Journal of Structural Biology 170 (2010) 420–425

Contents lists available at ScienceDirect

Journal of Structural Biology

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

Characterization of recombinantly produced spider flagelliform silk domains

Markus Heim [a], Christian B. Ackerschott [b,1], Thomas Scheibel [a,*]

[a] University of Bayreuth, FAN, Lehrstuhl für Biomaterialien, 95440 Bayreuth, Germany
[b] Technische Universität München, Department Chemie, Lehrstuhl für Biotechnology, 85747 Garching, Germany

ARTICLE INFO

Article history:
Received 5 November 2009
Received in revised form 23 December 2009
Accepted 27 December 2009
Available online 4 January 2010

Keywords:
Engineered spider silk
Protein folding
Protein oligomerization
Spider silk protein assembly
Spider silk protein solubility
Disulphide formation

ABSTRACT

The capture spiral of a spider's orb web is made of flagelliform silk, providing high elasticity and an outstanding toughness, perfectly suited for trapping prey. Flagelliform silk comprises mainly one single protein (FLAG) with an estimated molecular weight of 360 kDa. We engineered constructs mimicking distinct domains of FLAG (eFLAG) and produced them recombinantly to analyze the structure–function relationship of FLAG domains and assembly properties of FLAG. While in solution the small carboxy-terminal domain is structured, domains from the repetitive core region adopt a conformation typical for intrinsically unstructured proteins. To investigate the influence of the respective domains on solubility and assembly, we tested the aggregation behaviour of individual domains and domain ensembles in presence of conditions known to trigger silk assembly. Both, the length of the repetitive core domain as well as the presence of the carboxy-terminal non-repetitive domain showed impact on eFLAG aggregation.

1. Introduction

Spider silks originate from one ancestral silk of the Mid-Devonian period (Zschokke, 2003). Within the last 400 million years several silk types have evolved for special tasks like wrapping or prey capturing (Vollrath, 2000). Most spider silks provide outstanding mechanical properties, such as high stiffness, strength, extensibility, and an enormous toughness, the later exceeding that of most other natural or synthetic materials including synthetic rubber, Kevlar, or high-tensile steel (Gosline et al., 1999; Hardy et al., 2008; Heim et al., 2009; Lammel et al., 2008; Römer and Scheibel, 2008).

Flagelliform silk serves as the core fiber of the capture spiral of an orb web and is mainly composed of one protein component. The protein sequence of *Nephila clavipes* flagelliform silk has been analyzed in detail, and the molecular weight (MW) of natural flagelliform silk protein (FLAG) was estimated after mRNA analysis (Hayashi and Lewis, 1998). Hence, mRNA and cDNA analysis indicated a primary structure of FLAG with a calculated molecular weight of 360 kDa (Hayashi and Lewis, 2001).

While the amino-terminal non-repetitive domain seems to be involved in cellular secretion (Motriuk-Smith et al., 2005; Rising et al., 2006), the carboxy-terminal non-repetitive domain is supposed to play an essential role in silk assembly (Sponner et al., 2004, 2005; Stark et al., 2007). Although the carboxy-terminal sequence of FLAG differs slightly from that of carboxy-terminal regions of other spider silks, it is expected to have similar properties and functions (Challis et al., 2006).

The repetitive core domain, covering ~90% of the sequence of native FLAG, can be best described by four amino acid consensus motifs which are repeated up to several 100 times (Hayashi and Lewis, 2001). These motifs are supposed to be involved in the formation of defined structures in the assembled silk thread: $(GGX)_n$ presumably folds into 3_{10} helices, while $(GPGGX)_n$ (with X usually being A, S, Y, or V in both cases) and $(GPGGAGGPY)_n$ motifs are supposed to form β-turn structures (Hayashi, 2002; Hayashi and Lewis, 2000). The fourth motif represents a spacer region with 28 amino acids including seven negatively charged side chains, indicating a possible role in solubilising FLAG. Besides in the minor ampullate spidroins (MIS) in orb weaving spiders, such spacer regions are only found in FLAG silk (no alignments in blast databases), separating the consensus motifs of the repetitive core into clusters (Fu et al., 2009). The spacer region is the only consensus motif within the repetitive core of FLAG that seems to be

Abbreviations: CD, circular dichroism; DTT_{ox}, trans-4,5-dihydroxy-1,2-dithiane; DTT_{red}, (2*S*,3*S*)-1,4-dimercaptobutane-2,3-diol; EDTA, ethylenediaminetetraacetic acid; eFLAG, engineered spider flagelliform silk; FLAG, (natural) spider flagelliform silk; HEPES, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid; IB, inclusion body; IMAC, immobilized metal ion affinity chromatography; IPTG, isopropyl β-D-1-thiogalactopyranoside; MAS, major ampullate spidroin; MIS, minor ampullate spidroin; MW, molecular weight; PAGE, polyacrylamide gel electrophoresis; SDS, sodium dodecyl sulfate; TCEP, tris(2-carboxyethyl)phosphine; Tris, tris(hydroxymethyl)aminomethane; UV, ultraviolet.

