Construction of a synthetic *Araneus ventricosus* dragline silk gene multimer and its expression in *Escherichia coli*

Tingting Liu¹ · Anwen Liang¹ · Ziqiang Liang¹ · Guanghong Li¹ · Fanghai Wang¹

Received: 2 August 2017 / Accepted: 7 May 2018 / Published online: 11 May 2018
© Springer-Verlag GmbH Germany, part of Springer Nature 2018

Abstract

One of the most representative core gene sequence of *Araneus ventricosus* dragline silk protein partial cDNA monomer (JN857964.2) was selected and multimerized using a “head-to-tail” strategy by compatible but nonregenerable sites at both ends resulting in a concatemer of 16 contiguous monomers. This concatemer was cloned into pET-28a(+) expression vector and transformed into *Escherichia coli*. A 52.6 kDa silk protein was successfully expressed and detected by SDS–PAGE and confirmed by Western blotting. A maximum yield of the silk protein was expressed with 7.06 mM IPTG after 5 h incubation. This is the first report on the construction and overexpression of a *A. ventricosus* dragline silk multimeric gene construct and the results from our study will provide a reference point for further exploration and development of large-scale production of spider silk protein.

Keywords *Araneus ventricosus* · Spider dragline silk protein · Multimer · Expression

Introduction

Spider dragline silk is a protein fiber, which possess numerous excellent characteristics, such as, extreme lightness and softness, high strength and extensibility, good toughness and elasticity, strong shock resistance ability, resistance to low temperature, biodegradability, etc. Due to these unique properties, spider dragline silk has numerous applications. In military, it can be used to make bullet-proof vests, parachutes, armors in aircrafts and tanks. In medicine, it is useful in manufacturing artificial joints and limbs, as sutures for stitching wounds, as scaffolds for cell culture, and many biological tissue materials. Industrial applications of dragline silk include manufacturing of wheels, tyres and fishing nets. In architecture, it may be used as structural and composites for buildings and in textile manufacturing it can be used to make clothes, scarves, hats, etc. (Pan et al 2004; Zhao et al 2007).

Spiders spin a very small amount of silk, and the process of silk extraction from spiders is complex. Additionally,

spiders can not be reared in high-density farming due to their cannibalistic and aggressive nature and hence is difficult to obtain sufficient quantities of natural spider dragline silk to satisfy the demands for practical applications (Wang and Hu 2007). Therefore, research has focused on producing spider silk using alternate bioengineering technology. A spider dragline silk protein gene from *Nephila clavipes*, has been successfully cloned and expressed in *Escherichia coli* (Fahnestock and Irwin 1997; Xia et al. 2010; Qiao et al. 2014) and *Saccharomyces cerevisiae* (Fahnestock and Bedzyk 1997). Several other recombinant spider silk proteins have been produced from many species using genetic engineering strategies, including *Argiope aurantia*, *Nephilengys cruentata*, *Parawixia bistriata* and *Avicularia juruensis* (Tokareva et al. 2013).

We report in this paper for the first time cloning and overexpression of a spider silk protein from *Araneus ventricosus*. *A. ventricosus* is one of the common spiders in China, which has much larger webs and a better spider silk compared to other kind of spider silks.

✉ Fanghai Wang
lsswfh@mail.sysu.edu.cn

¹ State Key Laboratory for Biocontrol and Institute of Entomology, Sun Yat-sen University, Guangzhou 510275, China

Materials and methods

Selection of *A. ventricosus* dragline silk gene and preparation of DNA monomer

One of the most representative core sequence from *A. ventricosus* Masp1 partial cDNA (JN857964.2) (Kwang et al. 2012) was selected as a DNA monomer. Its nucleotide sequence is 108 bp: GCCGCGGCAGCCGAGCAGCA GCTGGTGGACAAGGAGGTCAAGGTGGATATGGA GGATTAGGTTCCCAAGGAGCTGGACAAGGAGGA TATGGAGCAGGACAAGGTGGT.

The following two synthetic oligonucleotides (Takara) were used to construct the artificial spider silk gene DNA monomer:

1. 72-base oligonucleotide, named as A strand,

5'-GCTAGCGCCGCGGCAGCCGAGCAGCAGCT GGTGGACAAGGAGGTCAAGGTGGATATGGAGGA TTAGGTTCC-3'.

