To assess the effect of different host strains on expression levels, the SUMO-pET-21a-3xNLS-SpCas9 and IF2-pET-21a-3xNLS-SpCas9 vectors were transformed into different *E. coli* expression strains, including BL21(DE3), Rosetta(DE3), and BL21(DE3) pLysS. Samples were collected before and after induction with 0.2 mM IPTG to evaluate and compare their expression levels (Figure 7). The expression level of IF2-Cas9 was 18.50 mg/L, compared to 13.65 mg/L for Cas9 alone in Rosetta(DE3) at 37°C. The highest SUMO-Cas9 expression was observed in Rosetta(DE3) at 37°C, yielding 20 mg/L, compared to 13.65 mg/L for Cas9 alone. Furthermore, the results showed that the host strain did not alter the relative trends in expression and solubility.

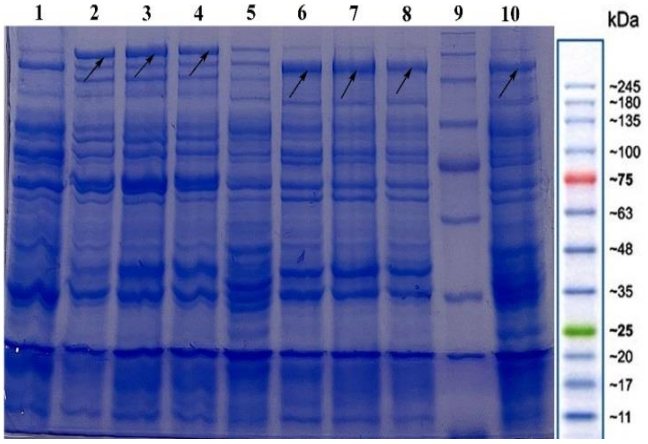

Fig. 7. SDS-PAGE gel analysis of the effect of different strains on SUMO-Cas9 and IF2-Cas9 expression. Lane 1: uninduced control; Lane 2: expression of IF2-Cas9 in BL21(DE3) (~176 kDa); Lane 3: expression of IF2-Cas9 in Rosetta(DE3); Lane 4: expression of IF2-Cas9 in BL21(DE3) pLysS; Lane 5: uninduced control; Lane 6: expression of SUMO-Cas9 in BL21(DE3) (~166 kDa); Lane 7: expression of SUMO-Cas9 in Rosetta(DE3); Lane 8: expression of SUMO-Cas9 in BL21(DE3) pLysS; Lane 9: protein marker; Lane 10: expression of Cas9 in BL21(DE3) (~158 kDa). As shown, the host strain did not alter the relative trends in expression and solubility.

Comparison of expression and solubility of Cas9, SUMO-Cas9, and IF2-Cas9 proteins. To determine whether IF2 and SUMO fusion proteins improve the solubility of Cas9, each of these proteins was individually overexpressed in 10 mL of LB broth containing ampicillin, under optimized expression conditions.

As shown in Figure 8, the solubility of the IF2-Cas9 protein in the supernatant was found to be 6.99%, compared to 5.51% for Cas9 alone, representing a 1.48-percentage-point increase in solubility (n = 3) following densitometric analysis using ImageJ software. Similarly, the solubility of the SUMO-Cas9 protein in the supernatant was 9.25%, compared to 5.51% for Cas9 alone, indicating a 67.9 ± 8.2% improvement in solubility (n = 3) (Figure 9).

[DOI: 10.61882/JoMMID.14.2.149]

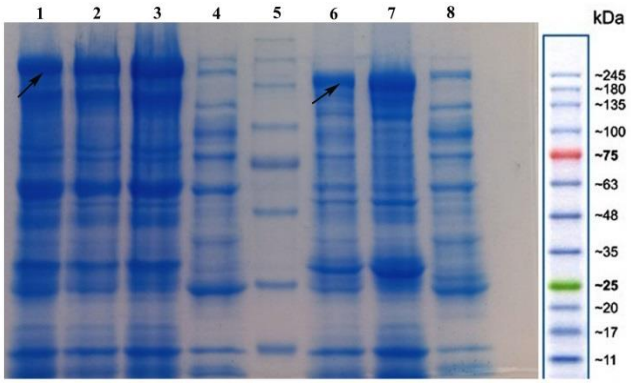

Fig. 8. SDS-PAGE analysis of the effect of IF2 fusion on Cas9 protein solubility. Lanes 1 and 2: total samples after IF2-Cas9 protein expression; Lane 3: cell pellet after IF2-Cas9 protein expression; Lane 4: supernatant after IF2-Cas9 protein expression; Lane 5: protein marker; Lane 6: total sample after Cas9 protein expression; Lane 7: cell pellet after Cas9 protein expression; Lane 8: supernatant after Cas9 protein expression. The IF2-Cas9 and Cas9 bands correspond to ~176 and ~158 kDa, respectively. The solubility of the IF2-Cas9 protein in the supernatant was found to be 6.99%, compared to 5.51% for Cas9 alone, representing a 1.48-percentage-point increase in solubility (n = 3) following densitometric analysis using ImageJ software.

