Published in final edited form as:

Biomol NMR Assign. 2011 October ; 5(2): 131–133. doi:10.1007/s12104-010-9284-z.

NMR assignments of the N-terminal domain of *Nephila clavipes* spidroin 1

Stuart Parnham,

Department of Biochemistry and Molecular Biology, Medical University of South Carolina, 173 Ashley Ave., BSB 535D, PO Box 250509, Charleston, SC 29425, USA

William A. Gaines,

Department of Genetics and Biochemistry, Clemson University, Clemson, SC 29634, USA

Brendan M. Duggan,

Department of Biochemistry and Molecular Biology, Medical University of South Carolina, 173 Ashley Ave., BSB 535D, PO Box 250509, Charleston, SC 29425, USA

William R. Marcotte Jr, and

Department of Genetics and Biochemistry, Clemson University, Clemson, SC 29634, USA

Mirko Hennig

Department of Biochemistry and Molecular Biology, Medical University of South Carolina, 173 Ashley Ave., BSB 535D, PO Box 250509, Charleston, SC 29425, USA

Mirko Hennig: hennig@musc.edu

Abstract

The building blocks of spider dragline silk are two fibrous proteins secreted from the major ampullate gland named spidroins 1 and 2 (MaSp1, MaSp2). These proteins consist of a large central domain composed of approximately 100 tandem copies of a 35–40 amino acid repeat sequence. Non-repetitive N and C-terminal domains, of which the C-terminal domain has been implicated to transition from soluble and insoluble states during spinning, flank the repetitive core. The N-terminal domain until recently has been largely unknown due to difficulties in cloning and expression. Here, we report nearly complete assignment for all ^1H , ^{13}C , and ^{15}N resonances in the 14 kDa N-terminal domain of major ampullate spidroin 1 (MaSp1-N) of the golden orb-web spider *Nephila clavipes*.

Keywords

Spider dragline silk; Major ampullate spidroin; N-terminal domain; *Nephila clavipes*

Biological context

Orb-weaving spiders (Araneidae) such as *Nephila clavipes* can produce a variety of high performance structural fibers that originate from different abdominal glands (Lewis 2006). These fibers have mechanical properties that are unparalleled in the environment and comparable to the best synthetic fibers produced by modern technology. Due to the superior intrinsic properties of spider silk, there is considerable interest in the production of protein-based structural polymers through genetic engineering for a wide range of different commercial and biomedical applications.

Spider dragline silk displays a unique combination of strength and extensibility. It is used not only to construct the outer frame and radii of the orb-shaped web but also as a hanging lifeline that allows the spider to evade and/or escape from predators. Molecular studies have revealed that the different fiber types consist of distinct structural proteins called spidroins. Analysis of the amino acid sequences of dragline silk identified its core constituents to be two fibrous proteins produced in the major ampullate gland that are called major ampullate spidroins 1 and 2 (MaSp1 and MaSp2) (Hinman and Lewis 1992; Xu and Lewis 1990). One characteristic feature of spider silk proteins is the presence of a large, highly repetitive central domain [~ 100 copies of an ~ 30 aa (MaSp1) or ~ 40 aa (MaSp2) repeat] within the interior of the silk protein. MaSp1 and MaSp2 contain short non-repetitive N- and C-terminal domains (~ 150 aa and ~ 100 aa, respectively) that flank the long, repetitive core sequences. All three domains (N-, repeat and C-) are present in dragline fibers indicating that all contribute in one way or another to the impressive tensile characteristics of the final product. Increasing our understanding of the diverse molecular modules found in silk proteins, along with a particular attention to their role in fiber assembly, will provide a more efficient pathway for using recombinant versions of these proteins in the field of biomimetics. Here we report ^1H , ^{13}C , and ^{15}N assignments as well as the positions of elements of regular secondary structure of the 14 kDa N-terminal domain of MaSp1 of the golden orb-web spider *Nephila clavipes*, which are prerequisites for the determination of its solution structure and to define MaSp1-N's mechanistic role in fiber formation.

Methods and experiments

Based on our recently isolated sequences (Gaines and Marcotte 2008), we have successfully cloned and over-expressed the N-terminal domain of *N. clavipes* MaSp1 (Nc.MaSp1-N) as a cleavable GST fusion in *E. coli*. The Nc.MaSp1-N protein was purified by glutathione-Sepharose affinity and the GST moiety cleaved by treatment with PreScission Protease® leaving nine non-native amino acids at the N-terminus. For uniform ^{15}N , and ^{15}N , ^{13}C -labeling, recombinant Nc.MaSp1-N protein was expressed in M9 minimal medium supplemented with $^{15}\text{NH}_4\text{Cl}$ with and without ^{13}C -labeled glucose, respectively.

The purified protein samples were concentrated and exchanged into the final NMR buffer containing 10 mM sodium phosphate, pH 7.0, 500 mM NaCl, and 3 mM NaN_3 in 500 μl of 90% H_2O /10% D_2O . NMR samples containing 0.5–1 mM protein were centrifuged and the supernatant passed through a 0.2 μm filter to ensure no insoluble material was in the final sample. For measurements in 100% D_2O , the protein was lyophilized and redissolved in D_2O .

NMR experiments were recorded on Bruker Avance III 600 MHz and Avance II 800 MHz spectrometers. The Avance 800 is equipped with a 5 mm ^1H [$^{13}\text{C}/^{15}\text{N}$] triple resonance probe featuring cryogenically cooled preamplifiers and r.f. coils on ^1H and ^{13}C channels and z-axis gradient while the Avance 600 features a conventional 5 mm ^1H [$^{13}\text{C}/^{15}\text{N}$] triple resonance probe. All NMR experiments were carried out at 298 K. The NMR data were processed using nmrPipe (Delaglio et al. 1995), and analyzed with CcpNMR Analysis software (Vranken et al. 2005).

