Journal of Structural Biology 186 (2014) 431–437

Contents lists available at ScienceDirect

Journal of Structural Biology

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

Influence of repeat numbers on self-assembly rates of repetitive recombinant spider silk proteins

Martin Humenik, Michael Magdeburg, Thomas Scheibel *

Lehrstuhl Biomaterialien, Fakultät für Ingenieurwissenschaften, Universität Bayreuth, Universitätsstraße 30, 95440 Bayreuth, Germany

ARTICLE INFO

Article history:
Received 17 January 2014
Received in revised form 11 March 2014
Accepted 12 March 2014
Available online 20 March 2014

Keywords:
Self-assembly
Cross-β fibril
Kinetics
Polyalanine
Spider silk
Structure

ABSTRACT

Assembly of recombinant spider silk variants eADF4(C*n*) comprising different numbers (*n*) of the consensus sequence motif C, derived from the natural *Araneus diadematus* dragline silk ADF4, yielded indistinguishable nanofibrils in cases of $n \geqslant 2$. The C-module comprises 35 amino acids rich in glycine and proline residues (in GPGXY repeats) and one polyalanine stretch $(Ala)_8$. All variants were found to be intrinsically disordered in solution, and upon fibril formation they converted into a cross-β structure. Heterologous seeding indicated high structural compatibility between the different eADF4(C*n*) variants, however, their assembly kinetics differed in dependence of the number of repeats. Kinetic analysis revealed a nucleation-growth mechanism typical for the formation of cross-β-fibrils, with nucleation rates as well as growth rates increasing with increasing numbers of repeats. Strikingly, the single C-module did not self-assemble into fibrils, but upon addition of heterologous seeds fibril growth could be observed. Apparently, interconnecting of at least two C-modules significantly facilitates the structural transformation from a disordered state into β-sheet structures, which is necessary for nucleation and beneficial for fibril growth.

1. Introduction

Spider silk represents a fascinating material being mechanically strong and, at the same time, biocompatible (Humenik et al., 2011). One of the best understood spider silks, the so-called dragline of orb weaving spiders, exceeds any natural or artificial fiber in toughness (Gosline et al., 1999). Natural spider silks are, however, limited in availability based on both the cannibalistic behavior of spiders hampering farming and a complicated "silking" process to retrieve fibers. To circumvent the limitations in availability, different strategies have been established for biotechnological production of spider silk proteins (Heidebrecht and Scheibel, 2013; Tokareva et al., 2013). Recombinant spider silk proteins enabled development of silk materials like fibers (Xia et al., 2010), particles (Blüm and Scheibel, 2012; Numata et al., 2011), capsules (Blüm et al., 2014; Rabotyagova et al., 2010) and nanofibrils (Humenik and Scheibel, 2014; Numata and Kaplan, 2011; Rammensee et al., 2006; Slotta et al., 2007).

Spider silk proteins (spidroins) typically have high molecular weights (250–350 kDa), with primary sequences dominated by a repetitive core (>1000 amino acids) flanked by small globular amino- and carboxyterminal domains (100–150 amino acids) (Fig. 1). Consensus motifs of the dragline core repeats are composed of polyalanine stretches $(Ala)_n$ ($n = 4–12$) flanked by GA, GGX and GPGXY sequences (Rising et al., 2005). Interplay between the terminal domains and folding of the core domain results in hierarchical fiber assembling during an explicit spinning process (Eisoldt et al., 2011), including ionic exchange of chaotropic Na^+ and Cl^- against kosmotropic K^+ and PO_4^{3-}, a concomitant pH drop and water removal (Vollrath and Knight, 2001). In combination with shear forces, the spinning conditions result in the formation of antiparallel inter- and intrachain β-sheets being preferentially aligned in parallel to the fiber axis (Papadopoulos et al., 2009; van Beek et al., 2002).