* Corresponding author. Fax: +49 (0) 921 557346.
E-mail address: Thomas.Scheibel@uni-bayreuth.de (T. Scheibel).
URL: http://www.fiberlab.de (T. Scheibel).
[1] Present address: Wyatt Technology Europe GmbH, 56307 Dernbach, Germany.

1047-8477/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.jsb.2009.12.025

intrinsically unstructured (Fu et al., 2009). Further, FLAG silk characteristically lacks polyalanine or poly(glycine–alanine) stretches that are known to form tight β-sheet crystallites in *Bombyx mori* fibroin $(GAGAGS)_n$ (Ha et al., 2005) and in orb weaving spiders' major ampullate spidroins (MAS) $(A_n/(GA)_n)$ (Holland et al., 2004). This renders FLAG less stiff and rigid than the two other well investigated silks with importance for its outstanding elasticity.

While *B. mori* fibroin and spider major ampullate silk are heterogenous fibers that are composed of three or two different protein species (Fu et al., 2009), respectively, FLAG silk is only comprising a single protein species. Therefore, FLAG silk seems to be a very suitable model system for studying the triggering of silk protein assembly and silk thread formation.

Since most information on the structure–function relationship of FLAG domains has been obtained from short model peptides, we employed a biotechnological strategy to produce recombinant proteins comprising distinct FLAG domains (Fig. 1) to gain better insights into FLAG's properties (Huemmerich et al., 2004; Vendrely et al., 2008).

The analysis of the secondary structure of our engineered constructs revealed that the carboxy-terminal domain adopts an α-helical conformation in solution. This conformation can be readily regained by refolding after denaturation. The folded carboxy-terminal domain is further able to form intermolecular disulphide bonds. These disulphides apparently have no influence on secondary or tertiary structure. In this study, effects of intermolecular disulphide formation, the size of the repetitive core domain, and the presence of ions on the stability and aggregation behaviour of certain engineered FLAG constructs have been investigated.

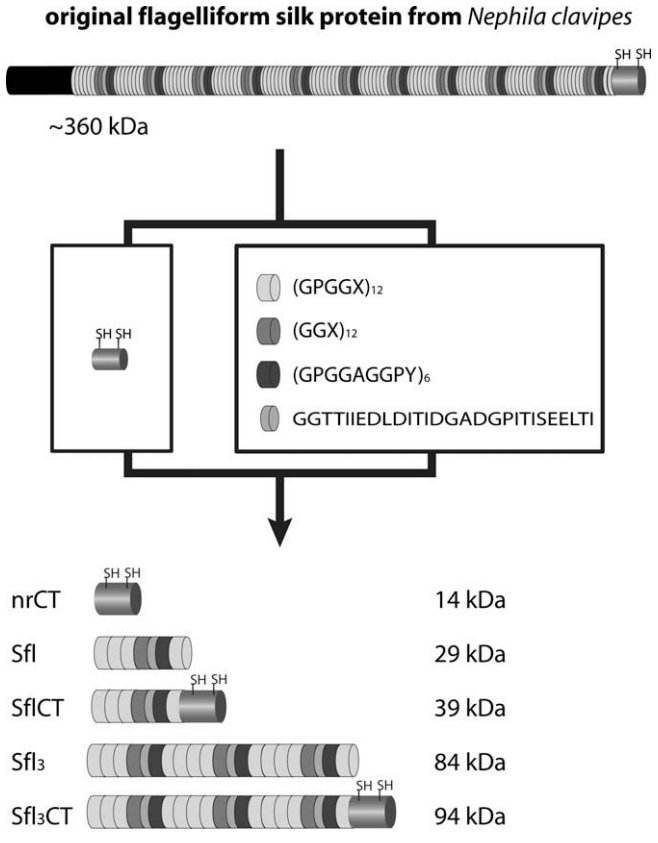

Fig. 1. Engineered recombinant flagelliform silk proteins (eFLAG). FLAG from *Nephila clavipes* consists of a large repetitive core domain and non-repetitive amino- and carboxy-terminal regions. The core domain comprises eleven repeats of an ensemble unit which contains repetitive sequences and is mimicked by the engineered construct Sfl. nrCT reflects the non-repetitive carboxy-terminal region, that contains the two sulfhydryl side chains. Subscript numbers indicate the number of Sfl-repeats.

