J Biomol NMR (2012) 54:415–420
DOI 10.1007/s10858-012-9679-5

NMR STRUCTURE NOTE

NMR structure note: repetitive domain of aciniform spidroin 1 from *Nephila antipodiana*

Shujing Wang · Weidong Huang · Daiwen Yang

Received: 19 September 2012 / Accepted: 25 October 2012 / Published online: 6 November 2012

Biological context

Orb weaving spiders can generate up to 7 types of spider silks with their silk toolkits to realize specific tasks, such as locomotion, prey capture, egg protection and prey wrapping (Lewis 2006). Aciniform (Ac) silk, spun with aciniform spidroin (AcSp1), participates in prey wrapping, inner egg case construction and web decoration (Hayashi et al. 2004; Lewis 2006; Vasanthavada et al. 2007). It is reported of the highest toughness among the 7 spider silks and with mechanical properties superior to most synthetic materials including Kevlar, Nylon and high tensile steel (Hayashi et al. 2004).

Spidroins are large silk proteins (250–350 kDa) with two relatively conserved terminal domains and a core characteristic repetitive region that amounts to ~90 % of the full length sequence and is believed to dominate the mechanical properties of spider silk (Gatesy et al. 2001; Lewis 2006). Spidroins are secreted into silk glands and then stably stored there with a concentration of up to 50 % (w/w) before orderly self-assembled into silk fibers (Hijirida et al. 1996). The first partial AcSp1 sequence comprising 14 identical repetitive (RP) domains was obtained from banded garden spider, *Argiope trifasciata* (Hayashi et al. 2004). This AcSp1-RP domain (~200 amino acids (a.a.)) was speculated to comprise mainly α-helices and random coils based on the chemical shift index (CSI) plot (Xu et al. 2011). A partial AcSp1-like sequence from black widow spider, *Latrodectus Hesperus*, was also reported, which contains a homogeneous RP domain (~187 a.a.) sharing 45 % sequence identity with the AcSp1-RP domain (Vasanthavada et al. 2007). The AcSp1-RP domain is distinguished from the repetitive domains of other spidroins, since it does not contain multiple short repetitive motifs like $(GA)_n$, $(GGX)_n$, $(GPGGX)_n$ and $(A)_n$ (Lewis 2006).

To understand the structure–function relationship of silk proteins, great efforts have been put on the terminal domains (Askarieh et al. 2010; Hagn et al. 2010; He et al. 2012), but relatively less information is revealed on the tertiary structure of RP domains. So far only the structures of RP domains from an egg case silk protein (tubliform spidroin, TuSp1) have been solved by us (Lin et al. 2009). Previous studies have shown that the minimum sequence requirement for a silk protein fragment to form silk fibers is that the fragment should contain a RP region and a terminal domain (Hagn et al. 2010; Lin et al. 2009; Stark et al. 2007). It has also been demonstrated that the C-terminal domain (CTD) plays important roles in stabilizing the solution state of silk proteins and in forming well defined silk fibers and the N-terminal domain (NTD) regulates the assembly of silk proteins in a pH-dependent manner (Askarieh et al. 2010). However, the exact roles of the RP domains in silk fiber formation and silk protein storage at a high concentration without premature aggregation are still not clear.

Aciniform Spidroin 1 was reported to contain ~50 % α-helices and ~50 % disordered regions in the liquid state in silk glands, and contain ~24 % α-helices and 20–30 % moderately oriented β-sheets in solid silk fibers (Lefevre et al. 2011). The combination of α-helices, β-crystalline and disordered (amorphous) structure in Ac silk is believed

S. Wang · W. Huang · D. Yang (✉)
Department of Biological Sciences, National University of Singapore, 14 Science Drive 4, Singapore 117543, Singapore
e-mail: dbsydw@nus.edu.sg

Present Address:
W. Huang
Ningxia Medical University, Yinchuan 750004, Ningxia, People's Republic of China

to correlate with its great toughness based on the comparison with other six types of spider silks (Lefevre et al. 2011). Nevertheless, current findings of spidroins, which are mainly from major ampullate spidroin (MaSp) and TuSp, are quite deficient in explaining the property of AcSp1. Therefore, to understand the molecular mechanisms of protein storage at a high concentration and structural transition during self-assembly of AcSp1 molecules, it is necessary to solve the solution structure of AcSp1 domains.

