Enhanced Expression of Rabies Virus Surface G-Protein in *Escherichia coli* using SUMO Fusion

Ankit Singh · Dinesh Yadav · Krishan Mohan Rai · Meenal Srivastava · Praveen C. Verma · Pradhyumna K. Singh · Rakesh Tuli

Abstract Fusion systems are known to increase the expression of difficult to express recombinant proteins in soluble form to facilitate their purification. Rabies glycoprotein was also tough to express at sufficient level in soluble form in both *E. coli* and plant. The present work was aimed to over-express and purify this membrane protein from soluble extract of *E. coli*. Fusion of Small Ubiquitin like Modifier (SUMO) with rabies glycoprotein increased ~1.5 fold higher expression and ~3.0 fold solubility in comparison to non-fused in *E. coli*. The SUMO fusion also simplified the purification process. Previously engineered rabies glycoprotein gene in tobacco plants provides complete protection to mice, but the expression was very low for purification. Our finding demonstrated that the SUMO-fusion was useful for enhancing expression and solubility of the membrane protein and again proves to be a good alternative technology for applications in biomedical and pharmaceutical research.

Keywords Enhanced expression · Rabies glycoprotein · Recombinant protein · Solubility · SUMO fusion

Abbreviations

- FLAP 5-lipoxygenase-activating protein
- rgp Rabies glycoprotein gene
- RGP Rabies glycoprotein
- SARS-CoV Severe acute respiratory syndrome coronavirus
- SUMO Small ubiquitin-like modifier
- S-RGP SUMO fusion with rabies glycoprotein
- UBL Ubiquitin-like modifiers
- UDPs Ubiquitin-domain proteins

1 Introduction

Rabies is an acute contagious infection of the central nervous system caused by the rabies virus. At least, 60,000 human deaths occur worldwide annually due to rabies [18]. Rabies virus genome encodes five major proteins- nucleoprotein (N), phosphoprotein (P), matrix protein (M), glycoprotein (G) and RNA-dependent RNA polymerase (L) [3, 5]. The rabies glycoprotein (G) plays an important role in viral pathogenesis and functions as a protective antigen [28]. Rabies virus surface glycoprotein gene (*rgp*) was engineered and transformed in tobacco plants previously in our laboratory. This gives complete protection to mice against the live virus challenge and provides a foundation for plant based vaccine [1]. However, the expression was very low which make it an intricate task to purify in functionally active form in sufficient quantity from the plant. There is a need of milligram quantities of rabies glycoprotein (RGP) in its native form to understand its physical properties which remains unclear previously.
Covalent modification of cellular proteins by the ubiquitin-like modifier SUMO (small ubiquitin-like modifier) regulates various cellular processes, such as nuclear transport, signal transduction and stabilization of proteins. Ubiquitin-like proteins fall into two classes: the first class, ubiquitin-like modifiers (UBLs) functions as modifiers in a manner analogous to that of ubiquitin. Examples of UBLs are SUMO, Rub1 (also called Nedd8), Apg8 and Apg12. Proteins of the second class include parkin, RAD23 and DSK2, are designated as ubiquitin-domain proteins (UDPs) [2, 9]. These proteins contain domains that are related to ubiquitin but are otherwise unrelated to each other. In contrast to UBLs, UDPs are not conjugated to other proteins. Once covalently attached to cellular targets, SUMO regulates protein: protein and protein: DNA interactions, as well as localization and stability of the target protein. Sumoylation occurs in most eukaryotic systems and SUMO is highly conserved from yeast to humans. SUMO has an apparent molecular weight of ~12 kDa. SUMO and ubiquitin show about 18% homology only, but both possess a common three-dimensional structure characterized by a tightly packed globular fold with β-sheets wrapped around the α-helix [11, 17].

