



A novel silk fibroin protein–based fusion system for enhancing the expression of nanobodies in *Escherichia coli*

Jianli Yu^{1,2} · Yang Guo¹ · Yi Gu^{1,2} · Xiyang Fan¹ · Fei Li³ · Haipeng Song³ · Rui Nian¹ · Wenshuai Liu¹

Received: 1 December 2021 / Revised: 22 February 2022 / Accepted: 26 February 2022 / Published online: 4 March 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract

Nanobodies show a great potential in biomedical and biotechnology applications. Bacterial expression is the most widely used expression system for nanobody production. However, the yield of nanobodies is relatively low compared to that of eukaryotic systems. In this study, the repetitive amino acid sequence motifs (GAGAGS) found in silk fibroin protein (SFP) were developed as a novel fusion tag (SF-tag) to enhance the expression of nanobodies in *Escherichia coli*. SF-tags of 1 to 5 hexapeptide units were fused to the C-terminus of 4G8, a nanobody against human epididymis protein 4 (HE4). The protein yield of 4G8 variants was increased by the extension of hexapeptide units and achieved a 2.5~7.1-fold increase compared with that of untagged 4G8 (protein yield of 4G8-5C=0.307 mg/g vs that of untagged 4G8=0.043 mg/g). Moreover, the fusion of SF-tags not only had no significant effect on the affinity of 4G8, but also showed a slight increase in the thermal stability of SF-tag-fused 4G8 variants. The fusion of SF-tags increased the transcription of 4G8 by 2.3~7.0-fold, indicating SF-tags enhanced the protein expression at the transcriptional level. To verify the applicability of the SF-tags for other nanobody expression, we further investigated the protein expression of two other anti-HE4 nanobodies 1G8 and 3A3 upon fusion with the SF-tags. Results indicated that the SF-tags enhanced the protein expression up to 5.2-fold and 5.7-fold for 1G8 and 3A3, respectively. For the first time, this study reported a novel and versatile fusion tag system based on the SFP for improving nanobody expression in *Escherichia coli*, which may enhance its potential for wider applications.

Key points

- A silk fibroin protein–based fusion tag (SF-tag) was developed to enhance the expression of nanobodies in *Escherichia coli*.
- The SF-tag enhanced the nanobody expression at the transcriptional level.
- The fusion of SF-tag had no significant effect on the affinity of nanobodies and could slightly increase the thermal stability of nanobodies.

Keywords Nanobody · *Escherichia coli* · Expression-enhancing tag · Silk fibroin protein · Hexapeptide

Introduction

Nanobodies are derived from the heavy-chain antibodies (HCABs), which are naturally present in all Camelidae. These HCABs have an antigen recognition part composed of single variable domains, referred to as the variable domain of the heavy chain of a heavy-chain antibody (VHH). Owing to a small diameter of 2.5 nm and a height of 4 nm, these single variable domains are accordingly named as nanobodies by its original Belgian developer Ablynx®. In recent years, nanobodies have attracted extensive attention in the biotechnology and biomedical industries owing to their nanoscale size, high stability, strong affinity and specificity, water solubility, superior cryptic cleft accessibility, and sustainable

Jianli Yu and Yang Guo contributed equally to this work

✉ Wenshuai Liu
liuws@qibebt.ac.cn

¹ CAS Key Laboratory of Biobased Materials, Qingdao Institute of Bioenergy and Bioprocess Technology, Chinese Academy of Sciences, No. 189 Songling Road, Qingdao 266101, China

² University of Chinese Academy of Sciences, No. 19(A) Yuquan Road, Beijing 100049, China

³ Shenzhen Innova Nanobodi Co., Ltd, No. 7018 Caitian Road, Shenzhen 518000, China

source. With the peculiar properties, nanobodies have been developed into various research tools for basic research, including affinity purification, chaperone-assisted crystallization, protein–protein interaction, and cellular bioimaging (Wang et al. 2016). The small structure of nanobodies facilitates their fast extravasation, deep tumor penetration, and rapid clearance in vivo, which highly benefits tumor-targeted molecular imaging. Several nanobody targeting probes for cancer diagnosis have entered clinic trials (Liu et al. 2021). Apart from the application in vivo tumor imaging, nanobodies have been developed to detect disease biomarkers in the blood or other biopsies. Further, nanobodies have also been developed for therapeutic use. Caplacizumab, a bivalent nanobody, was approved by the European Medicines Agency (EMA) and the US Food and Drug Administration (FDA) for the treatment of patients with thrombotic thrombocytopenic purpura (Scully et al. 2019). Moreover, several nanobody therapeutics are under clinical investigations for a variety of human diseases including breast cancer, lung diseases, brain tumors, inflammatory diseases, and infectious diseases (Jovcevska and Muyldermans 2020). Recently, nanobodies that neutralize SARS-CoV-2 were identified by several independent research groups to assist in combating the pandemic diseases (Custodio et al. 2020; Hanke et al. 2020; Huo et al. 2020; Li et al. 2021; Schoof et al. 2020; Xiang et al. 2020). Hence, the nanobody is considered a next generation of antibody-derived tool in antigen-related recognition, and will occupy a momentous portion of the antibody reagent market.

The expression of recombinant proteins is one of the most critical parts for protein applications in biotechnology and pharmaceuticals as it influences the bioactivity, cost, and yields of products. Compared to conventional antibodies, nanobodies can be expressed in *Escherichia coli* cells instead of mammalian cells owing to their single gene format as well as non-requirement for post-translational modifications (Liu and Huang 2018). As nanobodies contain a disulfide linkage between complementarity-determining regions (CDRs) 1 and 3, they are generally expressed in the oxidative environment of the *E. coli* periplasm to facilitate correct folding and disulfide bond formation (Billen et al. 2017; Salema and Fernandez 2013). Thus far, many nanobodies are successfully expressed in the periplasm of *E. coli* with the strains TG1 (Clontech), Shuffle® T7 (New England Biolabs), Origami (Novagen), Rosetta-gami (Novagen), and BL21 (DE3) (Novagen) (Liu and Huang 2018). However, a large amount of nanobodies is still only expressed at a very low yield due to the insufficiency of chaperones in the periplasm. Although various protein fusion tags (e.g., GST, MBP, Trx, NusA, SUMO, Fh8) have been developed to enhance protein expression and solubility (Ki and Pack 2020), large fusion tags may interfere with the protein structures or impair the

affinities of nanobodies. Therefore, short peptide fusion tags to enhance nanobody expression have continued to be investigated and developed.

Silk fibroin protein (SFP) is a natural polymer-based protein from the silkworm *Bombyx mori*. The *B. mori* SFP is comprised of a heavy chain (HC) and a light chain (LC) linked by a disulfide bond (Shimura et al. 1976). P25, a 25-kDa glycoprotein, is non-covalently linked to these chains to maintain the integrity of the complex (Tanaka et al. 1999). The LC is hydrophilic and relatively elastic. The HC consists of 12 crystalline domains interspersed with amorphous regions. Each crystalline domain consists of subdomains of ~70 residues, which primarily starts with the repeated glycine-, alanine-, and serine-rich hexapeptides (GAGAGS) and terminates with a GAAS tetrapeptide (Zhou et al. 2001). SFP exhibits numerous attractive properties, such as excellent mechanical property, good biocompatibility, controllable biodegradability, aqueous processability, and ease of functionalization and patterning. Therefore, SFP has been employed in a multitude of biomaterial applications (Kundu et al. 2013). In the study by Wang et al., SFP showed antiproliferative effects on lung cancer cells both in vivo and in vitro by inducing cell apoptosis (Wang et al. 2019).

In this study, we investigated the potential of SFP hexapeptide (GAGAGS) blocks (SF-tag) as an expression enhancement tag for nanobody production in the periplasm of *E. coli* cells. We selected 4G8, a nanobody against human epididymis protein 4 (HE4) that was screened from an immune library, as the model nanobody (Liu et al. 2018). The SF-tag was fused to the C-terminus of 4G8. The resulting recombinant fusion proteins (termed 4G8-SF-tag fusions) were assessed for the production yields, affinities, protein secondary structures, thermal stabilities, and transcription levels compared to the untagged form. Meanwhile, the expression enhancement effect of SF-tag was also tested on two other anti-HE4 nanobodies 1G8 and 3A3. The results imply that the SF-tag is an excellent and versatile fusion tag to increase the expression levels of nanobodies in *E. coli* cells.