2. 72-base oligonucleotide, named as B strand,

5'-AAGCTTACTAGTACCACCTTGTCTGCTCCATA TCCTCCTTGTCCAGCTCCTTGGGAACCTAATCCTCC ATA-3'.

At the 5' end of A strand was *Nhe* I restriction enzyme cutting site (GCTAGC), and at the 5' end of B strand were *Hind* III (AAGCTT) and *Spe* I (ACTAGT) restriction enzyme cutting sites, respectively, *Nhe* I and *Spe* I were isocandamers. A strand and B strands were designed to hybridize with each other in a 18-base region of complementarity at the 3' end of each strand.

After annealing for 30 s at 49 °C, the two oligonucleotide single chains were extended to form a double-stranded DNA through its complementary sticky ends using PrimeSTAR HS DNA polymerase (Takara). The product was the double chain DNA monomer.

Construction of a synthetic *A. ventricosus* dragline silk gene multimer

The monomer was phosphorylated and ligated into pSIMPLE-19 EcoRV/BAP (dephosphorylated) plasmid (Takara), then transformed into *E. coli* DH5 α competent cells (Tiangen). The white colonies with ampicillin resistance on the LB amp plates were tested to confirm the plasmids of the correct size inserts by PCR, and finally, three colonies were selected and sequenced by Beijing Genomics Institute (BGI).

The DNA monomer was doubled in size by manipulation of the restriction sites at the ends of the unit. Correct plasmids were separately digested with *Nhe* I-*Hind* III and *Spe* I-*Hind* III. The digests were electrophoresed on agarose gel and the appropriate DNA bands were extracted by MiniBEST Agarose Gel DNA Extraction Kit Ver.4.0 (Takara). The *Nhe* I-*Hind* III segment containing one copy of the DNA monomer was then ligated to the *Spe* I-*Hind* III segment which also contained one copy of linear plasmid, thus effectively regenerating a complete plasmid with ampicillin resistance while doubling the size of the monomer. The regenerated plasmid containing the doubled monomer was cloned into *E. coli* DH5 α and used as a template in the next cloning step. After four rounds of cloning, this strategy increased the size of the monomer to 16 unit repeats (Fig. 1).

Construction of pET-28a(+)-16 multimers dragline silk gene expression vector and transformation

The synthetic 16 multimers gene was released from the recombinant pSIMPLE-19 EcoRV/BAP vector by restriction digestion of 5' end with *Nhe* I and 3' end with *Hind* III. The purified insert was cloned into the frame of the expression vector pET-28a(+) at the *Nhe* I / *Hind* III sites. The ligation reaction system was as following: the synthetic 16 multimers gene 7 μ L, the pET-28a(+) digested with *Nhe* I and *Hind* III 1 μ L, 10 $\times$ T4 DNA ligase buffer 1 μ L, T4 DNA ligase 1 μ L. The above reaction system was placed at 16 °C for 8 h. Then the final complete plasmid pET-28a(+)-16 multimers dragline silk gene expression vector [pET-28a(+)-16u] was transformed into *E. coli* DH5 α competent cells. The colonies were screened for correct 16 multimers gene inserts by PCR, and the plasmid was extracted and confirmed by restriction enzyme.

Protein expression, SDS-PAGE electrophoresis and western blot analysis

For protein expression, the strong T7 promoter in the expression vector pET-28a(+) drives gene expression efficiently. The pET-28a(+) plasmid attached 6 $\times$ His tags to the N-terminal and the C-terminal of the protein, making detection of the expressed protein by western blotting very simple. The pET-28a(+)-16u plasmid from miniprep was then transformed into Rosetta(DE3) competent cells (Promega). Positive transformants on the LB plates with kanamycin were picked up and inoculated to 10 mL LB liquid medium with kanamycin. When they grew to an OD of 0.6 at 37 °C, added IPTG to the final concentration of 1 mM and induced for 5 h. After that, cells were harvested by centrifugation (10 min at 12,000 rpm), resuspended in lysis buffer (1.5% SDS, 1 mM PMSF, 1% TritonX-100, 1 mg/mL Lysozyme).