Attachment of a highly stable structure (such as that of ubiquitin or SUMO) at the N-terminus of a partner protein increases the yield by increasing stability [10]. SUMO fused with proteins of interest dramatically enhances the expression and promotes solubility by helping in correct folding of the proteins. It is known that ubiquitin exerts chaperoning effect on fused proteins in E. coli and increases their yield and solubility [12]. The solubilizing effect of ubiquitin and ubiquitin-like proteins was also explained in part by the outer hydrophilicity and inner hydrophobicity of the core structure of ubiquitin and SUMO, exerting a detergent-like effect on otherwise insoluble proteins [2, 9, 17]. The presence of a highly soluble fusion tag also enables increased soluble expression of protein products (up to 50% of the total protein) in E. coli, thereby preventing product aggregation or the formation of inclusion bodies [25]. In addition, toxicity of soluble expression products against the host microorganisms is often prevented when using a fusion partner, as was earlier noticed by our group during ASAL (garlic leaf lectin) expression and purification [27].

Optimizing conditions for expression and purification of membrane proteins for functional and structural studies is laborious and time-consuming process. This also requires an efficient expression and purification process. This process can be accelerated using fusion technology. Membrane proteins such as severe acute respiratory syndrome coronavirus (SARS-CoV) protease and 5-lipoxygenase-activating protein (FLAP) [31] have successfully been expressed and purified using fusion technology. Thus, we cloned SUMO fragment and constructed an expression cassette containing SUMO-rgp to utilize this fusion system to increase the expression level of difficult to express rabies glycoprotein in soluble and active form. In this study the SUMO fusion enhanced the solubility as well as simplified the purification of rabies glycoprotein from E. coli.

2 Materials and Methods

2.1 Cloning of Rabies Glycoprotein Gene

In our laboratory, rgp was previously cloned in pET 19b vector (Novagen) known as pSA33 [14]. pET-SUMO vector was obtained from LifeSensors and utilized for cloning of rgp gene. pET-SUMO and pET 19b vector were under the control of IPTG-inducible T7 and T7 lac promoters, respectively. Forward primer 5’ AGCGGCTCTCC AGGTAAGTTCCCTATCTACACTATCT 3’ and reverse primer 5’ TTACTAGGAATCTACATCAACTCATCT TCTCG 3’ were used to introduce Bsal and BamHI restriction site at 5’ and 3’ ends, respectively. The rgp gene was amplified with Deep Vent DNA polymerase (NEB, England) from pSA33 [14] using above primers. Amplified fragment of ~1.56 kb was digested with Bsal and BamHI, cleaned with PCR cleanup kit (Sigma chemicals) and ligated with linear SUMO vector according to manufacturer’s instruction. Ligation mixture was transformed into DH5α strain of E. coli (Strategene, India). Positive clones were screened by restriction digestion with NcoI and BamHI. The rgp gene cloned into the pET-SUMO vector was named as pSRGP.

2.2 Over-Expression of Rabies Glycoprotein

The construct of pSRGP and pSA33 were transformed in BL21 (RIL) codon plus strain of E. coli (Strategene, India). Primary culture was raised by inoculating single isolated colony in 5 ml Luria Bertini (LB) media (HiMedia, India) amended with 50 μg/ml each of kanamycin and chloramphenicol antibiotics (Sigma) for pSRGP transformants; while 100 and 50 μg/ml of ampicillin (Sigma) and chloramphenicol, respectively, for pSA33 transformants. Primary culture was grown at 37 °C, overnight at 200 rpm in an orbital shaker. Secondary culture (500 ml) was inoculated with primary at ratio 1:1000 in LB and incubated at 37 °C and 200 rpm until the OD600 nm reached 0.8. The culture was induced with 0.5 mM IPTG for pSRGP and 1.0 mM IPTG for pSA33 and was further grown at 20 °C for 6 h. Cells were harvested by centrifugation at 7000Xg. The pellet was washed twice with ST buffer (20 mM Tris–Cl, pH 8.0; 150 mM NaCl) and suspended in 20 mM Tris–Cl, pH 8.0 (1/40 volume of original culture) and digested with 0.5 mg/ml of lysozyme. Spheroplasts were sonicated on Sonic Vibra cell for 5 min (with 5 s pulse and 2 s
After sonication 1% sodium deoxycholate was added and incubated on shaker for 2 h and then centrifuged at 25,000×g. Supernatant was then dialyzed overnight with buffer containing 20 mM Tris–Cl, pH 8.0; 0.1% sodium deoxycholate and used for RGP purification.