Methods and results

Sequence identification

One clone D110 with ~800 base pairs, was identified as the C-terminal region of AcSp1 from our home built expressed sequence tag (EST) database of *N. antipodiana* (Huang et al. 2006) by online blasting. On the basis of the sequence of D110, a DNA fragment corresponding to more than one RP domain but less than two was identified using genomic DNA as template by PCR. The sequence of one complete RP domain was found to contain 202 a.a., which is divided into two regions: RP_{tr} (the truncated RP used here for structure determination) with 160 a.a. and RP_{ln} (a linker region connecting adjacent RP_{tr}s) that was predicted to be disordered (Fig. 1a). Pair-wise alignment shows that the repetitive sequence has an identity of 16–22 % with the two previously reported AcSp1 and AcSp1-like sequences.

Protein expression and purification

The RP_{tr} (160 a.a.) and full length RP (202 a.a.) were subcloned to a modified pET-32a (Novagen) vector respectively, with an uncleavable His-Tag (MHHHHHH) at the N-terminus of the target protein. The sequences were confirmed by DNA sequencing. The constructs were transformed into BL21(DE3) *E. coli* cells for expression. The unlabeled proteins were cultured in LB medium, while the ^{13}C, ^{15}N-labeled RP_{tr} was cultured in M9 medium supplied with 1 g/L $^{15}NH_4Cl$ and 2 g/L ^{13}C-D-glucose. The proteins were purified with a Ni–NTA affinity column (Qiagen) and then further with a gel filtration column using

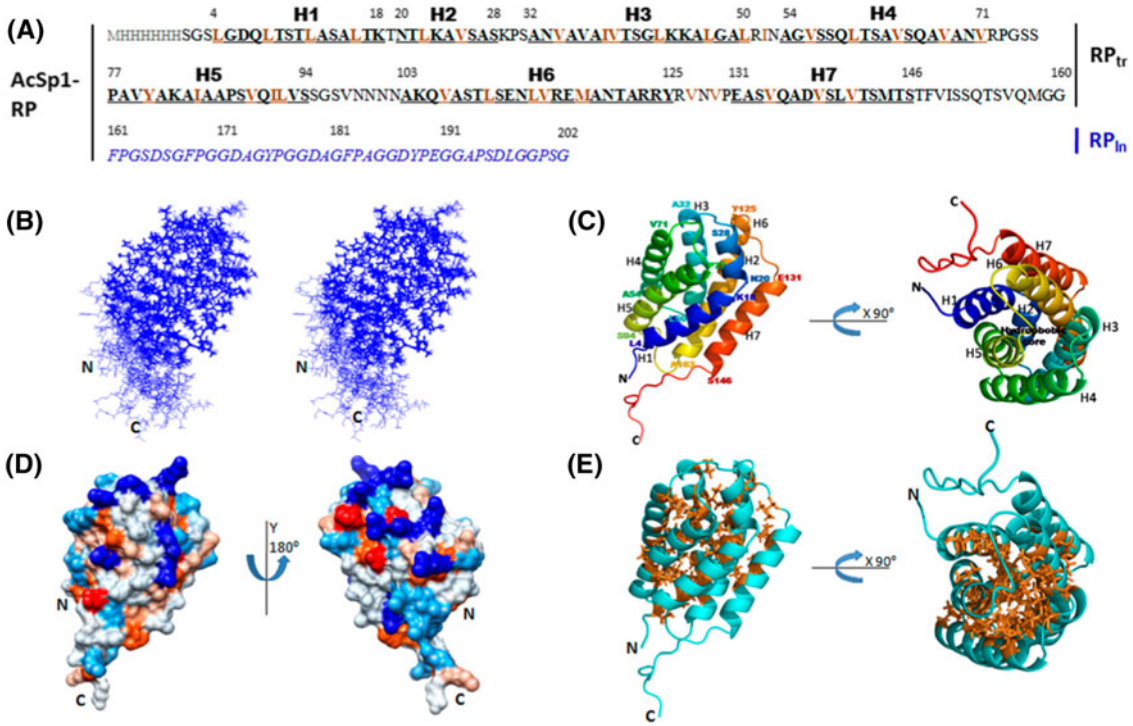

Fig. 1 Amino acid sequence of AcSp1-RP and solution structure of AcSp1-RP_{tr}. **a** Sequences of AcSp1-RP_{tr} and AcSp1-RP_{ln}. The 7 helices are indicated by H1-H7. The starting and ending residue number of each helix is labeled, and the residues participating in the hydrophobic core in **e** are colored tan, including 11 Leu, 15 Val, 4 Ile, 1 Met and 1 Tyr. **b** Stereo view of 10 structures of RP_{tr} with lowest target function values in a line mode including all atoms. **c** Representative ribbon structure with rainbow color viewed from one side (*left*) and the top (*right*) of the helix barrel. **d** Stereo view of hydrophobicity surface plot: hydrophobic (*tan*), less hydrophobic (*light tan*), neutral (*white*), hydrophilic (*less sky blue*), positively charged (*dark blue*) and negatively charged (*red*). **e** Stick view of side chains in the hydrophobic core (*tan*), which correspond to the residues highlighted in tan in **a**. The hydroxyl group of Tyr80 located in the hydrophobic core is outside the core