Materials and methods

Strains, plasmids, and reagents

E. coli strain TG1 (Clontech) was used as the host for phage display. *E. coli* strain DH5 α (Invitrogen) and BL21 (DE3) (Novagen) cells were used as the host for gene cloning and nanobody expression, respectively. HEK293 cells (Thermo Fisher Scientific) were used as the host for HE4-Fc expression. A pMES4 vector containing a signal peptide pelB at the N-terminus of multiple cloning sites (MCSs) was used

for expressing recombinant nanobodies in *E. coli*. The pcDNATM3.1(+) vector was used for expressing recombinant HE4-Fc protein in HEK293 cells.

DNA molecules were synthesized by Tsingke (China). Lipofectamine 3000 was purchased from Invitrogen (USA). Restriction endonucleases and Gibco® FreeStyleTM 293 Expression Medium were purchased from Thermo Fisher Scientific (USA). Ni Sepharose 6 Fast Flow, HiTrapTM Protein G HP column, and CM5 sensor chip were purchased from Cytiva (USA). eStain L1 Protein Staining Kit was obtained from GenScript (China). BCA Protein Assay Kit was obtained from Tiangen Biotech (China). TriPure Isolation Reagent was purchased from Merck (Germany). 5×All-In-One RT MasterMix Kit was purchased from ABM (Canada). SYBR Green qPCR mix and 96-well fluorescent qPCR plates were purchased from Monad (China).

Expression and purification of HE4-Fc

HE4 (NCBI Accession No. NP_006094.3) was fused with an Fc fragment of *Camelus bactrianus* (NCBI Accession No. ALB75438.1) at the C-terminus for immunization (Supplementary Fig. S1). A whole HE4-Fc gene was synthesized and cloned into the pcDNATM3.1(+) vector by *EcoRI/HindIII* restriction enzymes sites. HEK293 cells (5×10^7) were transfected with 100 µg of the pcDNATM3.1(+)-HE4-Fc expression vector using Lipofectamine 3000 following the manufacturer's recommended procedure. After transfection, cells were cultured in Gibco® FreeStyleTM 293 Expression Medium at 37 °C, 100 rpm, and 5% CO₂ for 3 days. After cells were removed by centrifugation, the supernatant was filtered through 0.22-µm filters for purification of HE4-Fc.

Purification of HE4-Fc was performed using a HiTrapTM Protein G HP column (1 mL). The column was equilibrated with 10 mM PBS (pH 7.2) at 1 mL/min, and the supernatant was injected. After a wash step with 10 mM PBS (pH 7.2), the bound protein was eluted using 0.1 M glycine-HCl (pH 2.7). The UV absorbance was monitored at 280 nm. The protein peak was neutralized using 2 M Tris and dialyzed into 10 mM PBS (pH 7.2). The purity of HE4-Fc was evaluated by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE), and the protein concentration was determined through BCA Protein Assay Kit as mentioned above.

Isolation of HE4-specific nanobodies

Alpaca immunization and phage display were performed as described previously (Pardon et al. 2014). In brief, an adult alpaca was immunized five times at 2-week intervals. The immunogen contained a total of ~400 µg of HE4-Fc mixed with an equal volume of Gerbu adjuvant for each immunization. Four days after the final immunization, RNA was isolated from peripheral blood lymphocytes (PBLs)

with TriPure Isolation Reagent and reverse transcribed into cDNA. Nanobody sequences were amplified and cloned into a pMES4 phagemid. TG1 cells were transformed with this library by electroporation. Then, cells were inoculated with the VCSM13 helper phage, and the resulting phage was enriched in four consecutive rounds of phage display on HE4 immobilized on 96-well plates. After the fourth round of phage display, individual colonies were randomly selected to identify HE4-specific nanobodies by performing phage ELISA.

Construction of expression plasmids

Three anti-HE4 nanobodies (termed 3A3, 1G8, and 4G8) were selected for the expression research in this work. Among them, 4G8 was designed with an in-framed (GAGAGS)_n ($n = 1, 2, 3, 4, \text{ or } 5$) sequence at the C-terminus of 4G8. The 4G8-(GAGAGS)_n nucleotide sequences were chemically synthesized and inserted between the pelB signal peptide and His₆ tag of the pMES4 plasmid by Tsingke (China), yielding expression plasmids pMES4-4G8-(GAGAGS)_n ($n = 1, 2, 3, 4, 5$). Additionally, 3A3 and 1G8-(GAGAGS)_n expression plasmids were constructed by restriction enzyme digestion of pMES4-4G8-(GAGAGS)_n plasmids and ligated with digested 3A3 or 1G8 sequences. Briefly, the synthesized pMES4-4G8-(GAGAGS)_n plasmids were digested with restriction enzymes *PstI* and *BstEII* at 37 °C for 3 h to obtain a linearized vector pMES4-(GAGAGS)_n. Next, restriction enzyme *PstI*- and *BstEII*-digested 3A3 and 1G8 were ligated into the linearized pMES4-(GAGAGS)_n vector by T4 ligase at 16 °C overnight and the ligated products were transformed into *E. coli* DH5α competent cells, yielding expression plasmids pMES4-1G8-(GAGAGS)_n and pMES4-3A3-(GAGAGS)_n, respectively ($n = 1, 2, 3, 4, 5$).

Expression and purification of recombinant nanobodies

All recombinant nanobodies were expressed in *E. coli* BL21 (DE3) and purified using nickel immobilized metal affinity chromatography. To express 4G8-SF-tag fusions, 1G8-SF-tag fusions, and 3A3-SF-tag fusions, the nanobody expression plasmids were transformed into *E. coli* BL21 (DE3). The *E. coli* clones containing the expression vectors were cultivated in LB medium (10 g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl) supplemented with ampicillin (100 µg/mL), and grown at 37 °C until OD₆₀₀ reached 1.0. Following induction by 1 mM IPTG at 30 °C for 20 h, the cell pellets were harvested by centrifugation at 12,000 g for 10 min at room temperature (RT). The periplasmic proteins were extracted as described previously (Fan et al. 2019) and separated by 15% SDS-PAGE. To compare protein expression

levels with and without SF-tags, each sample was loaded with the same volume (20 μL) from 20 mL of periplasmic protein extracts, recovered from 200 mL culture broths. The resolved protein bands were visualized using eStain L1 Protein Staining Kit.

The recombinant proteins were purified by immobilized metal affinity chromatography using a 1-mL gravity column prepacked with Ni Sepharose 6 Fast Flow. In short, the column was pre-equilibrated with the binding buffer (20 mM Tris-HCl, 0.3 M NaCl, pH 8.0). Then, periplasmic protein extracts were loaded onto the column at a constant flow rate of 10 drops per minute. After washing with the washing buffer (20 mM imidazole, 20 mM Tris-HCl, 0.3 M NaCl, pH 8.0), the bound proteins were eluted using the elution buffer (250 mM imidazole, 20 mM Tris-HCl, 0.3 M NaCl, pH 8.0). The purified proteins were evaluated by SDS-PAGE and dialyzed into 10 mM PBS (pH 7.2). Protein concentrations were determined using BCA Protein Assay Kit. In brief, 25 μL of the standard solution (bovine serum albumin, BSA) and samples were added into a 96-well microplate, and then 200 μL working solution was added to each well. After incubation at 37 °C for 30 min, the absorbance was measured at 562 nm on a microplate reader (Spark, Switzerland). A standard curve of BSA was prepared to determine the protein concentration of each sample.