2. Materials and methods

2.1. Cloning

Constructs coding for distinct domains of the FLAG protein of *N. clavipes* were engineered and optimized for expression in *Escherichia coli* as described previously (Huemmerich et al., 2004; Vendrely et al., 2008). Natural FLAG comprises a highly repetitive core domain with 11 ensemble repeats flanked by non-repetitive amino- and carboxy-terminal regions. In order to mimic its domain structure, the gene sequence of one of the ensemble repeats (named Sfl) was reconstituted using four different synthetic oligonucleotides. Using a previously established cloning technique, single ensemble repeat genes were multimerized (Vendrely et al., 2008). Since in bacteria production of proteins up to 150 kDa is highly efficient, we engineered proteins of up to three repeats (Sfl_3). The gene coding for the non-repetitive carboxy-terminal region, named nrCT, was amplified by PCR from a cDNA clone (kindly provided by Cheryl Hayashi, University of Riverside, CA, USA; GenBank Accession Number AF027973) and was optionally ligated with the sequences coding for Sfl and Sfl_3.

2.2. Protein production and purification

E. coli BLR (DE3) cells (Novagen) transformed with the expression plasmids were induced with 1 mM IPTG at an OD_{600} of 0.6 and harvested after 3 h of incubation at 30 °C. Cell pellets were resuspended in 20 mM HEPES, 5 mM NaCl, pH 7.5 (5 mL per 1 g of cells) at 4 °C. Upon addition of 0.2 mg lysozyme (Sigma–Aldrich, St. Louis, MO) per mL, the suspension was incubated at 4 °C until it was viscous. Protease inhibitor (Serva, Heidelberg, Germany) was added, and the suspension was sonicated five times (each cycle being 15 s) on ice using an HD/UW2200/KE76 ultrasonicator (Bandelin, Berlin, Germany). DNA was digested using 10 μg/mL DNase in 3 mM $MgCl_2$ for 30 min at room temperature (RT). While nrCT was produced solubly, all other Flag variants formed inclusion bodies (IBs).

For IB preparation, 0.5 vol of 60 mM EDTA, 2–3% Triton X-100 (v/v), 1.5 M NaCl, pH 7 was added, and the suspension was incubated for 30 min at 4 °C. IBs were sedimented at 6000*g* for 20 min at 6 °C and washed twice in 100 mM Tris–HCl, 20 mM EDTA, pH 7. IBs were dissolved in 20 mM HEPES, 5 mM NaCl, 8 M urea, pH 7.5, and the solution was then applied to an equilibrated Q-Sepharose Fast Flow column (20 mL, 1 mL/min; GE healthcare, Munich, Germany). The proteins were eluted using a linear gradient of NaCl. All variants of FLAG eluted at NaCl concentrations between 200 and 250 mM.

All pooled fractions of Sfl, Sfl_3, SflCT, and Sfl_3CT were precipitated using a final ammonium sulphate concentration of 1.2 M. After sedimentation, the protein pellet was dissolved in 20 mM HEPES, 5 mM NaCl, 8 M urea, pH 7.5 and dialysed against 10 mM ammonium hydrogen carbonate. Finally, Sfl, Sfl_3, SflCT, and Sfl_3CT (all with purities >98%) were lyophilized. Afterwards, the protein yields per liter of bacterial suspension culture (OD_{600} ~0.8) were calculated to be 10 mg/L for Sfl, 7 mg/L for SflCT, and 2.5–3 mg/L for Sfl_3 and Sfl_3CT.

nrCT was purified using a Ni-IMAC column (5 mL Fast Flow, 2 mL/min, GE healthcare) equilibrated with 20 mM HEPES, 100 mM NaCl, 20 mM imidazole, pH 7.5. The proteins were eluted with 20 mM HEPES, 100 mM NaCl, 500 mM imidazole, pH 7.5. Afterwards, nrCT was dialysed against 100 mM Tris–HCl pH 7. Additionally, some of the purified, native nrCT was precipitated,