FPLC (GE Healthcare) in a buffer with 20 mM Tris–HCl, 100 mM NaCl and 1 mM EDTA at pH 7.0. Protein samples were then dialyzed against a buffer with 10 mM sodium phosphate and 1 mM EDTA at pH 6.0. The protein purity (>95 %) was determined by SDS-PAGE.

NMR spectroscopy and NMR resonance assignment

NMR Spectra were acquired on a Bruker 800 MHz NMR spectrometer equipped with a cryogenic probe at 25 °C. The NMR sample contained ~2 mM protein in a buffer with 10 mM sodium phosphate, 1 mM EDTA, 5 % D_2O and 0.01 % NaN_3 at pH 6.0. Besides 2D ^{1}H-^{15}N HSQC and ^{1}H–^{13}C HSQC, 3D HNCA, 3D HN(CO)CA, 3D MQ-(H)CCH-TOCSY and 4D time shared $^{13}C/^{15}N$-edited NOESY were recorded.

All the NMR data were processed with NMRPipe (Delaglio et al. 1995). Resonance assignment was obtained using our previously developed 4D NOESY-based strategy (Xu et al. 2006) with the in-house software NMRspy together with the XYZ4D extension (http://yangdw.science.nus.edu.sg/Software&Scripts/XYZ4D/index.htm). 156 out of 160 residues in the RP_{tr} region were assigned in backbone resonances. Besides 6 histidine residues in the His-Tag, S1, A54, S152 and T154 were not assigned in amides since they were not observable in the 2D ^{1}H–^{15}N HSQC. Side-chain assignment was realized based on the $^{13}C,^{15}N$-edited sub-spectrum of the 4D NOESY and 3D MQ-(H)CCH-TOCSY with the aid of the ^{1}H-^{15}N HSQC and ^{1}H–^{13}C HSQC spectra. More than 95 % of the aliphatic and aromatic side chain resonances, together with ~90 % of the NH and NH_2 side chain resonances, were assigned. Unambiguous NOEs were obtained from 3 subspectra: $^{13}C,^{15}N$-edited, $^{13}C,^{13}C$-edited and $^{15}N,^{15}N$-edited 4D NOESY. The ambiguous NOEs were further assigned during iterated structure calculation and refinement.

NMR structure calculation

Distance constraints were obtained from the NOEs assigned, while dihedral angle restraints of Φ and ψ were calculated with the TALOS+program using the assigned chemical shifts of $^{13}C_\alpha$, $^{13}C_\beta$, $^{1}H_\alpha$, $^{1}H^N$ and $^{15}N^H$ (Cornilescu et al. 1999). 100 conformers were calculated with CYANA using the standard simulated annealing method (Güntert et al. 1997). Ten conformers with the lowest target function values were selected for analysis from the 100 calculated ones. The quality of the calculated structures was analyzed by MolMol software (Koradi et al. 1996) and PROCHECK program (Laskowski et al. 1996) in the form of root mean square deviation (RMSD) and Ramachandran plot parameters. The experimental and structural statistics are summarized in Table 1.