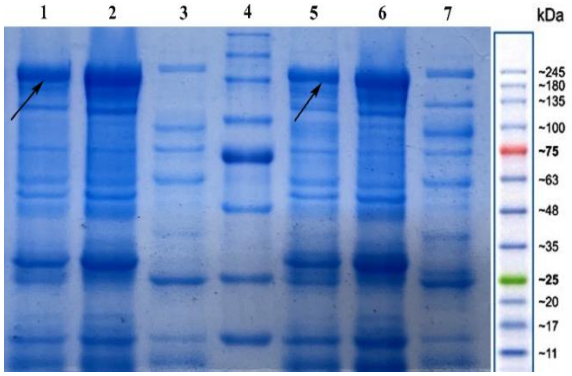

Fig. 9. SDS-PAGE analysis of the effect of SUMO fusion on Cas9 protein solubility. Lane 1: total samples after SUMO-Cas9 protein expression (~166 kDa); Lane 2: cell pellet after SUMO-Cas9 protein expression; Lane 3: supernatant after SUMO-Cas9 protein expression; Lane 4: protein marker; Lane 5: total sample after Cas9 protein expression (~158 kDa); Lane 6: cell pellet after Cas9 protein expression; Lane 7: supernatant after Cas9 protein expression. The solubility of the SUMO-Cas9 protein in the supernatant was 9.25%, compared to 5.51% for Cas9 alone, indicating a 67.9 ± 8.2% improvement in solubility (n = 3).

DISCUSSION

Approximately half of bacterial species possess an adaptive immune mechanism called the CRISPR-Cas system, which defends against invading phages. This system can adapt, rapidly generating immunity against new phage infections [1]. Its ability to induce double-strand breaks at specific locations, particularly near PAM sequences, has made the CRISPR-Cas system an invaluable tool for genome editing [3]. Studies have demonstrated that the CRISPR-Cas9 system can be delivered to target cells in three different forms: plasmid DNA (pDNA), messenger RNA (mRNA), or RNP, each with its own advantages and disadvantages. The delivery method of CRISPR tools is a crucial factor that determines both therapeutic safety and efficacy. The RNP method has been characterized by rapid activity, decreased immune responses, and reduced off-target effects. Research has shown that the RNP complex of Cas9 and gRNA minimizes off-target effects and enhances the efficiency of targeted insertions [30]. The complex consists of purified Cas9 protein combined with synthetic gRNA, which is directly delivered into the cells, bypassing the need for transcription and translation. Once inside the cell, the RNP complex is immediately active, thereby increasing the efficiency of gene insertion [30]. Because the complex is preassembled *in vitro*, it is immediately active upon delivery, bypassing the need for intracellular transcription and translation [30].

To overcome solubility challenges, researchers have fused solubility partners to recombinant proteins to enhance their solubility and stability, particularly for proteins prone to aggregation. Although commonly used solubility tags such as MBP, glutathione S-transferase (GST), and NusA have relatively large molecular weights (40, 30, and 55 kDa, respectively), studies have shown that smaller solubility partners, such as SUMO (11 kDa), can also yield soluble recombinant proteins comparable to those obtained with larger fusion partner [24].

[DOI: 10.61882/JoMMID.14.2.149]

Research suggests that smaller solubility partners, such as domain I of IF2 (18 kDa), can achieve solubility levels similar to their larger counterparts [17, 20]. In a study by Sørensen *et al*. [17], fusing domain I of IF2 with streptavidin in *E. coli* BL21(DE3) resulted in a three-fold increase in soluble active streptavidin expression compared with that achieved using MBP or NusA. Likewise, Yang *et al*. [20] demonstrated that pluripotency transcription factors fused with domain I of IF2 were successfully expressed and purified in an active, soluble form.

The Cas9 protein is relatively large (158 kDa) and consists of an alpha-helical lobe and a nuclease lobe. Upon expression in *E. coli*, Cas9 often forms inclusion bodies due to its insolubility and aggregation. The recovery of active Cas9 requires a time-consuming and challenging refolding process [31].