Affinity analysis by surface plasmon resonance binding assay

The affinity values (K_D) of 4G8 variants, 1G8 variants, and 3A3 variants were estimated by surface plasmon resonance (SPR) technology using the BIACORE T100 system (GE, USA). HE4-Fc (100 $\mu\text{g}/\text{mL}$) was immobilized on the CM5 sensor chip surface through amine coupling. Afterwards, the chip was incubated in 1 M ethanolamine-HCl (pH 8.5) to completely block the remaining activated NHS groups. Running buffer HBS-N (10 mM HEPES, 150 mM NaCl, pH 7.4) was pumped at a constant flow rate for setting up the assay. Nanobody variants (5 $\mu\text{g}/\text{mL}$) were sequentially injected at a flow rate of 30 $\mu\text{L}/\text{min}$ for 120 s and then dissociated for 1000 s. After each injection, the chip was regenerated with 10 mM glycine-HCl (pH 1.5) for 30 s and then 0.1 M NaOH for 60 s at a flow rate of 30 $\mu\text{L}/\text{min}$. SPR signals were measured in response unit (RU) and recorded as a sensorgram. The kinetic parameters, including the association rate constant (k_{on}), the dissociation rate constant (k_{off}), and the equilibrium dissociation constant (K_D), were calculated based on a 1:1 fitting model.

Circular dichroism

A Chirascan circular dichroism spectrometer (Applied Photophysics, UK) was utilized to determine the secondary structure

and the melting temperature (T_m) of nanobody variants. The proteins were diluted to 0.25 mg/mL by 50 μM PBS (pH 7.4). Far-UV (190–250 nm) circular dichroism (CD) measurements were carried out in a quartz cuvette with a 1-mm optical path length at 1-nm resolution. Raw ellipticity data, given in millidegrees (mdeg), was converted to molar ellipticity [θ] in deg cm^2/dmol by the following equation (Greenfield 2006):

$$[\theta] = (\text{mdeg} \times MRW) \div (\text{pathlength} \times [\text{nanobody}])$$

where the mean residue weight MRW = (molecular weight of the nanobody in Da/number of backbone amino acids), pathlength = cell path length in millimeters, and $[\text{nanobody}]$ = concentration of nanobody variants in milligrams per milliliter. Thermal unfolding experiments was followed at 200 nm with CD measurements taken every 2 °C from 20 to 85 °C with a temperature increase of 1 °C/min using a Peltier junction temperature controller. The melting temperature analysis was performed in triplicate, and thermal denaturation data were fitted using the Global 3 software (version 3.1.0.1, Applied Photophysics, UK).

Real-time RT-PCR analysis

After inducing with 1 mM IPTG at 30 °C for 6 h, the *E. coli* cells were pelleted and lysed by 1 mL TriPure Isolation Reagent. These samples were mixed with 200 μL chloroform and incubated at room temperature for 5 min, and then centrifuged at 4 °C 12,000 rpm for 15 min. RNA was extracted from the water phase through isopropanol precipitation. After washing by 1 mL 75% ethanol, the RNA pellets were allowed to air-dry at room temperature and dissolved in 15 μL nuclease-free water. The quality and quantity of the RNAs were examined using a NanoVue Plus spectrophotometer (GE Healthcare, UK) and confirmed by electrophoresis.

cDNAs were synthesized using 5 \times All-In-One RT MasterMix Kit according to the manufacturer's procedure. The transcription levels of 4G8 were determined by real-time RT-PCR using SYBR Green qPCR Mix on a LightCycler 480 II thermal cycler system (Roche, Germany) (forward primer: TGGAGGCACCTTCAGTAACTATAAC; reverse primer: GATCCTTCACGGAGTCTGCATAG). The internal standard gene, 16S rRNA, was used as an internal control to normalize the transcription levels of the samples (forward primer: CTC TTGCCATCGGATGTGCCCA; reverse primer: CCAGTG TGGCTGGTCATCCTCTCA). The real-time RT-PCR was set up as follows: 95 °C for 30 s and 40 cycles of 95 °C for 10 s, 60 °C for 30 s.

Results

Effect of the repetitive SFP hexapeptide motifs on nanobody production

We immunized one alpaca with HE4-Fc. A phage display library was generated, and we performed four consecutive rounds of phage display, followed by a phage ELISA-based binding screen. We isolated three nanobodies (termed 4G8, 1G8, and 3A3) that bind specifically to HE4 (Supplementary Fig. S2).

To investigate the influence of the SF-tag in nanobody expression, nanobody 4G8 against HE4 was selected as the model nanobody. Previous studies showed that N-terminal fusion tags might affect the affinities of nanobodies. Thus, the SF-tag was fused at the C-terminus of 4G8. To facilitate target purification, we also added a His₆ tag at the C-terminus. To examine the effect of the numbers of SFP hexapeptide units on nanobody production, 1 to 5 SFP hexapeptide units were fused to the C-terminus of 4G8 to form 4G8-1C, 4G8-2C, 4G8-3C, 4G8-4C, and 4G8-5C, respectively (Fig. 1). 4G8 without SFP hexapeptide units was used as a reference. The presence of the SF-tags on the protein molecular weight was minor owing to the small

size of one SFP hexapeptide unit (0.4 kDa) (Table 1). The 4G8-SF-tag fusion genes were inserted into pMES4 expression plasmids and successfully expressed in the periplasm of *E. coli* BL21 (DE3) with a molecular weight identical to their expected sizes (Fig. 2A). There was no significant difference in *E. coli* cell growth among these recombinants, indicating that the SF-tags were suitable for protein fusion labels (Fig. 2B). As shown in SDS-PAGE analysis (Fig. 2A), the 4G8-SF-tag fusions showed a seeming yield increase compared with that of untagged-4G8. To assess whether the SFP hexapeptide units could be used as an expression-enhancing tag, the 4G8-SF-tag fusions were purified using nickel immobilized metal affinity chromatography (Fig. 2C) and the yields of purified proteins were calculated. As shown in Fig. 2D, all the 4G8-SF-tag fusions showed remarkable increases in the protein yield compared with that of untagged-4G8, resulting in fold increases in the protein yield of 4G8-1C, 4G8-2C, 4G8-3C, 4G8-4C, and 4G8-5C by 2.5-fold, 4.0-fold, 5.8-fold, 5.2-fold, and 7.1-fold, respectively. Thus, the protein yields of 4G8-SF-tag fusions were increased greatly by the extension of SFP hexapeptide units, indicating that the length of SFP hexapeptide units had a significant effect on protein production.

Fig. 1 Schematic diagrams of the target proteins with or without the C-terminal SF-tags. SF-tags were fused with 4G8 via the flexible (GGGS)₃ linkers (red broken lines)