Solution structure of AcSp1-RP_{tr}

The solution structure of AcSp1-RP_{tr} was visualized using PyMol (http://www.pymol.org). In the $^{13}C,^{13}C$-edited 4D NOESY, we observed no NOEs arising from interactions of the His-Tag and RP_{tr} residues, indicating that the His-Tag is unstructured and has no specific interactions with RP_{tr}. The disorder of the His-Tag is further supported by the absence of its amide correlations in the HSQC. Excluding the N-terminal His-Tag region, RP_{tr} domain is composed of a compact seven-helix bundle (L4-S146) and a disordered C-terminal region (T147–G160). The 7 helices (H1–H7) span residues L4-K18 (H1), N20–A27 (H2), A32–L50 (H3), A54–N70 (H4), P77–S94 (H5), A103–Y125 (H6) and E131–S146 (H7) respectively, and are intercepted by short turns (Fig. 1a–c). This is consistent with the secondary structure characterized by CSI. The four residues with unassigned backbone NHs are all from the relatively flexible regions: S1 connects to the His-Tag; A54 is located at the beginning of helix 4; and S152 and T154 lie in the disordered C-terminal region. The tight packing of these 7 helices is mediated by strong hydrophobic interactions. The 7 helices shield a hydrophobic core in a barrel-like manner, which is formed predominantly by long hydrophobic side chains of Val, Leu and Ile (Fig. 1a, e). Besides this core, hydrophobic interactions are also found in between two adjacent helices. All positively (6 Arg and 7 Lys) and negatively (2 Asp and 3 Glu) charged residues are located on the surface of the structure, with their side chains exposed to the solvent. In addition, the surface of the RP_{tr} domain is very hydrophilic without large hydrophobic

Table 1 NMR experimental and structural statistics of RP_{tr}

NOE constraints	3,440
Sequential (|i − j| = 1)	1,125
Medium-range (1 < |i − j| <5)	1,231
Long-range (|i − j| ≥5)	1,084
Dihedral angle restraints	186
Ramachandran plot	
Most favored region	85.5 %
Additionally allowed region	13.2 %
Generously allowed region	0.7 %
Disallowed region	0.7 %
Average maximum violations per structure	
Distance (Å)	0.22 ± 0.02
Van der waals (Å)	0.32 ± 0.03
Torsion angles (°)	3.1 ± 0.29
CYANA target function value ($Å^2$)	3.74 ± 0.09
Average RMSD to mean structure (Å)	
Backbone heavy atoms (4–146)	0.30 ± 0.08
Heavy atoms (4–146)	0.71 ± 0.12

patches (Fig. 1d). The calculated solution structure of AcSp1-RP_{tr} has been deposited in PDB data bank (accession code: 2LYI).

Discussion and conclusion

Comparison with structural homologues

Online homologue search of AcSp1-RP_{tr} with DALI server (Holm and Rosenstrom 2010) gave several hits, all of which are helix-rich domains. The top matched structures are TuSp1-RP_2 and TuSp1-RP_1 from the same species, *N. antipodiana* (Lin et al. 2009). Although the sequence identities between TuSp1-RP_2 and AcSp1-RP_{tr} and between TuSp1-RP_1 and AcSp1-RP_{tr} are only 26 and 21 %, AcSp1-RP_{tr} adopts a similar topology to TuSp1-RP_2 and TuSp1-RP_1 in the first six helices with backbone RMSDs of 2.2 and 2.7 Å, respectively (Fig. 2). The orientation of the seventh helix in TuSp1-RP_1 was not determined due to its weak interactions with other helices, but the seventh helix in AcSp1-RP_{tr} has strong interactions with H1 and H6. Different from AcSp1-RP_{tr} and TuSp1-RP_1, TuSp1-RP_2 is lack of the seventh helix.

Although the sequence similarity of AcSp1-RP and TuSp1-RP domains is low, the high similarity in structure suggests that these two spidroins may share some common characters in fiber formation mechanism and silk fiber properties. Furthermore, they both participate in the egg case formation: AcSp1 forms the inner egg case, while TuSp1 forms the outer egg case (Hayashi et al. 2004; Lin et al. 2009). The structure conservation may result from the common functional selection during thousands of years of evolution.

Solubility and stability

The RP_{tr} domain could be concentrated to ~500–600 mg/ml (50–60 %, w/w) without any precipitation in a buffer with 20 mM Tris–HCl, 300 mM NaCl and 1 mM EDTA at pH 8.0. The extremely high solubility can be explained by the hydrophilic and largely positively charged surface. The excellent solubility of the RP_{tr} domain should greatly contribute to the high concentration of AcSp1 in the silk gland since the RP domains account for a majority of the full length protein.

Temperature induced denaturation monitored by circular dicroism demonstrated that RP_{tr} is highly stable with a T_m value of 82 °C in 5 mM sodium phosphate at pH 6.0. The high thermal stability is most probably due to the strong hydrophobic interactions among helices in the tightly packed hydrophobic core, which is mediated through a large number of long hydrophobic side chains of Val, Leu, and Ile.