Despite advancements in non-bacterial recombinant expression systems (such as yeast, baculovirus, mammalian cells, and cell-free systems), *E. coli* remains the preferred host for recombinant protein expression due to its fast growth rate, cost-effectiveness, ease of genetic manipulation, and ability to produce high yields of recombinant proteins [32]. However, *E. coli* also has several limitations as a host system. Large proteins or those with multiple membrane domains often fail to express properly, leading to insoluble inclusion bodies. Additionally, *E. coli* lacks the machinery for eukaryotic post-translational modifications, such as glycosylation, which can be crucial for proper protein folding and activity of certain targets [33].

To improve soluble protein expression in *E. coli*, various strategies have been explored, including the use of fusion proteins (solubility tags), co-expression with molecular chaperones, low-temperature cultivation, and engineering of specialized expression vectors and host strains [33]. Among the various *E. coli* strains available for recombinant protein expression, BL21(DE3) is the most widely used [34]. This study utilized the pET-21a-3xNLS-SpCas9 vector, which expresses the target gene under the control of the T7 promoter, a highly efficient system for protein expression in *E. coli* [31].

In this study, the *IF2* gene was cloned upstream of the *cas9* gene in the pET-21a-3xNLS-SpCas9 vector. This led to the expression of the IF2-Cas9 protein (176 kDa) in the *E. coli* BL21(DE3) strain. The protein was subsequently purified using affinity chromatography. To determine the optimal expression conditions, bacterial cultures were grown at 30°C and 37°C, using IPTG concentrations ranging from 0.2 to 1 mM. The highest IF2-Cas9 expression level was observed in Rosetta(DE3) at 37°C, induced with 0.2 mM IPTG for 4 h. The expression level of IF2-Cas9 was 18.50 mg/L, compared to 13.65 mg/L for Cas9 alone, demonstrating that IF2 enhanced expression. However, SDS-PAGE analysis revealed that a substantial fraction of IF2-Cas9 was retained in the cell pellet as inclusion bodies, with only 6.99% present in the supernatant, compared to 5.51% for Cas.

The highest SUMO-Cas9 expression was observed in Rosetta(DE3) at 37°C, yielding 20 mg/L, compared to 13.65 mg/L for Cas9 alone. In terms of solubility, SUMO-Cas9 exhibited higher solubility (9.25%) compared to Cas9 (5.51%), demonstrating a marked improvement in soluble expression. Marblestone *et al*. [35] demonstrated that SUMO is superior to commonly used fusion tags in enhancing expression and solubility, with the additional advantage of generating recombinant proteins with native sequences. A similar study showed that the N-terminal fusion of SUMO with IFNα2-Tα1 is effective for its soluble expression and makes its purification process more convenient [36]. SUMO acts as a solubility enhancer by providing a hydrophilic folding scaffold that directly prevents the aggregation of the target protein [37]. In contrast, the *E. coli* translation initiation factor IF2 primarily functions as a covalently attached fusion partner, improving expression yield and folding, with its solubility effect often being a beneficial secondary consequence. While both can increase soluble yield, SUMO is more universally employed as a dedicated solubility tag and offers the key advantage of being cleanly removable by SUMO-specific proteases to yield a native protein product [38].

In conclusion, this study evaluated the expression and solubility of recombinant fusion proteins (IF2-Cas9 and SUMO-Cas9) in *E. coli*. While both fusion partners enhanced expression, SUMO-Cas9 exhibited greater solubility than IF2-Cas9. The Rosetta(DE3) strain was identified as the most efficient host for producing soluble Cas9 fusion proteins. These findings highlight the potential of SUMO as a superior solubility-enhancing tag for Cas9 expression in *E. coli*; however, the functional activity of the purified proteins has yet to be confirmed. Future work will focus on cleaving the SUMO tag, assembling the RNP complex, and validating the functional genome-editing efficiency of the purified protein in mammalian cells.

ACKNOWLEDGMENT

The authors wish to thank the Biotechnology Research Center of the Pasteur Institute of Iran.

CONFLICTS OF INTEREST

The authors declare that there are no conflicts of interest associated with this manuscript.

FUNDING

This study was supported by grant No. 2141 from the Pasteur Institute of Iran.

DATA AVAILABILITY

The summary data supporting the findings of this study are presented within the article. Raw data,

[DOI: 10.61882/JoMMID.14.2.149]

including individual replicate measurements, are available from the corresponding author upon reasonable request.

AI DISCLOSURE

The authors declare that no artificial intelligence (AI) tools were used in the writing or data analysis for this manuscript.