2.3 Purification of the Rabies Glycoprotein

2.3.1 Ion-Exchange Chromatography

Soluble fraction of RGP did not bind onto the Sepharose-Q as well as Ni–NTA matrix (GE Healthcare, USA) in 20 mM Tris–Cl buffer pH 8.0 with 0.1% sodium deoxycholate, while other contaminating proteins bound on the columns. Therefore, we followed negative purification strategy for this protein in which RGP was enriched at this stage.

2.3.2 Size-Exclusion Chromatography

Five ml of enriched protein was loaded on Superdex-200 (GE Healthcare, USA) pre-equilibrated with 20 mM Tris–Cl buffer containing 20 mM Tris–Cl, pH 8.0; 0.1% sodium deoxycholate, and 150 mM NaCl and 2 ml fractions were collected during the entire run.

2.4 ELISA

Total soluble protein was estimated by the Bradford reagent (Bio-Rad, USA). A 96-well micro-titre plate was coated with Total soluble protein was estimated by the Bradford reagent (Bio-Rad, USA). A 96-well micro-titre plate was coated with Total soluble protein was estimated by the Bradford reagent and 2 ml fractions were collected during the entire run.

2.5 SDS-PAGE and Western Blotting Analysis

Samples were mixed with equal volume of sample loading buffer (25 mM Tris–Cl, pH 6.8, 2% SDS, 200 mM DTT, 20% glycerol and 0.25% Bromophenol Blue), boiled for 5 min and centrifuged at 13,000×g for 2 min. The supernatant was electrophoresed on a 10% discontinuous SDS-PAGE and transferred to PVDF membrane (Bio-Rad) in blotting buffer (25 mM Tris base, 192 mM glycine and 20% methanol). All washings, blocking and antibody dilutions were made in TBS-T buffer (100 mM Tris pH 7.5, 150 mM NaCl, 0.1% Tween-20). The membrane was blocked with 5% non fat dry milk powder (Bio-Rad) for 1 h. The primary and secondary antibodies were used at 5,000 fold and 20,000 fold dilutions, respectively. Blot was developed with DAB color developer kit (Bangalore Genie).

2.6 Immunization of Balb/c Mice, Titration of Antibody Response in Immunized Mice

Balb/c mice (five in each group) were immunized intraperitoneally with 25 μg purified glycoprotein. Commercial (Rabipur, Aventis Pharma Ltd., Mumbai) inactivated rabies virus vaccine and phosphate buffer saline were used as positive and negative control, respectively. In all cases, complete Freund’s adjuvant was used at the time of priming. Three boosters were given on 7th and 14th day with Freund’s incomplete adjuvant and on 28th day without adjuvant. The mice were bled on the 35th day, from the retro orbital sinus for the estimation of anti-rabies antibody titer in serum.

3 Results

3.1 Cloning of rgp Gene in pET-SUMO Vector

The amplified coding sequence of rgp was cloned in pET–SUMO vector in a way that the complete ORF produces the S-RGP protein (Fig. 1a). The recombinant plasmids were screened by PCR using rgp gene specific primers. PCR amplified the expected band of 1.56 kb (Fig. 1b). Restriction digestion with NcoI and BamHI produced the bands of pET SUMO vector 5.5 kb and rgp gene of 1.85 kb (Fig. 1c), which confirmed the correct orientation and insertion of rgp in pET-SUMO vector.

3.2 Over-Expression of Rabies Glycoprotein

The cloned rgp gene in pET 19b and pET-SUMO vector are predicted to encode recombinant protein with respective molecular masses of ~60.6 and ~72 kDa (Fig. 2a). In order to obtain highest possible quantity of soluble fusion protein, we optimized the expression conditions and found that solubility of fusion protein was dramatically increased, if the culture was grown at 37 °C until cell density reaches OD600 0.8–1.0, then induced with 0.5 and 1 mM IPTG, respectively and further grown at 20 °C for 6 h (Fig. 2). The SDS-PAGE analysis showed the induction of ~60 kDa RGP (Fig. 3a, lane 3) and ~72 kDa S-RGP (Fig. 3b, lane 3). On densitometric analysis an increased expression level (~1.5 folds higher) of fusion protein (S-RGP) was observed as compared to non-SUMO fused RGP in E. coli. Total soluble proteins were extracted in extraction buffer and their densitometric analysis establishes that the amount of soluble S-RGP (Fig. 3b, lane 4) is ~3.0 times more as compared to RGP (Fig. 3a, lane 4). Whereas, RGP was poorly soluble and most of the protein was in the form of inclusion bodies (Fig. 3a, lane 5).
3.3 Purification of S-RGP