Silk formation

When 3 ml RP samples containing ~5 mg/ml protein in a buffer with 20 mM Tris–HCl, 300 mM NaCl and 1 mM EDTA at pH 8.0 were swayed in 15 ml falcon tubes at a rate of 20 rpm, silk fibers were formed within 30 min. Besides silk fibers, precipitates were observed too. The fibers were sticky and stuck to one another. Thus it was difficult to obtain a single string of fiber (Fig. 3a). Most silk fibers had a diameter of 10–15 μm as observed under a scanning electron microscope (SEM) (JEOL JSM-6700F) (Fig. 3c, d). The result shows that the RP alone has the ability to form silk fibers. This is different from other investigated spider silk proteins for which both CTD/NTD and RP are required (Askarieh et al. 2010; Hagn et al. 2010; Lin et al. 2009; Stark et al. 2007). To test if the unstructured linker region is necessary for silk fiber formation, the same silk formation experiment was conducted on RP_{tr} samples under the same conditions. Similar to RP, RP_{tr} self-assembled into silk fibers (Fig. 3b). Interestingly, no precipitates were observed for the RP_{tr} samples. Because both RP and RP_{tr} contained the same His-Tag at the N-terminus and the linker was located at the C-terminus, the observed differences in fiber formation and precipitation should result from the linker instead of the His-Tag. The result shows that the unstructured linker region containing three F/YPG/SGDA/SG motifs does not contribute to the silk fiber formation but to the random aggregation that leads to the observed precipitates.

Spider and silkworm silk proteins adopt α-helical and random coil structures in the liquid state, but they contain a significant amount of β-strand structure in the fiber state (Lefevre et al. 2011). Thus it is believed that structural transition occurs during the assembly of silk protein molecules to form silk fibers. In the absence of shear force, RP_{tr} did not give rise to any observable precipitates or fibers at protein concentrations less than 10 mg/ml even after the samples were kept at room temperature for 2 weeks. This result is consistent with the high stability and solubility of RP_{tr}. In the presence of shear force produced by swaying the sample tubes, however, RP_{tr} could form silk fibers within 30 min. It is reasonable to assume that RP_{tr} can be destabilized significantly or partially unfolded by mechanical force. According to the RP_{tr} structure (Fig. 2a, c), the second helix has weaker hydrophobic interactions with its proximal residues in comparison with other 6 helices. Hence, the second helix likely unfolds first in the presence of shear force. Unfolding of this helix will result in partial or complete exposure of many buried hydrophobic residues (including L16, L22, V25, V34, I38, V39, V68, I84, V114, M117, V127, V129, V134, V141) to the solvent, which can initiate oligomerization and further structural changes of RP_{tr} to form silk fibers. To fully