AUTHORS' CONTRIBUTIONS

FB: Methodology; Conceptualization. MHSh: Methodology. SN: Methodology. MR: Writing – review & editing. YK: Writing – original draft. EB: Investigation. NSHA: Formal analysis. AA: Supervision. FD: Conceptualization; Project administration. All authors read and approved the final manuscript.

ETHICS STATEMENT

This work does not contain any studies involving human or animal subjects.

REFERENCES

1. Mojica FJ, Díez-Villaseñor C, García-Martínez J, Soria E. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. J Mol Evol. 2005; 60 (2): 174-82.
2. Gagnon JA, Valen E, Thyme SB, Huang P, Akhmetova L, Pauli A, et al. Efficient mutagenesis by Cas9 protein-mediated oligonucleotide insertion and large-scale assessment of single-guide RNAs. PLoS One. 2014; 9 (5): e98186.
3. Doudna JA, Charpentier E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. Science. 2014; 346 (6213): 1258096.
4. Nishimasu H, Ran FA, Hsu PD, Konermann S, Arora S, Kocha JS, et al. Crystal structure of Cas9 in complex with guide RNA and target DNA. Cell. 2014; 156 (5): 935-49.
5. Wiedenheft B, Sternberg SH, Doudna JA. RNA-guided genetic silencing systems in bacteria and archaea. Nature. 2012; 482 (7385): 331-8.
6. Frangoul H, Altshuler D, Cappellini MD, Chen YS, Domm J, Eustace BK, et al. CRISPR-Cas9 gene editing for sickle cell disease and β-thalassemia. N Engl J Med. 2021; 384 (3): 252-60.
7. Joung J, Ladha A, Saito M, Kim NG, Woolley AE, Segel M, et al. Detection of SARS-CoV-2 with SHERLOCK one-pot testing. N Engl J Med. 2020; 383 (15): 1492-4.
8. Dash PK, Kaminski R, Bella R, Su H, Mathews S, Ahooyi TM, et al. Sequential LASER ART and CRISPR treatments eliminate HIV-1 in a subset of infected humanized mice. Nat Commun. 2019; 10 (1): 2753.
9. Kim S, Kim D, Cho SW, Kim J, Kim JS. Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. Genome Res. 2014; 24 (6): 1012-9.
10. Vakulskas CA, Behlke MA. Evaluation and reduction of CRISPR off-target cleavage events. Nucleic Acid Ther. 2019; 29 (4): 167-74.
11. DeWitt MA, Corn JE, Carroll D. Genome editing via delivery of Cas9 ribonucleoprotein. Methods. 2017; 121-122: 9-15.
12. Liang X, Potter J, Kumar S, Ravinder N, Chesnut JD. Enhanced CRISPR/Cas9-mediated precise genome editing by improved design and delivery of gRNA, Cas9 nuclease, and donor DNA. J Biotechnol. 2017; 241: 136-46.
13. Mout R, Ray M, Yesilbag Tonga G, Lee YW, Tay T, Sasaki K, et al. Direct cytosolic delivery of CRISPR/Cas9-ribonucleoprotein for efficient gene editing. ACS Nano. 2017; 11 (3): 2452-8.
14. Gundry MC, Brunetti L, Lin A, Mayle AE, Kitano A, Wagner D, et al. Highly efficient genome editing of murine and human hematopoietic progenitor cells by CRISPR/Cas9. Cell Rep. 2016; 17 (5): 1453-61.
15. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science. 2012; 337 (6096): 816-21.
16. Liu C, Zhang L, Liu H, Cheng K. Delivery strategies of the CRISPR-Cas9 gene-editing system for therapeutic applications. J Control Release. 2017; 266: 17-26.
17. Sørensen HP, Sperling-Petersen HU, Mortensen KK. A favorable solubility partner for the recombinant expression of streptavidin. Protein Expr Purif. 2003; 32 (2): 252-9.
18. Dyson MR, Shadbolt SP, Vincent KJ, Perera RL, McCafferty J. Production of soluble mammalian proteins in *Escherichia coli*: identification of protein features that correlate with successful expression. BMC Biotechnol. 2004; 4: 32.
19. Klock HE, Koesema EJ, Knuth MW, Lesley SA. Combining the polymerase incomplete primer extension method for cloning and mutagenesis with microscreening to accelerate structural genomics efforts. Proteins. 2008; 71 (2): 982-94.
20. Yang WC, Welsh JP, Lee J, Cooke JP, Swartz JR. Solubility partner IF2 domain I enables high yield synthesis of transducible transcription factors in *Escherichia coli*. Protein Expr Purif. 2011; 80 (1): 145-51.
21. Sati SP, Singh SK, Kumar N, Sharma A. Extra terminal residues have a profound effect on the folding and solubility of a *Plasmodium falciparum* sexual stage-specific protein over-expressed in *Escherichia coli*. Eur J Biochem. 2002; 269 (21): 5259-63.
22. Malhotra A. Tagging for protein expression. Methods Enzymol. 2009; 463: 239-58.
23. Moreno JM, Dyrskjøt L, Kristensen JE, Mortensen KK, Sperling-Petersen HU. Characterization of the domains of *Escherichia coli* initiation factor IF2 responsible for recognition of the ribosome. FEBS Lett. 1999; 455 (1-2): 130-4.