The soluble proteins extracted from *E. coli* were purified under native conditions on Biologic Duo Flow FPLC system (Bio-Rad). We tried to purify S-RGP using Ni–NTA under native conditions but S-RGP didn’t bind to the column, while it bound under denaturing conditions and purified up to homogeneity. So, we tried an anion exchanger matrix Sepharose Q fast flow (GE healthcare, USA). The column was equilibrated with 20 mM Tris pH 8.0 and 0.1% sodium deoxycholate. The protein was loaded on to the column, but most of S-RGP remains unbound while most of contaminating proteins bind to the column. The unbound S-RGP was concentrated up to 5 ml in Amicon ultra filtration tubes and loaded on to Superdex-200 (GE healthcare, USA) pre-equilibrated with 20 mM Tris pH 8.0, 150 mM NaCl and 0.1% sodium deoxycholate. The S-RGP eluted near to void volume. It shows that protein is in complex form and contaminating proteins were completely removed (Fig. 4a). This protein was probed either with anti-His antibody or polyclonal anti-rabies glycoprotein antibody (Fig. 4b). Besides this contamination, S-RGP was purified up the level of more than 80% (Table 1).

4 Discussion

Rabies is acute progressive encephalitis. It is caused by a promiscuous neurotropic virus of Lyssavirus genus of family Rhabdoviridae. The disease spreads through domestic and wildlife animals. Rabies is a major zoonosis...
of significant public health concern in many parts of the world, especially in developing countries where rabies is endemic among dogs [18]. The first rabies vaccine [21], consisted of subcutaneous inoculation of spinal cord suspension, derived from rabid rabbits. Since then, a continuous effort is going on for the improvement of the vaccine [22, 23]. The virus genome encodes five major proteins of which the G-protein of rabies virus has been identified as the major viral antigen that induces protective immunity [6]. In recent years, plants are emerging as a promising alternative to produce safe and effective therapeutic proteins in plants [4, 14, 26]. Recombinant proteins expressed in plants have shown sufficient promise to warrant human clinical trials [26]. In case of rabies glycoprotein, stable expression in tomato plants was reported [16], while complete protection was also shown when mice were injected with rabies G-proteins expressed and purified from tobacco plants [1]. We need X-ray crystallographic and NMR data to elucidate the structural details of rabies glycoprotein. The prerequisite of such kind of studies are the milligram quantities of glycoprotein and obtaining this much amount of protein from plant bioreactor/mammalian system is a very tough job as it is a membrane protein. However, still there is a scope for the modification to be made in the gene for the expression of large quantities in various hosts. Although, the glycoprotein expressed in E. coli is non-glycosylated, and doesn’t give protection against challenge [13, 30], yet it is able to induce the IgG production (~1.5 times of negative control; data not shown) in mice, if administered in intra-peritoneum along with adjuvant. An increase in the rabies virus neutralizing antibody titers, after priming with any authentic vaccine and boosting with E. coli expressed, non-glycosylated G-protein has also been reported [24].

The expression and purification of membrane proteins at high levels is an arduous task because the membrane proteins are strongly hydrophobic and they fold incorrectly and may form aggregate, leading to either rapid degradation or accumulation as inclusion bodies. But if the membrane proteins are expressed as a fusion protein, this problem is alleviated up to some extent [31]. Ubiquitin has been reported to have chaperon like activity when fused with some protein. This property of ubiquitin not only increases the expression of fusion protein [2, 19], but also increases the solubility of fusion protein. SUMO is an ubiquitin like protein consist of approximately 100 AA which is highly conserved in eukaryotes and is not yet

Fig. 2 SDS-PAGE profile of induction of SRGP at different time interval at 20 °C. Lane 1st standard molecular mass marker; 2nd un-induced SRGP culture; 3rd induced with 0.5 mM IPTG and harvested after 2 h; 4th harvested after 4 h; 5th harvested 5 h; 6th harvested after 6 h; and 7th sample was harvested at overnight

Fig. 3 a SDS-PAGE profile of induction of pSA33. Lane 1st was standard molecular mass marker; 2nd un-induced RGP culture; 3rd induced with 1.0 mM IPTG and harvested after 6 h; 4th total soluble protein obtained after sonication; 5th was protein remain in pellet.