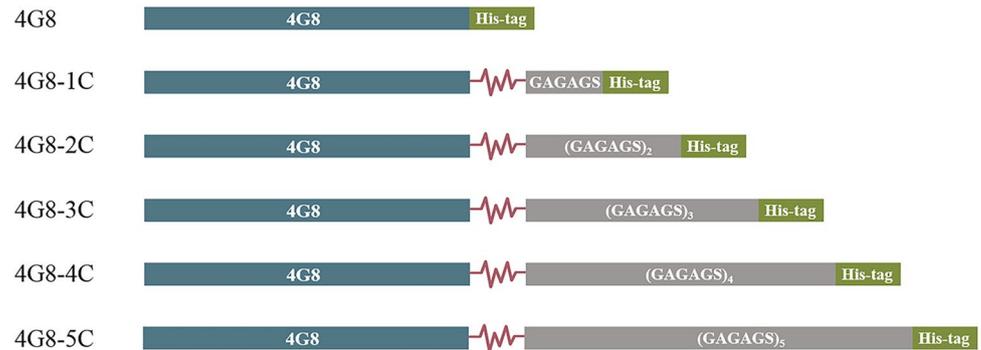


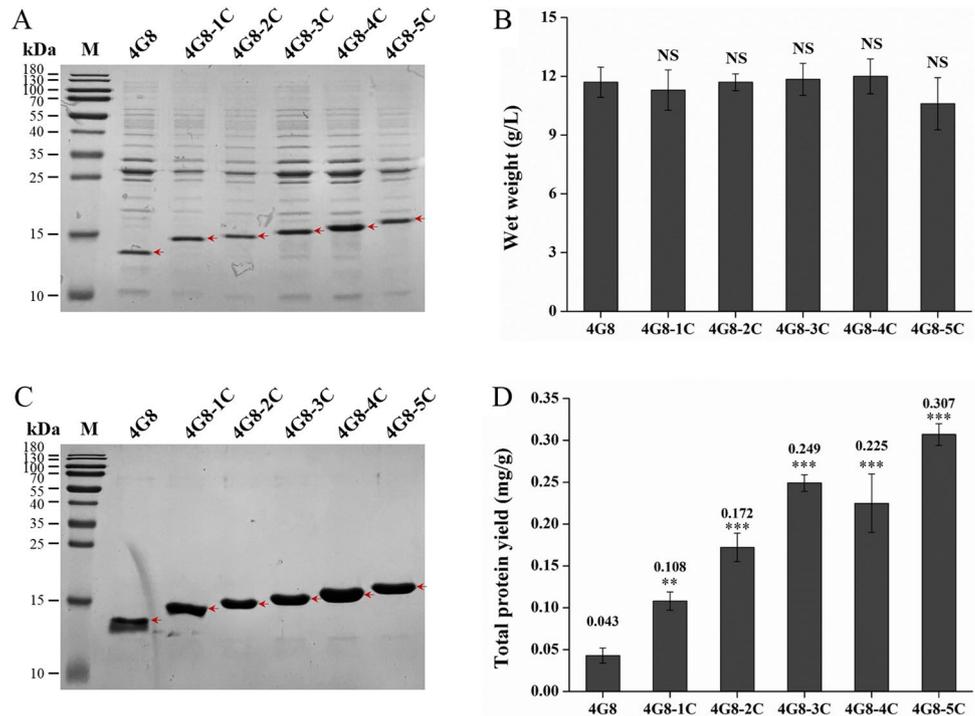
Table 1 Properties of 4G8-SF-tag fusions

Protein	pI	Molecular weight (Da)	GRAVY	mRNA free energy (kcal/mol)	Protein yield (mg/g)	<i>T_m</i> (°C)	<i>k_{on}</i> (M ⁻¹ s ⁻¹)	<i>k_{off}</i> (s ⁻¹)	<i>K_D</i> (M)
4G8	8.67	13,951.7	-0.678	-62.1	0.043	61.8	5.938 × 10 ⁴	1.413 × 10 ⁻³	2.379 × 10 ⁻⁸
4G8-1C	8.67	15,126.5	-0.624	-80.4	0.108	62.2	2.792 × 10 ⁵	2.764 × 10 ⁻³	9.900 × 10 ⁻⁹
4G8-2C	8.67	15,526.7	-0.588	-83.7	0.172	63.4	4.437 × 10 ⁵	4.333 × 10 ⁻³	9.767 × 10 ⁻⁹
4G8-3C	8.67	15,927.0	-0.555	-84.7	0.249	63.4	2.091 × 10 ⁵	2.926 × 10 ⁻³	1.400 × 10 ⁻⁸
4G8-4C	8.67	16,327.3	-0.525	-90.1	0.225	63.7	3.799 × 10 ⁵	3.563 × 10 ⁻³	9.381 × 10 ⁻⁹
4G8-5C	8.67	16,727.5	-0.496	-89.8	0.307	63.3	2.797 × 10 ⁵	3.226 × 10 ⁻³	1.153 × 10 ⁻⁸

The mRNA free energy was analyzed by an mRNA structure analyzer (<http://rna.urmc.rochester.edu/RNAstructureWeb>). The GRAVY was analyzed by a GRAVY calculator (http://bioinformatics.org/sms2/protein_gravy.html). The association and dissociation rate constants (*k_{on}* and *k_{off}*) were measured by SPR and used to calculate the equilibrium dissociation constant (*K_D*)

pI isoelectric point, GRAVY grand average of hydrophathy

Fig. 2 Effect of the SF-tags on the production of 4G8. **A** SDS-PAGE analysis of 4G8-SF-tag fusion expression in the periplasm of *E. coli* BL21 (DE3). M, marker. **B** Wet weight of cells that expressed 4G8-SF-tag fusion in shake flask fermentation. Error bars, standard deviations from three replicate flasks. *P* values were determined by Student's *t* test. NS, not significant. **C** SDS-PAGE analysis of purified 4G8-SF-tag fusions. M, marker. **D** Comparison of the protein yields in mg protein per gram of wet weight of cells. Error bars, standard deviations from three independent experiments. *P* values were determined by Student's *t* test. ***P* < 0.01; ****P* < 0.001



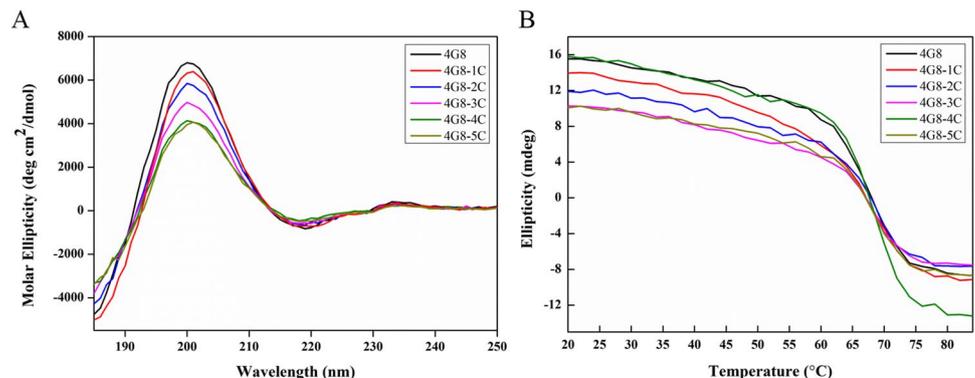
Affinity and stability characterization of the 4G8-SF-tag fusions

The antibody–antigen binding characteristic is generally evaluated by affinity analysis. To investigate whether the SF-tags interfere with the affinity of 4G8, we assessed the binding of 4G8-SF-tag fusions to HE4 by SPR binding analysis. The antigen, HE4 with an Fc fusion (HE4-Fc), was expressed by transient expression in HEK293 cells and purified by protein G chromatography. The binding of 4G8-SF-tag fusions to HE4-Fc was analyzed with a series of 4G8-SF-tag fusion concentrations ranging from 0.05 to 2 μ M. The binding affinities were 23.8 nM for 4G8, 9.9 nM for 4G8-1C, 9.8 nM for 4G8-2C, 14.0 nM for 4G8-3C, 9.4 nM for 4G8-4C, and 11.5 nM for 4G8-5C (Table 1). Thus, the SPR values showed no significant difference in

binding affinities between 4G8 and 4G8-SF-tag fusions, indicating that SF-tags did not affect the binding affinity of 4G8.

CD spectroscopy is a tool to analyze the secondary structure of protein. We firstly examined various secondary structures (α -helix, β -sheet, and random coils) of 4G8-SF-tag fusions by far-UV CD (190–250 nm). The spectra displayed one positive peak centered at 200 nm and one negative peak centered at 220 nm, which are the characteristic peaks for the β -sheet structure (Fig. 3A). Although the overall shapes of the far-UV CD spectra from 4G8 and 4G8-SF-tag fusions were similar, the intensity of the peak at 200 nm was lower for 4G8-SF-tag fusions, and the peak intensity was lower with the extension of the SFP hexapeptide unit length, indicating that the structure of 4G8 might change upon the addition of SF-tags.

Fig. 3 Representative far-UV CD spectra of 4G8 in the presence and absence of SF-tags. **A** Far-UV scans (185–250 nm) were performed at 24 °C. **B** *T*_ms were determined by CD spectroscopy at 200 nm



Thermal stability is the unique property of nanobodies. The melting temperature (T_m) of a protein represents the temperature at which it exists in a state of equilibrium between unfolded and folded states. Thus, temperature-induced unfolding experiments were conducted to determine the T_m of 4G8-SF-tag fusions. Figure 3B shows the ellipticity at 200 nm plotted versus temperature for 4G8 with and without SF-tags. The T_m values of 4G8, 4G8-1C, 4G8-2C, 4G8-3C, 4G8-4C, and 4G8-5C were 61.8 °C, 62.2 °C, 63.4 °C, 63.4 °C, 63.7 °C, and 63.3 °C, respectively (Table 1). In the presence of SF-tags, the T_m values showed a minor increase, indicating a slight increase in the thermal stability of the protein. The hydrophobic interaction was reported to play an important role in increasing the thermal stability of proteins (Stepanenko et al. 2008). The GRAVY value of a protein is a measure of its hydrophobicity or hydrophilicity. Positive GRAVY values indicate hydrophobicity, and negative values mean hydrophilicity. Consistent with the T_m values determined by CD, the GRAVY values of 4G8-SF-tag fusions become less negative, indicating a stronger hydrophobicity of proteins which may contribute to the higher thermal stability of 4G8-SF-tag fusions (Table 1).