Since the core domain of spider silk proteins lacks a defined tertiary structure in solution (Lefevre et al., 2011) they can be referred to as intrinsically disordered proteins (IDPs) (Uversky and Dunker, 2010). Intrinsic disorder of entire or partial protein sequences has been recognized as an important functional feature for many eukaryotic proteins (Chouard, 2011). Additionally, IDPs are better suited to convert their soluble structure into one of insoluble cross-β fibrils (Jahn et al., 2010) in comparison to tightly folded

* Corresponding author. Address: Bayreuther Zentrum für Kolloide und Grenzflächen (BZKG), Institut für Bio-Makromoleküle (bio-mac), Bayreuther Zentrum für Molekulare Biowissenschaften (BZMB), Bayreuther Materialzentrum (BayMAT), Universität Bayreuth, Universitätsstraße 30, 95440 Bayreuth, Germany.
E-mail address: thomas.scheibel@bm.uni-bayreuth.de (T. Scheibel).
URL: http://www.fiberlab.de (T. Scheibel).

http://dx.doi.org/10.1016/j.jsb.2014.03.010
1047-8477/

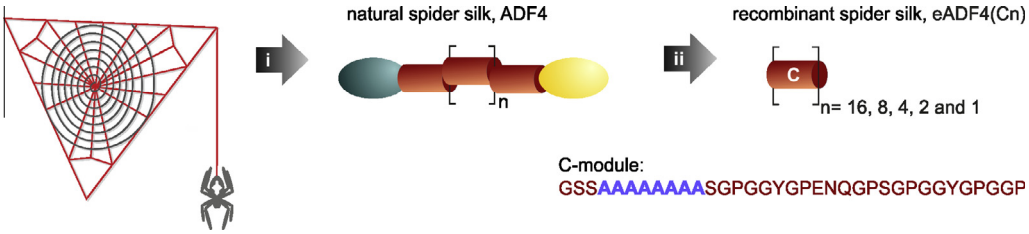

Fig.1. Composition of natural and recombinant spider silk proteins; (i) Dragline silk of the European garden spider (*A. diadematus)* comprises at least two major components ADF3 and ADF4, both showing a repetitive core domain (red) and small non-repetitive termini (green and yellow). (ii) The repeated consensus motif of the core domain of ADF4 (C-module) has been optimized for the production in *E. coli* (Huemmerich et al., 2004), and variants with different numbers of C-modules, eADF4(C*n*), were generated.

globular proteins due to higher conformational flexibility (Uversky, 2013).

The engineered spider silk protein eADF4(C16), comprising 16 times the consensus motif (C-module) of the core domain of dragline protein *Araneus diadematus* fibroin 4 (ADF4), self-assembles from disordered structure in solution into β-sheet rich nanofibrils exhibiting X-ray diffraction patterns typical for cross-β structures (Slotta et al., 2007). This self-assembly is triggered by cosmotropic phosphate ions similar to that of native spider silk. The C-module comprises 35 amino acids (Fig. 1) with one $(Ala)_8$ stretch forming β-sheets as well as glycine/proline rich GPGXY repeats remaining unstructured or helical upon assembly (Lefevre et al., 2011; Spiess et al., 2010).

Here, variants eADF4(C*n*) (Fig. 1) differing in number (n = 1, 2, 4, 8 and 16) of C-modules were analyzed concerning their assembly kinetics in order to determine the critical number of repeats required for nucleation and elongation of nanofibrils.

2. Material and methods

All chemicals used in this study had analytical grade purity. Sinapinic acid for MALDI TOF was obtained from Aldrich (Taufkirchen, Germany). All other chemicals were acquired form Carl Roth GmbH & Co. KG (Karlsruhe, Germany). Double-distilled water was prepared using a Millipore system (Merck Millipore, Merck KGaA, Darmstadt, Germany).

2.1. Protein production and purification

Proteins eADF4(C8) and eADF4(C16) were produced and purified as described previously (Huemmerich et al., 2004).

Variants eADF4(C4), (C2) and (C1) were produced and purified in fusion with a His_6-SUMO tag (Suhre and Scheibel, 2014). The cloning vectors pCS eADF4(C4), (C2) and (C1) were digested using BamHI and HindIII (New England BioLabs (NEB)) and designated inserts were ligated with a correspondingly digested pET28a(+)-vector (Novagen) encoding the His_6-SUMO tag using T4 Ligase (NEB). Constructs pET His_6-SUMO-eADF4(C*n*), n = 1, 2 and 4, were transformed into *E. coli* BL21(DE3)-GOLD (Novagen). The respective *E. coli* strain was fermented in LB Medium containing 35 μg/ml kanamycin and 0.001% (v/v) Breox FMT 30 antifoam (Cognis) using a 1.3 L Minifors reactor (Infors HT, Bottmingen, Switzerland). Gene expression was induced at an optical density OD_{600} = 45 using 1 mM IPTG, and fermentation continued at 30 °C for 4 h.