Fig. 1 Schematic map of cloning and multimerizing of the monomer encoding dragline silk protein of *Araneus ventricosus*

The suspension was lysed at room temperature for 1 h, frozen at -80°C , and then thawed at 37°C , and was repeated three times. The lysate was centrifuged and the supernatant was taken as samples for SDS-PAGE analysis.

SDS-PAGE gel electrophoresis was dyed using 0.25% coomassie bright blue R250 dye solution. The gels were scanned and the overexpressed protein quantities calculated

using Bandscan software. All data were analyzed using the SPSS Statistics package, version 22.0.

Following SDS-PAGE, protein bands were transferred electrophoretically to a sheet of PVDF membrane, and stained immunochemically (Sambrook et al. 1989) with mouse anti-His tag monoclonal antibody (UPSTATE, 1:10,000), followed by goat anti-mouse monoclonal antibody with HRP label (Protein-tech, 1:5000). The membrane was stained using a DAB developing kit in the dark for 5 min, washed, and photos taken.

Optimization of the expression condition

The rosetta (DE3) cells with pET-28a(+)-16u plasmids were cultured in a shaking incubator at 220 rpm, 37°C to an $\text{OD}_{660\text{nm}}$ of 0.6 following which 1 mM IPTG was added and the incubation continued for a further 6 h. During this period 1 mL samples was collected at 1 h intervals and the samples and the overexpressed proteins were determined as described above. Different IPTG concentrations (0.5–10 mM) were tested over a 5 h incubation period to determine the optimal concentration of IPTG for induction. Samples were analyzed on SDS-PAGE gels as described above scanned by Bandscan software to calculate the overexpressed protein quantity.

Results

Construction of monomeric and multimeric spider dragline silk gene

Plasmids containing correctly constructed tandem repeats of the synthetic DNA sequences were identified by double digestion with *Nhe* I-*Hind* III and *Spe* I-*Hind* III. Multimers were formed by insertion of *Nhe* I-*Hind* III fragments into *Spe* I-*Hind* III digested vectors. These enzymes generated identical cohesive ends, and when the ligation joined a *Nhe* I and *Spe* I site together in a “head to tail” fashion, the internal restriction sites were destroyed. This strategy was exploited to control construction of multimers ranging in size from 1 to 16 repeats of the monomer. These constructs are shown on agarose gel (Fig. 2).

Construction of pET-28a(+)-16 multimer dragline silk gene expression vector

The plasmid pET-28a(+)-16u extracted from the positive colony was digested with mono-restriction endonuclease enzyme *Nhe* I, then separated by agar gel electrophoresis. The 16 multimer dragline silk gene was 1834 bp, the pET-28a(+) plasmid was 5369 bp, so the pET-28a(+)-16u was 7203 bp. The results of agar gel electrophoresis

Fig. 2 Agarose gel analyses showing the synthetic spider silk DNA multimers at different doubling stages. Note: The sequential recombinant plasmids containing the different silk-like gene insert fragments were subjected to restriction digestion with *Nhe* I and *Hind* III to release the silk-like gene insert. The restriction digestion products were separated on 1% agarose gels. M: marker; $\times 1$ – $\times 16$: repetitive synthetic silk sequences after sequential insert doubling. The size of the silk inserts was 124, 238, 466, 922 and 1834 bp for $\times 1$, $\times 2$, $\times 4$, $\times 8$ and $\times 16X$, respectively. The black arrows show the linearized pSIMPLE 19 EcoRV/BAP plasmid and the red arrows show the silk-like gene insert

(Fig. 3), confirmed that we have obtained the right size of pET-28a(+)-16u.

Protein expression and analysis

Our expression vector pET-28a(+) placed the synthetic genes under the control of the T7 strong promoter and placed $6 \times \text{His}$ tags to the N-terminal and the C-terminal of the recombinant protein for detection by western blotting. pET-28a(+)-16u were transformed into *E. coli* Rosetta (DE3) and the encoded protein was expressed after induction by IPTG. Crude extracts were analyzed by polyacrylamide gel electrophoresis. Staining of the gels with Coomassie Brilliant Blue showed clear expression of the protein, which was 52.6 kDa (Fig. 4). Western blotting analysis using an anti-his tag antibody further confirmed the expression of the protein, because the band corresponding to the molecular weight of the expected protein was clearly observed (Fig. 5).