SF-tags enhanced the transcription of 4G8 protein

To investigate the effect of SF-tag on nanobody expression at the transcription level, we performed real-time RT-PCR analysis on the encoding gene of 4G8. As expected, we observed a substantial increase in the mRNA level of the 4G8-SF-tag fusions (Fig. 4). In addition, the increase in 4G8 transcription was correlated with the size of the SFP hexapeptide unit. Compared with untagged 4G8, the mRNA levels of 4G8-1C, 4G8-2C, 4G8-3C, 4G8-4C, and 4G8-5C increased by 2.3-fold, 2.6-fold, 3.9-fold, 3.4-fold, and 7.0-fold, respectively. The results indicated that 4G8 with longer SFP hexapeptide units showed a higher transcription level, which was consistent with the protein yield results.

The structural stability of mRNA plays an important role in both transcription and translation of proteins (Mao et al. 2014; Shpaer 1985; Zur and Tuller 2012). The mRNA free energy for 4G8 with or without SF-tags was calculated by an mRNA structure analyzer (<http://rna.urmc.rochester.edu/RNAstructureWeb>). A lower free energy indicates a higher mRNA stability. As shown in Table 1, SF-tag fusions showed a significant decrease in the mRNA free energy, lowering the free energy of -18.3 kcal/mol for 4G8-1C, -21.6 kcal/mol for 4G8-2C, -22.6 kcal/mol for 4G8-3C, -28.0 kcal/mol for 4G8-4C, and -27.7 kcal/mol for 4G8-5C. Hence, there was a negative correlation between the mRNA abundance and mRNA free energy. Further, combined with protein expression results, it seemed that the more stable the mRNA of 4G8-SF-tag fusions become, its expression value increases.

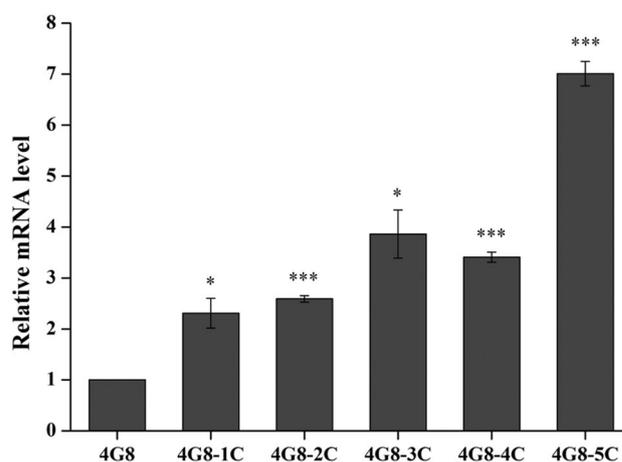


Fig. 4 Relative mRNA levels of 4G8 in SF-tag fusions. The mRNA level of untagged 4G8 was set to 1. Error bars, standard deviations from three independent experiments. P values were determined by Student's t test. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$

Fusion of the SF-tags is effective for other nanobodies

We isolated a series of nanobodies against HE4. To further explore the applicability of SF-tags on other nanobody expression, nanobodies 1G8 and 3A3 were fused with the SF-tags and expressed in *E. coli* cells. Without the SF-tag fusion, 1G8 showed a yield of 0.049 mg/g of wet *E. coli* cells, while the fusion with the SF-tags increased the yield level of 1G8 variants by 2.5-fold for 1G8-1C, 2.2-fold for 1G8-2C, 2.9-fold for 1G8-3C, 2.6-fold for 1G8-4C, and 5.2-fold for 1G8-5C (Fig. 5A, B). Nanobody 3A3 was expressed poorly in the periplasm of *E. coli* cells as indicated by the SDS-PAGE gel in Fig. 5C that no 3A3 protein band could be visualized. However, the expression level of 3A3 was greatly enhanced upon the fusion with SF-tags, resulting in the yield level increasing by 2.1-fold for 3A3-1C, 1.6-fold for 3A3-2C, 2.3-fold for 3A3-3C, 4.9-fold for 3A3-4C, and 5.7-fold for 3A3-5C (Fig. 5D). These results indicated that the SF-tags could enhance the expression of other nanobodies.

Further, we examined the binding affinities of 1G8 variants and 3A3 variants with antigen HE4-Fc by SPR binding analysis. As shown in Table 2, the dissociation constant (K_D) values were 0.7 nM for 1G8, 0.9 nM for 1G8-1C, 1.3 nM for 1G8-2C, 1.6 nM for 1G8-3C, 2.1 nM for 1G8-4C, 2.2 nM for 1G8-5C, 64.9 nM for 3A3, 55.8 nM for 3A3-1C, 28.7 nM for 3A3-2C, 54.0 nM for 3A3-3C, 36.5 nM for 3A3-4C, and 42.6 nM for 3A3-5C. As expected, there was no significant difference in binding affinities between the untagged nanobodies and nanobodies with SF-tag fusions, suggesting that the SF-tags did not affect the affinities of nanobodies.

Fig. 5 Effect of the SF-tags on the production of 1G8 and 3A3. **A** SDS-PAGE analysis of 1G8-SF-tag fusion expression in the periplasm of *E. coli* BL21 (DE3). M, marker. **B** Comparison of the protein yields of the purified 1G8-SF-tag fusions. Error bars, standard deviations from three independent experiments. *P* values were determined by Student's *t* test. ***P* < 0.01; ****P* < 0.001. **C** SDS-PAGE analysis of 3A3-SF-tag fusion expression in the periplasm of *E. coli* BL21 (DE3). M, marker. **D** Comparison of the protein yields of the purified 3A3-SF-tag fusions. Error bars, standard deviations from three independent experiments. *P* values were determined by Student's *t* test. ****P* < 0.001