Fig. 3 b SDS-PAGE profile of induction of pSRGP. Lane 1st was standard molecular mass marker; 2nd un-induced S-RGP culture; 3rd induced with 0.5 mM IPTG and harvested after 3 h; 4th total soluble protein obtained after sonication; 5th was protein remain in pellet
reported from prokaryotes [29]. A rationale role of SUMO in facilitating the solubility of insoluble protein is that the inner SUMO core is a dense hydrophobic globular structure and the protein surface is hydrophilic and water soluble which is analogous to amphipathic detergents [2, 15].

Detailed studies on the properties of RGP have been hindered by a lack of availability of purified protein in many cases. The purification of RGP from *E. coli* is cumbersome and expensive. The expression of *rgp* without fusion in *E. coli* is also very difficult due to its low expression and solubility. During this study we have expressed the rabies G-protein as a SUMO fusion in *E. coli*. Milligram quantities of S-RGP protein was produced and purified under native conditions with a novel method using negative purification where most of the contaminating proteins bind on to the column and protein of interest remains in the unbound fraction. By this method we were able to enrich the G-protein up to 80% purity (Table 1).

Table 1 Percent of purity and fold purification of S-RGP

| Purification step               | Volume (ml) | Total protein (mg) | Amount of SRGP | Purity (%) | Purification factor |
|--------------------------------|-------------|--------------------|----------------|------------|--------------------|
| Crude cell free extract        | 40          | 140                | 9.2            | 6.57       | 1                  |
| Unbound from Sepharose Q       | 10          | 25                 | 7.9            | 31.6       | 4.8                |
| Gel filtration (Superdex 200)  | 10.5        | 5.2                | 4.2            | 80.77      | 12.3               |

This work opens the possibility of using SUMO fusions for the expression of RGP in *E. coli* to express these in soluble form. The SUMO fusion, expression and purification protocols described here can greatly reduce the cost of producing RGP in good quantity. Complete solubility and high production of S-RGP in *E. coli* without affecting its

Fig. 4 Purification profile of S-RGP on SDS-PAGE. a The gel is stained with Coomassie brilliant blue. Lane 1st was standard molecular marker; 2nd un-induced total soluble protein obtained after sonication; 3rd induced total soluble protein obtained after sonication; 4th Sample after elution from IEC; 5th concentrated sample and 6th sample purified from gel filtration column (Superdex 200). b Western blot of S-RGP purification. Lane 1st was standard pre-stained molecular marker; 2nd Total soluble protein obtained after sonication; 3rd unbound protein from Sephrose Q; 4th eluted sample from IEC; 5th concentrated sample and 6th was sample purified from gel filtration column (Superdex 200).

Fig. 5 Protein profile of proteolytic digestion of SUMO-RGP. a Protein profile of proteolytic digestion of SUMO-RGP on 10% SDS-PAGE. 2 U of sumo protease enzyme was incubated overnight with 1 μg SUMO-RGP proteins. Lane 1st Marker, 2nd SUMO-RGP undigested, 3rd SUMO-RGP digested. b Protein profile of proteolytic digestion of SUMO-GFP on 12% SDS-PAGE. 0.2 U of sumo protease enzyme was incubated overnight with 1 μg SUMO-GFP proteins. Lane 1st SUMO-GFP undigested, 2nd SUMO-GFP digested.
property suggested that the fusion protein can be a preferred candidate molecule for producing large quantity of RGP. This result is consistent with previous reports that ubiquitin fusion does not inhibit biological activity of the protein [7, 20].

However, in our case we were not able to cleave the SUMO tag. This problem with membrane proteins is especially pronounced since they require detergents for their proper folding and detergents may decrease or diminish the proteolytic activity of SUMO protease 1 (Fig. 5a, b). Our results suggest the need to examine the SUMO cleavage facilitation for the purification and biosafety of deploying RGP vs. SUMO–RGP in the development of transgenic plants for edible vaccine production.