Optimization of the time and concentration of IPTG-inducible expression

The 16 multimer spider dragline silk protein was induced to be expressed by IPTG from 2 to 6 h, all samples are shown on the SDS–PAGE gel (Fig. 6). The gels were scanned by Bandsan software to calculate the expression percentage of the protein (Fig. 7). When the IPTG inducing time was 5 h, the percentage of the protein expression was the highest (Fig. 7). The data were dealt with SPSS Statistics

Fig. 3 The size of pET-28a(+)-16u is shown on the agarose gel. Note: lane M: marker; lane 1: the pET-28a(+)-16u plasmid digested by *Nhe* I. The black arrow shows the linearized pET-28a(+)-16u

22.0. The cubic models of 16 multimer spider dragline silk protein expression levels were obtained at different induction time course of IPTG, $y = 3.91 - 3.66x + 1.11x^2 - 0.1x^3$ ($R^2 = 0.779$). From the equation, it was confirmed that the optimized time for IPTG to induce the protein expression is 4.92 h (Fig. 8).

The 16 multimer spider dragline silk protein was induced to be expressed by different IPTG concentrations from 0.5 to 10 mM. All samples are shown on the SDS–PAGE gel (Fig. 9). The gels were scanned by Bandsan software to calculate the expression percentage of the protein (Fig. 10). When the concentration of IPTG induction was 6 mM, the percentage of the protein expression was the highest (Fig. 10). The data were dealt with SPSS Statistics 22.0. The cubic models of 16 multimer spider dragline silk protein expression levels were obtained at different induction

Fig. 4 The SDS-PAGE analysis of *Araneus ventricosus* dragline silk protein expressed by pET-28a(+)-16u. Note: lane M, protein marker; lane Con, the sample from pET-28a(+); lane 1, the sample from pET-28a(+)-16u. The black arrow shows the 16 multimer dragline silk protein

Fig. 5 Western blot analysis of *Araneus ventricosus* dragline silk protein expressed by pET-28a(+)-16u. Note: lane M, protein marker; lane Con, the sample from pET-28a(+); lane 1, the sample from pET-28a(+)-16u

Fig. 6 The SDS-PAGE analysis of the protein expression of pET-28a(+)-16u induced by IPTG at different time. Note: lane M, protein mark; lane Con, no IPTG induction; lane 1–5, the time of IPTG induction was 2, 3, 4, 5, and 6 h respectively. The black arrow shows the 16 multimer dragline silk protein

Fig. 7 Different expression levels of 16 multimer dragline silk protein at different induction time courses of IPTG

concentration of IPTG, $y = 5.83 - 2.27x + 0.69x^2 - 0.05x^3$ ($R^2 = 0.934$). From the equation, it was confirmed that the optimized concentration for IPTG to induce the protein expression is 7.06 mM (Fig. 11).

Discussion

Spider dragline silk gene is highly repetitive and rich in GC, thus it is difficult to clone the complete gene by PCR. In this study, we have indicated that application of compatible, nonregenerable restriction sites allows precise control of the manipulation of DNA repeat number. We have made the *A. ventricosus* dragline silk gene monomer repeat 16 times and its encoded protein has reached 52.6 kDa, and if we want the larger gene to encode the larger protein, we can continue

Fig. 8 The cubic model shows the different expression level of 16 multimer dragline silk protein at different induction time courses of IPTG

Fig. 9 The SDS-PAGE analysis of the protein expression of pET-28a(+)-16u induced by IPTG at different concentrations. Note: lane M, protein marker; lane Con, induction by 0 mM IPTG; lane 1–6, induction by 0.5, 1, 2, 4, 6, 10 mM IPTG respectively. The black arrow shows the 16 multimer dragline silk protein

using this doubling method. This strategy have opened the way for engineering proteins with different peptide domains in a controlled manner.