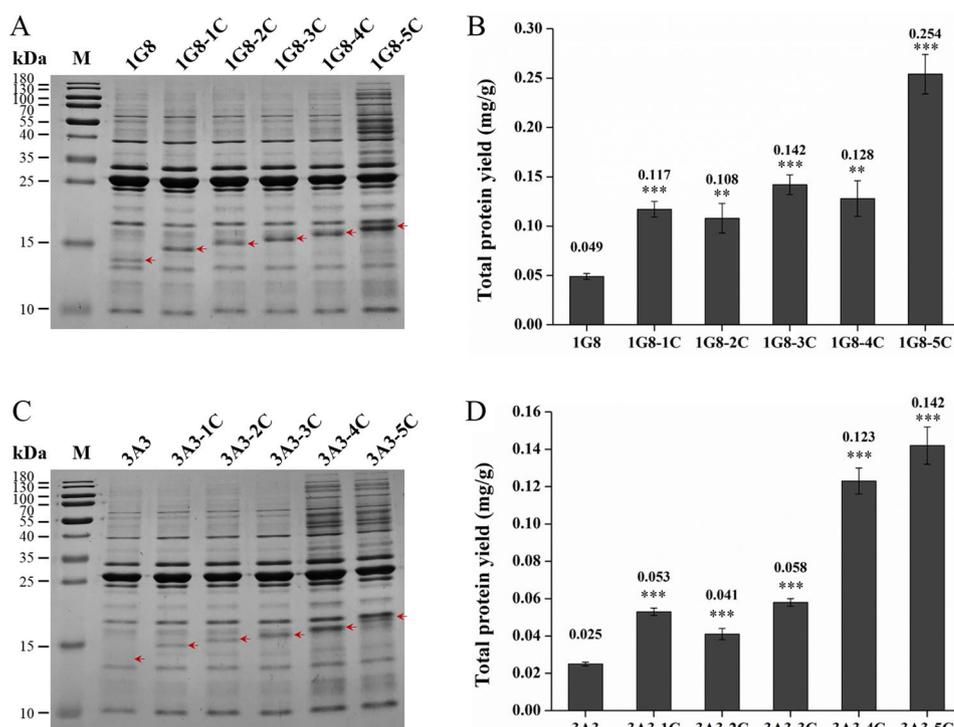


Table 2 Properties of 1G8-SF-tag fusions and 3A3-SF-tag fusions

Protein	Protein yield (mg/g)	k_{on} ($M^{-1} s^{-1}$)	k_{off} (s^{-1})	K_D (M)
1G8	0.049	3.519×10^6	2.585×10^{-3}	7.346×10^{-10}
1G8-1C	0.117	3.791×10^6	3.586×10^{-3}	9.460×10^{-10}
1G8-2C	0.108	2.548×10^6	3.222×10^{-3}	1.264×10^{-9}
1G8-3C	0.142	3.415×10^6	5.533×10^{-3}	1.620×10^{-9}
1G8-4C	0.128	1.420×10^6	3.025×10^{-3}	2.131×10^{-9}
1G8-5C	0.254	1.543×10^6	3.335×10^{-3}	2.162×10^{-9}
3A3	0.025	7.854×10^4	5.097×10^{-3}	6.490×10^{-8}
3A3-1C	0.053	4.414×10^4	2.462×10^{-3}	5.577×10^{-8}
3A3-2C	0.041	1.396×10^5	4.007×10^{-3}	2.869×10^{-8}
3A3-3C	0.058	2.392×10^4	1.292×10^{-3}	5.399×10^{-8}
3A3-4C	0.123	4.337×10^4	1.585×10^{-3}	3.654×10^{-8}
3A3-5C	0.142	2.694×10^4	1.148×10^{-3}	4.261×10^{-8}

The association and dissociation rate constants (k_{on} and k_{off}) were measured by SPR and used to calculate the equilibrium dissociation constant (K_D)

Discussion

Over the decades, nanobodies have attracted an increasing amount of attention from the pharmaceutical and biotechnology industries owing to their peculiar properties. One of the main limitations in the successful commercialization of any recombinant protein is the production

yield and purification level. However, a large number of nanobodies are currently mainly produced in *E. coli* cells with a low yield. Although several fusion tags have been used to functionalize nanobodies (Gotzke et al. 2019; Vegiani et al. 2020), few of them demonstrated an expressing enhancement effect on nanobodies. Previously, we found an interesting result that the expression of nanobodies was enhanced by the fusion with an SFP hexapeptide (GAGAGS). The SFP is a widely abundant natural protein derived from the silkworm cocoons, which has been extensively used as biomaterials in biological and biomedical fields (Kundu et al. 2013). To our knowledge, there is no study on the effect of SFP on the expression level of recombinant proteins. We systematically assessed the effect of SFP hexapeptide on the expression of nanobodies in the present study. Three nanobodies against HE4 (4G8, 1G8, 3A3) were selected as the model nanobodies and designed with different numbers of SFP hexapeptide units (SF-tags) at the C-terminus. To avoid interfering with the protein structure and alleviate metabolic burden on *E. coli*, short peptide tags (up to 5 SFP hexapeptide units, 30 residues) were fused to the nanobodies. All the nanobodies with SF-tag fusions showed a significant increase in protein expression level compared to the untagged nanobodies. Meanwhile, there was a trend in enhanced expression level with an increasing number of the SFP hexapeptide fusion tags. Therefore, 10-, 30-, 40-, 50-SFP hexapeptide units were fused to another model nanobody 11C12, which was isolated from immune libraries to

target carcinoembryonic antigen (CEA) (Lin et al. 2020). As expected, the yield of the 11C12 variants was greatly enhanced by the extended SF-tags as compared to the untagged 11C12 (data not shown). Hence, these results indicated that the SF-tags can be used as a novel fusion tag to enhance nanobody expression in *E. coli* cells.

Several mechanisms have been proposed to explain the expression-enhancing effects of fusion tags. For instance, fusion tags can enhance the protein solubility by attracting or acting as the molecular chaperones (e.g., MBP, NusA, and SUMO tags), slowing down the protein translation (e.g., NusA tag), or regulating the formation of disulfide bonds (e.g., Trx and PDI tags) (Ki and Pack 2020). However, nanobodies themselves show an excellent solubility owing to a large number of hydrophilic residues in frame region 2 (FR2). Therefore, we speculated that the enhanced expression of nanobodies with SF-tag fusions may not be due to promoting protein solubility.

Furthermore, transcription is another key factor in protein expression. To investigate whether the increased yield of SF-tag fusions is due to an increased transcription, we performed real-time RT-PCR analysis on the encoding gene of 4G8. As expected, we observed a substantial increase in the mRNA level of 4G8 in the fusion constructs. In addition, the increase in 4G8 transcription was positively correlated with the size of tags, which was consistent with the expression yield results. The mRNA secondary structure and stability have been identified as the major factors that influence several cellular processes, including transcript splicing (Patterson et al. 2002), ribosome abundance and its binding (Mao et al. 2014), and regulation of gene expression level and its accuracy (Grunberg-Manago 1999; Mao et al. 2014; Shpaer 1985). Based on our prediction of the mRNA free energy using a structure analyzer, the mRNA sequence of untagged 4G8 had a higher free energy than those of the 4G8-SF-tag fusion sequences, indicating that the mRNA structure of SF-tagged proteins is more stable than that of the untagged one. The results indicated that the stability of the mRNA secondary structure holds a positive correlation with mRNA abundance, and high abundance of mRNA favors a greater protein abundance. These findings are homologous to the findings of Victor et al. that the mRNA stability positively regulated the mRNA expression level in *Saccharomyces cerevisiae* (Victor et al. 2019). Greater mRNA stability tends to have a higher protein abundance, presumably by preventing the aggregation of mRNA molecules (Zur and Tuller 2012), leading to a high translation efficiency (Mao et al. 2014) or increasing the accuracy of translation (Shpaer 1985). However, some studies found that the decrease in mRNA structure stability contributed to an increase in mRNA expression, as ribosomes must unwind every structure encountered during translation (Jia and Li 2005; Kudla et al. 2009; Nguyen et al. 2019). The counterintuitive findings suggest that the

stability of the mRNA structure affects the translation in a complex manner.

The SF-tags used in this study are mostly composed of hydrophobic amino acids such as glycine and alanine. Most hydrophobic residues are totally or partially buried in the protein matrix, and could create a strong hydrophobic interaction network in the protein (Stepanenko et al. 2008). Serine provides an extra hydroxyl group that is able to engage in the hydrogen bonds between the serine side chains (Mayen et al. 2015). Weak non-covalent forces such as hydrophobic interactions and hydrogen bonds have been identified to play an important role in increasing the structural and thermal stability of proteins (Kumar et al. 2000). Furthermore, the silk-like (GAGAGS) blocks have been shown to spontaneously form hydrogen-bonded β -sheet crystals, which impart thermal stability (Anderson 1998; Anderson et al. 1994). Hence, these may explain the thermostability-enhancing effect of SF-tags.

The advantage of SF-tags for nanobody expression is that the amino acid sequence is 30 residues or less, which does not impair the activity when fused to the protein of interest. Therefore, it may be not necessary to remove the fusion tags for the protein application compared with the large-sized tags. In addition, SFPs have been known to possess various biological attributes. As a biomaterial approved by FDA, SFPs show an excellent wound healing and antibacterial activity, which can be used for cell growth matrix and tissue engineering (Joseph and Raj 2012). Furthermore, the SFP-derived peptides were reported to have anticancer effects. In the study by Wang et al., silk fibroin peptides with a molecular weight of ~ 10 kDa inhibited the growth of human lung cancer cells by inducing the cell cycle arrest at the S phase and promoting cell apoptosis (Wang et al. 2019). In a later study, the melanin synthesis in B16 melanoma cells was inhibited with the silk fibroin peptide treatment via downregulation of microphthalmia-associated transcription factor (MITF) and tyrosinase expression (Wang et al. 2021). Therefore, apart from the expression-enhancing effect found in the present study, the fusion of the SF-tags might endow nanobodies with new functions for wider applications.