Cells pellets (100 g) were disrupted using a high pressure homogenisator Microfluidizer M110S (Microfluidics, USA-Newton), and 12.5 mg Protease-Inhibitor Mix-HP (SERVA, Heidelberg, Germany) was added to the suspension. Column chromatography was performed using an Äkta Purifier (GE Healthcare, Germany). After centrifugation, the supernatant was incubated with 20 mM imidazole and loaded onto a Ni–NTA chelating Sepharose column (50 mL, GE Healthcare) followed by washing with 50 mM Tris/HCl, 100 mM NaCl, 10 mM imidazole, pH 8 (buffer A) using three column volumes. The His_6-SUMO fusion proteins were eluted using 60% buffer B (50 mM Tris/HCl, 100 mM NaCl, 500 mM imidazole). The eluate was diluted six times using ddH_2O and the dilutant was loaded onto a Q-Sepharose column (150 mL, GE Healthcare), washed with 25 mM HEPES-Na, pH 8, and eluted using a gradient of 0–100% buffer B (25 mM HEPES-Na, 1 M NaCl) over five column volumes. Fusion proteins typically eluted at 150 mM NaCl, whereas remnants of bacterial DNA eluted at 700 mM NaCl. Pooled fractions were incubated subsequently with recombinant His_6-tagged *Saccharomyces cerevisiae* Ulp1 protease (1/1000 w/w) at 4 °C for 10 h. His_6-SUMO free eADF4(C*n*) was loaded onto a Ni–NTA column and collected in the flow through. eADF4(C*n*) was intensively dialyzed against 25 mM NH_4HCO_3 and lyophilized. The identity of the respective protein was confirmed by MALDI-TOF mass spectrometry analysis (Fig. S1; Table S1).

2.2. Protein characterization

2.2.1. FT-IR and CD spectroscopy

Soluble and assembled silk proteins were used at 1 mg/mL for FT-IR and at 0.2 mg/mL for CD-spectroscopy. FT-IR absorption spectra were recorded upon accumulation of 128 scans from 900 to 4000 cm^{-1} on a Bruker Tensor 27 (Bruker, Germany) spectrometer equipped with an AquaSpec™ Flow Cell (micro-biolytics GmbH, Germany), and spectral transformations were performed using OPUS software (version 6.5, Bruker Optik, GmbH). CD spectra were recorded upon accumulating five scans using a J-815 CD spectrometer (Jasco, Germany).

2.2.2. MALDI-TOF mass spectrometry

MALDI-TOF mass spectrometry was performed on a Bruker Reflex III (Bruker, Germany) equipped with a 337 nm N_2 laser in a linear mode and 20 kV acceleration voltages. Data were processed using mMass 5.1.0 software (Strohalm et al., 2008). Proteins, dialyzed in 25 mM NH_4HCO_3, were mixed in 1/1 ratio with a matrix solution (20 mg/ml sinapinic acid in 90% acetonitrile and 0.2% trifluoroacetic acid) before spotting onto a target plate.

2.3. Characterization of fibrils

eADF4(C*n*) was typically dissolved in 6 M GdmSCN, extensively dialyzed against 10 mM Tris/HCl, pH 8, and then centrifuged in a Beckman Optima ultracentrifuge at 185,000*g* for 1 h at 4 °C. Fresh protein solutions were prepared to exclude presence of oligomers before assembly. Protein self-assembly was initiated by addition of 100–200 mM potassium phosphate buffer (KPi), pH 8.

2.3.1. Congo Red binding experiments

Congo Red (CR) (5 μM) was mixed either with soluble or assembled proteins at 0.5 mg/mL in 10 mM Tris/HCl. Baselines of fibril suspensions were recorded at the same protein concentration, to