The *A. ventricosus* dragline silk partial cDNA which we selected as the monomer contains many rare codons, such as GGA, which encoded glycine, so we utilized Rosetta (DE3) as the expression host. Rosetta (DE3), which has rare codons, can enhance the expression of eukaryotic proteins. The strain has a plasmid with chloramphenicol resistance to supply the tRNAs of AUA, AGG, AGA, CUA, CCC and GGA, so it has achieved “universal” translation. The SDS-PAGE and western blotting confirmed the identity of the protein expressed by synthetic spider silk multimer. Some researchers found that the expression level decreased as the repeated times of the synthetic gene monomer increased,

Fig. 10 Different expression levels of 16 multimer dragline silk protein at different IPTG concentrations

Fig. 11 The cubic model shows different expression levels of 16 multimer dragline silk protein at different IPTG concentrations

apparently due to the gene rearrangement and the unusual secondary structure of mRNA (Du et al. 2011).

In the last part of this paper, we explored the optimal IPTG induction time and concentration for expression of the protein. We believe that it will help us to produce the spider dragline silk protein more efficiently and provide experience and reference for further exploration and development of large-scale production of spider silk protein.

In the past two decades, many researchers focused on dragline silk gene multimer construction of *N. clavipes* and its expression in *E. coli*. Although the 96 times concatemer has been successfully obtained by connecting synthetic spider fibroin gene monomer code in tandem using isocaudamer ligation method (Qiao et al. 2014), and 284.9 kDa recombinant protein of the spider *N. clavipes* was produced in

metabolically engineered *E. coli* and spun into a fiber displaying mechanical properties comparable to those of the native silk (Xia et al. 2010). The researcher also found that the yield of recombinant spider silk protein decreased with the increase of molecular weight (Fahnestock and Irwin 1997) and the recombinant proteins of lower molecular weight versions yielded inferior fiber properties (Xia et al. 2010). Here we first report about the spider *A. ventricosus* dragline silk gene multimer construction using a “head-to-tail” strategy and its successful expression in *E. coli*, however the mechanical properties of the silk protein are not involved in this report. We will investigate the mechanical properties and difference of the silk protein from biotechnology and spider in the future.

Acknowledgements This research was supported by the science and technology planning project of Guangdong Province (2014A010107009). Many thanks to Dr. Muhammad shakeel from College of Agriculture, South China Agricultural University, and Professor Bharat K. C. Patel from School of Earth, Environmental & Biological Sciences, Queensland University of Technology for correcting and improving the use of Grammar and the English language in the manuscript.

Compliance with ethical standards

Conflict of interest All authors declared that they have no conflict of interest.

References

- Du WH, Zhao TF, Zhu Y (2011) Advances in genetic engineering of spider silk proteins. *Sci Seric* 37:892–898
- Fahnestock SR, Bedzyk LA (1997) Production of synthetic spider dragline silk protein in *Pichia pastoris*. *Appl Microbiol Biotechnol* 47:33–39
- Fahnestock SR, Irwin SL (1997) Synthetic spider dragline silk proteins and their production in *Escherichia coli*. *Appl Microbiol Biotechnol* 47:23–32
- Kwang SL, Bo YK, Doh HK et al (2012) Molecular cloning and characterization of the partial major ampullate silk protein gene from the spider *Araneus ventricosus*. *J Asia Pac Entomol* 15:641–646
- Pan ZJ, Li CP, Sheng JY (2004) The research and application of high performance spider silk protein fiber(1). *Silk* 10:40–43
- Qiao X, Wang Y, Li JJ et al (2014) Construction of spider dragging silk protein MaSp1 prokaryotic expression vector and its expression and purification in *Escherichia coli*. *Mil Med Sci* 8:621–625
- Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor
- Tokareva O, Michalczechen-Lacerda VA, Rech EL et al (2013) Recombinant DNA production of spider silk proteins. *Microb Biotechnol* 6:651–663
- Wang LG, Hu P (2007) The research progress of spider silk and prospects. *Adv Mater Ind* 9:63–67
- Xia XX, Qian ZG, Ki CS et al (2010) Native-sized recombinant spider silk protein produced in metabolically engineered *Escherichia coli* results in a strong fiber. *Proc Natl Acad Sci USA* 107:14059–14063
- Zhao AC, Xia QY, Lu C et al (2007) Research trend in super fiber spider silks. *News Lett Seric Sci* 27:28–34