In conclusion, for the first time, the repetitive amino acid sequence motifs (GAGAGS) found in silk fibroin protein were developed as a novel and versatile fusion tag (SF-tag) to enhance the expression of nanobodies in *E. coli*. SF-tags of 1 to 5 hexapeptide units were fused to the C-terminus of three nanobodies against human epididymis protein 4 (4G8, 1G8, and 3A3). The protein yields were increased by the extension of hexapeptide units, resulting in a 2.5 ~ 7.1-fold increase for 4G8 variants, a 2.2 ~ 5.2-fold increase for 1G8 variants, and a 1.6 ~ 5.7-fold increase for 3A3 variants compared with those of untagged forms. Moreover, the fusion of SF-tags not only had no significant effect on the affinities of the nanobodies, but also showed a slight increase in thermal

stability. The fusion of SF-tags increased the transcription of 4G8 by 2.3~7.0-fold, indicating SF-tags enhanced the protein expression at the transcriptional level.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00253-022-11857-7>.

Author contribution YJL and LWS conceived the study and designed the experiments; YJL and GY conducted the experiments and analyzed the data; GUY and LF contributed to result collection and processing; LWS interpreted the data and wrote the manuscript; YJL and GY reviewed and edited the writing; and NR, SHP, and FXY reviewed the manuscript. All authors have read and approved the manuscript.

Funding This work was supported by Shandong Energy Institute (SEI) (SEI I202128).

Data availability All data generated or analyzed during this study are included in this published article.

Declarations

Ethics approval The animal study underwent approval by the Department of Science & Technology of Shandong Province (SYXK(Lu)20160004).

Conflict of interest The authors declare no competing interests.

References

- Anderson JP (1998) Morphology and crystal structure of a recombinant silk-like molecule, SLP4. *Biopolymers* 45(4):307–321. [https://doi.org/10.1002/\(SICI\)1097-0282\(19980405\)45:4%3c307::AID-BIP5%3e3.0.CO;2-P](https://doi.org/10.1002/(SICI)1097-0282(19980405)45:4%3c307::AID-BIP5%3e3.0.CO;2-P)
- Anderson JP, Cappello J, Martin DC (1994) Morphology and primary crystal-structure of a silk-like protein polymer synthesized by genetically-engineered *Escherichia coli* bacteria. *Biopolymers* 34(8):1049–1058. <https://doi.org/10.1002/bip.360340808>
- Billen B, Vincke C, Hansen R, Devoogdt N, Muyldermans S, Adriaensens P, Guedens W (2017) Cytoplasmic versus periplasmic expression of site-specifically and bioorthogonally functionalized nanobodies using expressed protein ligation. *Protein Expr Purif* 133:25–34. <https://doi.org/10.1016/j.pep.2017.02.009>
- Custodio TF, Das H, Sheward DJ, Hanke L, Pazicky S, Pieprzyk J, Sorgenfrei M, Schroer MA, Gruzinov AY, Jeffries CM, Graewert MA, Svergun DI, Dobrev N, Remans K, Seeger MA, McInerney GM, Murrell B, Hallberg BM, Low C (2020) Selection, biophysical and structural analysis of synthetic nanobodies that effectively neutralize SARS-CoV-2. *Nat Commun* 11(1):5588. <https://doi.org/10.1038/S41467-020-19204-Y>
- Fan XY, Zhou YH, Yu JL, Li F, Chen Q, Song HP, Feng DX, Liu WS, Nian R, Xian M (2019) Non-affinity purification of a nanobody by void-exclusion anion exchange chromatography and multimodal weak cation exchange chromatography. *Sep Purif Technol* 225:88–96. <https://doi.org/10.1016/j.seppur.2019.05.068>
- Gotzke H, Kilisch M, Martinez-Carranza M, Sograte-Idrissi S, Rajavel A, Schlichthaerle T, Engels N, Jungmann R, Stenmark P, Opazo F, Frey S (2019) The ALFA-tag is a highly versatile tool for nanobody-based bioscience applications. *Nat Commun* 10:4403. <https://doi.org/10.1038/S41467-019-12301-7>
- Greenfield NJ (2006) Using circular dichroism spectra to estimate protein secondary structure. *Nat Protoc* 1(6):2876–2890. <https://doi.org/10.1038/nprot.2006.202>
- Grunberg-Manago M (1999) Messenger RNA stability and its role in control of gene expression in bacteria and phages. *Annu Rev Genet* 33:193–227. <https://doi.org/10.1146/annurev.genet.33.1.193>
- Hanke L, Perez LV, Sheward DJ, Das H, Schulte T, Moliner-Morro A, Corcoran M, Achour A, Hedestam GBK, Hallberg BM, Murrell B, McInerney GM (2020) An alpaca nanobody neutralizes SARS-CoV-2 by blocking receptor interaction. *Nat Commun* 11(1):4420. <https://doi.org/10.1038/S41467-020-18174-5>
- Huo JD, Le Bas A, Ruza RR, Duyvesteyn HME, Mikolajek H, Malinauskas T, Tan TK, Rijal P, Dumoux M, Ward PN, Ren JS, Zhou DM, Harrison PJ, Weckener M, Clare DK, Vogirala VK, Radecke J, Moynie L, Zhao YG, Gilbert-Jaramillo J, Knight ML, Tree JA, Buttigieg KR, Coombes N, Elmore MJ, Carroll MW, Carrique L, Shah PNM, James W, Townsend AR, Stuart DI, Owens RJ, Naismith JH (2020) Neutralizing nanobodies bind SARS-CoV-2 spike RBD and block interaction with ACE2. *Nat Struct Mol Biol* 27:846–854. <https://doi.org/10.1038/s41594-020-0469-6>
- Jia MW, Li YD (2005) The relationship among gene expression, folding free energy and codon usage bias in *Escherichia coli*. *FEBS Lett* 579(24):5333–5337. <https://doi.org/10.1016/j.febslet.2005.08.059>
- Joseph B, Raj SJ (2012) Therapeutic applications and properties of silk proteins from *Bombyx mori*. *Front Life Sci* 6(3–4):55–60. <https://doi.org/10.1080/21553769.2012.760491>
- Jovcevska I, Muyldermans S (2020) The therapeutic potential of nanobodies. *BioDrugs* 34(1):11–26. <https://doi.org/10.1007/s40259-019-00392-z>
- Ki MR, Pack SP (2020) Fusion tags to enhance heterologous protein expression. *Appl Microbiol Biotechnol* 104(6):2411–2425. <https://doi.org/10.1007/s00253-020-10402-8>
- Kudla G, Murray AW, Tollervey D, Plotkin JB (2009) Coding-sequence determinants of gene expression in *Escherichia coli*. *Science* 324(5924):255–258. <https://doi.org/10.1126/science.1170160>
- Kumar S, Tsai CJ, Nussinov R (2000) Factors enhancing protein thermostability. *Protein Eng* 13(3):179–191. <https://doi.org/10.1093/protein/13.3.179>
- Kundu B, Rajkhowa R, Kundu SC, Wang XG (2013) Silk fibroin biomaterials for tissue regenerations. *Adv Drug Del Rev* 65(4):457–470. <https://doi.org/10.1016/j.addr.2012.09.043>
- Li T, Cai H, Yao H, Zhou B, Zhang N, van Vlissingen MF, Kuiken T, Han W, GeurtsvanKessel CH, Gong Y, Zhao Y, Shen Q, Qin W, Tian XX, Peng C, Lai Y, Wang Y, Hutter CAJ, Kuo SM, Bao J, Liu C, Wang Y, Richard AS, Raoul H, Lan J, Seeger MA, Cong Y, Rockx B, Wong G, Bi Y, Lavillette D, Li D (2021) A synthetic nanobody targeting RBD protects hamsters from SARS-CoV-2 infection. *Nat Commun* 12(1):4635. <https://doi.org/10.1038/s41467-021-24905-z>
- Lin JT, Yu JL, Wang H, Xu YR, Li F, Chen XH, Liang YL, Tang JS, Wu LL, Zhou ZW, Chen CL, Liu MJ, Chun X, Nian R, Song HP (2020) Development of a highly thermostable immunoassay based on a nanobody-alkaline phosphatase fusion protein for carcinoembryonic antigen detection. *Anal Bioanal Chem* 412(8):1723–1728. <https://doi.org/10.1007/s00216-020-02456-4>
- Liu MM, Li L, Jin D, Liu YZ (2021) Nanobody—a versatile tool for cancer diagnosis and therapeutics. *Wires Nanomed Nanobiotechnol* 13(4):e1697. <https://doi.org/10.1002/Wnan.1697>
- Liu W, Song H, Chen Q, Yu J, Xian M, Nian R, Feng D (2018) Recent advances in the selection and identification of antigen-specific nanobodies. *Mol Immunol* 96:37–47. <https://doi.org/10.1016/j.molimm.2018.02.012>