Nearly complete sequential backbone assignments of the 135 aa Nc.MaSp1-N protein were achieved using through-bond 3D HNCACB, CBCA(CO)NH, HNCA, HN(CO)CA, and HNCO experiments. HBHA(CO)NH, $\text{H}_{\text{ali}}(\text{CCO})\text{NH}$ and $\text{C}_{\text{ali}}(\text{CO})\text{NH}$ in combination with HCCH-based experiments were used to unambiguously assign side chain spin systems (Sattler et al. 1999). A 3D ^{13}C - and/or ^{15}N -resolved NOESY approach confirmed and completed unambiguous ^1H - and ^{13}C -resonance assignments. Resonance assignments of the aromatic protons of the Nc.MaSp1-N protein were obtained using a combination of 2D

H(CDCG)CB transferring $^1\text{H}\delta$ coherence to $^{13}\text{C}\beta$ in an out-and-back manner (Yamazaki et al. 1993) and homonuclear 2D NOESY experiments. Using these experiments, nearly complete assignments for the aromatic carbons and protons were achieved.

Assignments and data deposition

High-quality NMR data has been obtained for Nc.MaSp1-N, as illustrated by the 2D ^1H , ^{15}N HSQC spectrum shown in Fig. 1. Analysis of the backbone chemical shifts of MaSp1-N, using the program TALOS+ (Shen et al. 2009) and the chemical shift index (Wishart and Sykes 1994) identifies six helical regions ($\alpha 1$: residues 17–32, $\alpha 2$: residues 40–48, $\alpha 3$: residues 51–62, $\alpha 4$: residues 71–86, $\alpha 5$: residues 95–114, and $\alpha 6$: residues 120–135) within the N-terminal domain as shown in Fig. 2. Assignments were obtained for all 135 residues, including 100% of the exchangeable amide $^1\text{H}^{\text{N}}$ protons, 99% of all amide ^{15}N nitrogens, 96% of side chain $^1\text{H}^{\text{C}}$ and 87% of all side chain ^{13}C carbons. ^1H chemical shifts were externally referenced to DSS, with heteronuclear ^{13}C and ^{15}N chemical shifts referenced indirectly according to the $^1\text{X}/^1\text{H}$ ratio in DSS. The chemical shifts have been deposited in the BioMagResBank (<http://www.bmrb.wisc.edu>) under the accession number 17148.

Acknowledgments

We acknowledge the support of the Hollings Marine Laboratory NMR Facility for this work. This work was supported by a grant from the National Institutes of Health to WRM (R15 EB007403).

References

- Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR*. 1995; 6:277–293. [PubMed: 8520220]
- Gaines WAT, Marcotte WR Jr. Identification and characterization of multiple spidroin 1 genes encoding major ampullate silk proteins in *Nephila clavipes*. *Insect Mol Biol*. 2008; 17:465–474. [PubMed: 18828837]
- Hinman MB, Lewis RV. Isolation of a clone encoding a second dragline silk fibroin. *Nephila clavipes* dragline silk is a two-protein fiber. *J Biol Chem*. 1992; 267:19320–19324. [PubMed: 1527052]
- Lewis RV. Spider silk: ancient ideas for new biomaterials. *Chem Rev*. 2006; 106:3762–3774. [PubMed: 16967919]
- Sattler M, Schleucher J, Griesinger C. Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog Nucl Magn Reson Spectrosc*. 1999; 34:93–158.
- Shen Y, Delaglio F, Cornilescu G, Bax A. TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J Biomol NMR*. 2009; 44:213–223. [PubMed: 19548092]
- Vranken WF, Boucher W, Stevens TJ, Fogh RH, Pajon A, Llinas M, Ulrich EL, Markley JL, Ionides J, Laue ED. The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins*. 2005; 59:687–696. [PubMed: 15815974]
- Wishart DS, Sykes BD. The ^{13}C chemical-shift index: a simple method for the identification of protein secondary structure using ^{13}C chemical-shift data. *J Biomol NMR*. 1994; 4:171–180. [PubMed: 8019132]
- Xu M, Lewis RV. Structure of a protein superfiber: spider dragline silk. *Proc Natl Acad Sci USA*. 1990; 87:7120–7124. [PubMed: 2402494]
- Yamazaki T, Forman-Kay JD, Kay LE. Two-dimensional NMR experiments for correlating $^{13}\text{C}\beta$ and $^1\text{H}\delta/\epsilon$ chemical shifts of aromatic residues in ^{13}C -labeled proteins via scalar couplings. *J Am Chem Soc*. 1993; 115:11054–11055.

Fig. 1. Quality of the NMR spectra obtained for Nc.MaSp1-N. Panel A shows an 800 MHz ^1H , ^{15}N HSQC spectrum recorded at 25°C. The assignments of the signals from backbone amide groups are indicated by residue type and number

Fig. 2.

Summary of the NMR evidence used to derive the secondary structure of MaSp1-N. Locations of the α -helices are marked by $\alpha 1$ T17-A32, $\alpha 2$ A40-T48, $\alpha 3$ D51-R62, $\alpha 4$ L71-A86, $\alpha 5$ V95-T114 and $\alpha 6$ Q120-S135. Experimentally determined hydrogen bond donors are represented as *dots* above the helices and were found to be limited to helical regions