Acknowledgments Mr. Ankit Singh thanks Council of Scientific and Industrial Research (CSIR), New Delhi, India for the financial support in the form of Senior Research Fellowship for this research.

References

1. Ashraf S, Singh PK, Yadav DK, Shahnawaz M, Mishra S, Sawant SV, Tuli R (2005) J Biotechnol 119:1–14
2. Butt TR, Jonnalagadda S, Monia BP, Sternberg EJ, Marsh JA, Stadel JM, Ecker DJ, Crooke ST (1989) Proc Natl Acad Sci USA 86:2540–2544
3. Conzelmann K, Cox JH, Schneider LG, Thiel H (1990) Virology 175:485–499
4. Curtiss RI, Cardineau CA (1990) World Patent Application WO:90/02484
5. Coslett DG, Hollow BP, Obijeski JK (1980) J Gen Virol 49:161–180
6. Cox JH, Dietzschold B, Schneider LG (1977) Infect Immun 16:754–775
7. Guzzo CM, Yang DCH (2007) Protein Exp and Purif 54:166–175
8. Harlow E, Lane D (1999) Cold Spring Harbor, Cold Spring Harbor Laboratory Press NY
9. Jentsch S, Pyrowolakis G (2000) Trends Cell Biol 10:335–342
10. Kim KL, Baek SH, Chung SH (2002) J Cellu Physiol 191:257–268
11. Li SJ, Hochstrasser M (2000) Mol Cell Biol 20:2367–2377
12. Malakhov MP, Mattern MR, Malakhova OA, Drinker M, Weeks SD, Butt TR (2004) J Struct Funct Genom 5:75–86
13. Malek LT, Soostmeyer G, Garvin RT, James E (1984) In: Chanock RM and Lerner RA (ed) Modern Approaches to Vaccines. CSHL, New York, pp 203–208
14. Mason HS, Lam DMK, Amtzen CJ (1992) Proc Natl Acad Sci USA 89:11745–11749
15. McDonnell DP, Pike JW, Drutz DJ, Butt TR, O’Malley BW (1989) Mol Cell Biol 9:3517–3523
16. McGarvey PB, Hammond J, Dienelt MM, Hooper DC, Fu ZF, Dietzschold B (1995) Biotechnology 13:1484–1487
17. Melchior F (2000) Annu Rev Cell Dev Biol 16:591–626
18. Meslin FX, Stohr K (1997) In: Dodet B and Meslin FX (ed) Rabies Control in Asia. Elsevier, Paris, pp 15–18
19. Mishra S, Yadav DK, Tuli R (2006) J Biotechnol 127:95–108
20. Niiranen L, Espelid S, Karlsten CR, Mustonen M, Paulsen SM, Heikinheimo P, Willlassen NP (2007) Protein Exp and Purif 52:210–218
21. Pasteur L (1985) Compt Rendus Acad Sci Paris 101:765–774
22. Perrin P, Lafon M, Sureau P (1990) Adv Biotechnol Process 14:325–345
23. Plotkin S (1993) J Infect Dis 168:29–37
24. Rath A, Batra S, Choudhury D, Kapre SV, Rupprecht CE, Gupta SK (2005) Virus Res 113:143–152
25. Satakarni M, Curtis R (2011) Protein Exp and Purif. doi: 10.1016/j.pep.2011.04.015
26. Tiwari S, Verma PC, Singh PK, Tuli R (2009) Biotech Adv 27:449–467
27. Upadhyay SK, Rai P, Saurabh S, Singh R, Chandrashekhark V, Verma PC, Singh PK, Tuli R (2010) J Biotech 146:1–8
28. Wiktor TJ, Gyorgy E, Schlumberger D, Sokol F, Koprowski H (1973) J Immunol 110:269–276
29. Yeh ET, Gong L, Kamitani T (2000) Gene 248:1–14
30. Yelverton E, Norton S, Obijeski JF, Goeddel FV (1983) Science 219:614–620
31. Zuo X, Li S, Hall J, Mattern MR, Tran H, Shoo J, Tan R, Weiss SR, Butt TR (2005) J Struct Funct Genomics 6:103–111