[DOI: 10.61882/JoMMID.14.2.149]

24. Butt TR, Edavettal SC, Hall JP, Mattern MR. SUMO fusion technology for difficult-to-express proteins. Protein Expr Purif. 2005; 43 (1): 1-9.
25. Butt TR, Jonnalagadda S, Monia BP, Sternberg EJ, Marsh JA, Stadel JM, et al. Ubiquitin fusion augments the yield of cloned gene products in *Escherichia coli*. Proc Natl Acad Sci U S A. 1989; 86 (8): 2540-4.
26. Wittliff JL, Wenz LL, Dong J, Nawaz Z, Butt TR. Expression and characterization of an active human estrogen receptor as a ubiquitin fusion protein from *Escherichia coli*. J Biol Chem. 1990; 265 (35): 22016-22.
27. Wang Z, Li H, Guan W, Ling H, Wang Z, Mu T, et al. Human SUMO fusion systems enhance protein expression and solubility. Protein Expr Purif. 2010; 73 (2): 203-8.
28. Zuo X, Li S, Hall J, Mattern MR, Tran H, Shoo J, et al. Enhanced expression and purification of membrane proteins by SUMO fusion in *Escherichia coli*. J Struct Funct Genomics. 2005; 6 (2-3): 103-11.
29. Zuo X, Mattern MR, Tan R, Li S, Hall J, Sterner DE, et al. Expression and purification of SARS coronavirus proteins using SUMO-fusions. Protein Expr Purif. 2005; 42 (1): 100-10.
30. Zhang S, Shen J, Li D, Cheng Y. Strategies in the delivery of Cas9 ribonucleoprotein for CRISPR/Cas9 genome editing. Theranostics. 2021; 11 (2): 614-48.
31. Rosano GL, Ceccarelli EA. Recombinant protein expression in *Escherichia coli*: advances and challenges. Front Microbiol. 2014; 5: 172.
32. Peleg Y, Unger T. Resolving bottlenecks for recombinant protein expression in *E. coli*. Methods Mol Biol. 2012; 800: 173-86.
33. Francis DM, Page R. Strategies to optimize protein expression in *E. coli*. Curr Protoc Protein Sci. 2010; Chapter 5: Unit 5.24.
34. Ahmad I, Nawaz N, Darwesh NM, Ur Rahman S, Mustafa MZ, Khan SB, et al. Overcoming challenges for amplified expression of recombinant proteins using *Escherichia coli*. Protein Expr Purif. 2018; 144: 12-8.
35. Marblestone JG, Edavettal SC, Lim Y, Lim P, Zuo X, Butt TR. Comparison of SUMO fusion technology with traditional gene fusion systems: enhanced expression and solubility with SUMO. Protein Sci. 2006;15 (1): 182-9.
36. Aslam M, Gull I, Abbas Z, Athar M. Soluble expression of IFNα2-Tα1 fusion protein in *Escherichia coli* by N-terminal SUMO fusion and its anti-proliferative activity. Pak J Zool. 2018; 50 (6): 2291-8.
37. Bi J, Tiong E, Koo YS, Zhou W, Wong FT. Further characterization and engineering of an 11-amino acid motif for enhancing recombinant soluble protein expression. Microb Cell Fact. 2025; 24 (1): 122.
38. Arakawa T, Akuta T. Beyond purification: evolving roles of fusion tags in biotechnology. Curr Issues Mol Biol. 2025; 47 (9): 768.

Cite this article:
Barkhordari F, Hosseini Shalmaei M, Naderi S, Raigani M, Kayalha Y, Bayat E, Hashemi Aval NS, Adeli A, Davami F. The Influence of IF2 and SUMO Fusion Tags on the Expression and Solubility of the SpCas9 Protein. J Med Microbiol Infect Dis, 2026; 14 (2): 149-158. DOI: 10.61882/JoMMID.14.2.149.

[DOI: 10.61882/JoMMID.14.2.149]