- Liu YK, Huang H (2018) Expression of single-domain antibody in different systems. *Appl Microbiol Biotechnol* 102(2):539–551. <https://doi.org/10.1007/s00253-017-8644-3>
- Mao YH, Liu HL, Liu YL, Tao SH (2014) Deciphering the rules by which dynamics of mRNA secondary structure affect translation efficiency in *Saccharomyces cerevisiae*. *Nucleic Acids Res* 42(8):4813–4822. <https://doi.org/10.1007/s00253-017-8644-3>
- Mayen JFC, Lupan A, Cosar C, Kun AZ, Silaghi-Dumitrescu R (2015) On the roles of the alanine and serine in the beta-sheet structure of fibroin. *Biophys Chem* 197:10–17. <https://doi.org/10.1016/j.bpc.2014.11.001>
- Nguyen TKM, Ki MR, Son RG, Pack SP (2019) The NT11, a novel fusion tag for enhancing protein expression in *Escherichia coli*. *Appl Microbiol Biotechnol* 103(5):2205–2216. <https://doi.org/10.1007/s00253-018-09595-w>
- Pardon E, Laeremans T, Triest S, Rasmussen SGF, Wohlkoenig A, Ruf A, Muyldermans S, Hol WGJ, Kobilka BK, Steyaert J (2014) A general protocol for the generation of nanobodies for structural biology. *Nat Protoc* 9(3):674–693. <https://doi.org/10.1038/nprot.2014.039>
- Patterson DJ, Yasuhara K, Ruzzo WL (2002) Pre-mRNA secondary structure prediction aids splice site prediction. *Pac Symp Biocomput*:223–234. https://doi.org/10.1142/9789812799623_0021
- Salema V, Fernandez LA (2013) High yield purification of nanobodies from the periplasm of *E. coli* as fusions with the maltose binding protein. *Protein Expr Purif* 91(1):42–48. <https://doi.org/10.1016/j.pep.2013.07.001>
- Schoof M, Faust B, Saunders RA, Sangwan S, Rezelj V, Hoppe N, Boone M, Billesbolle CB, Puchades C, Azumaya CM, Kratochvil HT, Zimanyi M, Deshpande I, Liang JH, Dickinson S, Nguyen HC, Chio CM, Merz GE, Thompson MC, Diwanji D, Schaefer K, Anand AA, Dobzinski N, Zha BS, Simoneau CR, Leon K, White KM, Chio US, Gupta M, Jin ML, Li F, Liu YX, Zhang KH, Bulkley D, Sun M, Smith AM, Rizo AN, Moss F, Brilot AF, Pourmal S, Trenker R, Pospiech T, Gupta S, Barsi-Rhyne B, Belyy V, Barile-Hill AW, Nock S, Liu YW, Krogan NJ, Ralston CY, Swaney DL, Garcia-Sastre A, Ott M, Vignuzzi M, Walter P, Manglik A, Consortium QSB (2020) An ultrapotent synthetic nanobody neutralizes SARS-CoV-2 by stabilizing inactive Spike. *Science* 370(6523):1473–1479. <https://doi.org/10.1126/science.abe3255>
- Scully M, Cataland SR, Peyvandi F, Coppo P, Knobl P, Hovinga JAK, Metjian A, de la Rubia J, Pavenski K, Callewaert F, Biswas D, De Winter H, Zeldin RK, Investigators H (2019) Caplacizumab treatment for acquired thrombotic thrombocytopenic purpura. *N Engl J Med* 380(4):335–346. <https://doi.org/10.1056/Nejmoa1806311>
- Shimura K, Kikuchi A, Ohtomo K, Katagata Y, Hyodo A (1976) Studies on silk fibroin of *Bombyx mori*. I. Fractionation of fibroin prepared from posterior silk gland. *J Biochem* 80(4):693–702. <https://doi.org/10.1093/oxfordjournals.jbchem.a131328>
- Shpaer EG (1985) The secondary structure of mRNAs from *Escherichia coli*: its possible role in increasing the accuracy of translation. *Nucleic Acids Res* 13(1):275–288. <https://doi.org/10.1093/nar/13.1.275>
- Stepanenko OV, Marabotti A, Kuznetsova IM, Turoverov KK, Fini C, Varriale A, Staiano M, Rossi M, D'Auria S (2008) Hydrophobic interactions and ionic networks play an important role in thermal stability and denaturation mechanism of the porcine odorant-binding protein. *Proteins* 71(1):35–44. <https://doi.org/10.1002/prot.21658>
- Tanaka K, Inoue S, Mizuno S (1999) Hydrophobic interaction of P25, containing Asn-linked oligosaccharide chains, with the H-L complex of silk fibroin produced by *Bombyx mori*. *Insect Biochem Mol Biol* 29(3):269–276. [https://doi.org/10.1016/S0965-1748\(98\)00135-0](https://doi.org/10.1016/S0965-1748(98)00135-0)
- Veggiani G, Giabbai B, Semrau MS, Medagli B, Riccio V, Bajc G, Storici P, de Marco A (2020) Comparative analysis of fusion tags used to functionalize recombinant antibodies. *Protein Expr Purif* 166:105505. <https://doi.org/10.1016/J.Pep.2019.105505>
- Victor MP, Acharya D, Begum T, Ghosh TC (2019) The optimization of mRNA expression level by its intrinsic properties—insights from codon usage pattern and structural stability of mRNA. *Genomics* 111(6):1292–1297. <https://doi.org/10.1016/j.ygeno.2018.08.009>
- Wang MS, Du YB, Huang HM, Zhu ZL, Du SS, Chen SY, Zhao HP, Yan Z (2019) Silk fibroin peptide suppresses proliferation and induces apoptosis and cell cycle arrest in human lung cancer cells. *Acta Pharmacol Sin* 40(4):522–529. <https://doi.org/10.1038/s41401-018-0048-0>
- Wang Y, Fan Z, Shao L, Kong X, Hou X, Tian D, Sun Y, Xiao Y, Yu L (2016) Nanobody-derived nanobiotechnology tool kits for diverse biomedical and biotechnology applications. *Int J Nanomedicine* 11:3287–3302. <https://doi.org/10.2147/IJn.S107194>
- Wang YQ, Duan TB, Hong MH, Zhou YT, Huang H, Xiao X, Zheng J, Zhou H, Lu Z (2021) Quantitative proteomic analysis uncovers inhibition of melanin synthesis by silk fibroin via MITF/tyrosinase axis in B16 melanoma cells. *Life Sci* 284:119930. <https://doi.org/10.1016/J.Lfs.2021.119930>
- Xiang YF, Nambulli S, Xiao ZY, Liu H, Sang Z, Duprex WP, Schneidman-Duhovny D, Zhang C, Shi Y (2020) Versatile and multivalent nanobodies efficiently neutralize SARS-CoV-2. *Science* 370(6523):1479–1484. <https://doi.org/10.1126/science.abe4747>
- Zhou CZ, Confalonieri F, Jacquet M, Perasso R, Li ZG, Janin J (2001) Silk fibroin: structural implications of a remarkable amino acid sequence. *Proteins* 44(2):119–122. <https://doi.org/10.1002/Prot.1078>
- Zur H, Tuller T (2012) Strong association between mRNA folding strength and protein abundance in *S. cerevisiae*. *Embo Rep* 13(3):272–277. <https://doi.org/10.1038/embor.2011.262>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.