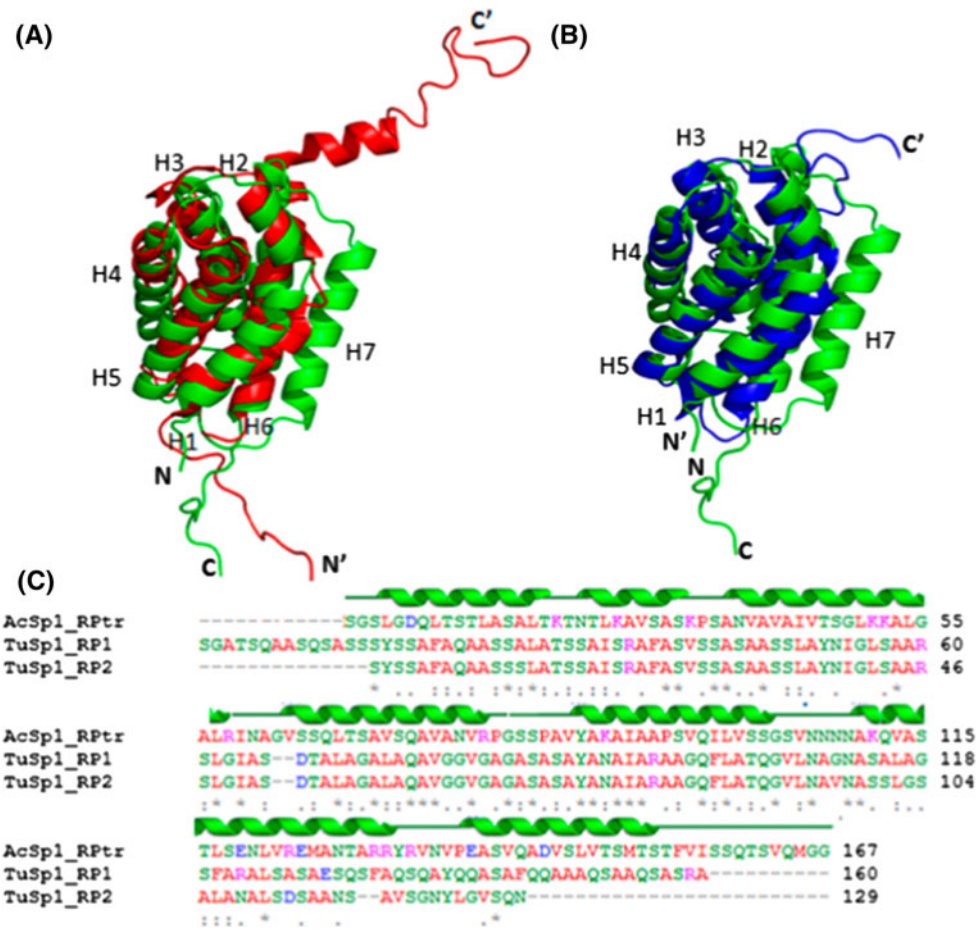

Fig. 2 Structure comparison and sequence alignment of AcSp1-RP_{tr} and TuSp1-RP. **a** Structure comparison of AcSp1-RP_{tr} (*green*, PDB code: 2LYI) and TuSp1-RP_1 (*red*, PDB code: 2K3N) with a Z score of 12.2 and RMSD of 2.7 Å. **b** Structure comparison of AcSp1-RP_{tr} (*green*) and TuSp1-RP_2 (*blue*, PDB code: 2K3O) with a Z score of 14.0 and RMSD of 2.2 Å. **c** Sequence alignment of RPs. Sequence identities between RP_{tr} and TuSp1-RP_1 and between RP_{tr} and TuSp1-RP_2 are 21 and 26 %, respectively. The secondary structure of AcSp1-RP_{tr} is indicated on the top of the sequences in **c**

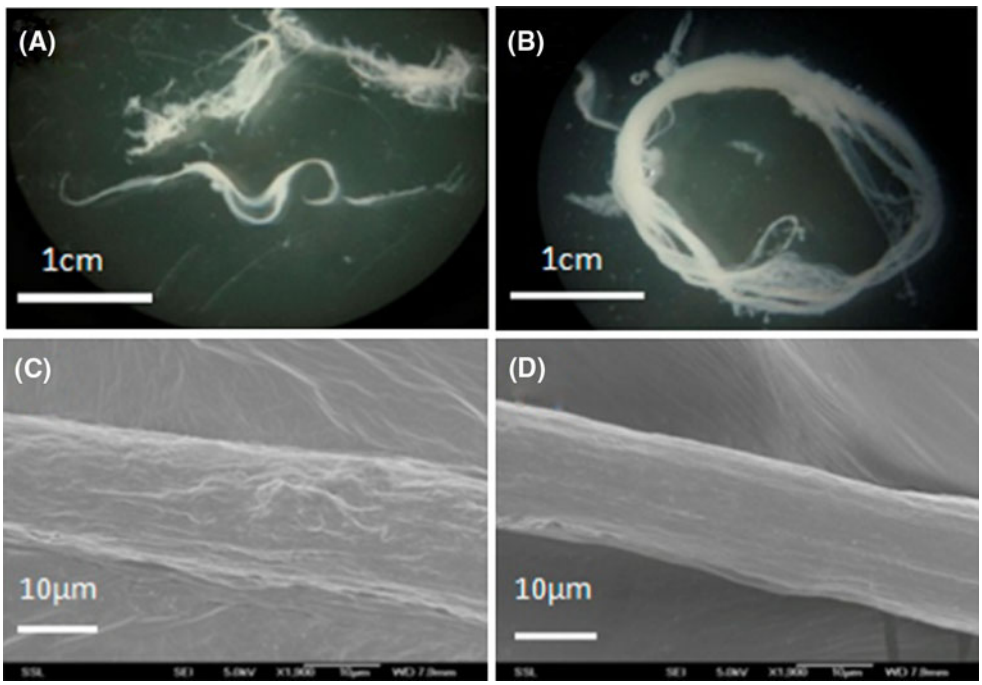

Fig. 3 Appearances of artificial Ac silk fibers. Light microscopic images of silk fibers formed from RP (*a*) and RP_{tr} (*b*) SEM images of single strings of silk fibers formed from RP (*c*) and RP_{tr} (*d*)

understand the ordered assembly process, besides the structures of the protein in the initial solution state and the final fiber state, one has to have the structures of the intermediate states. At present, no good experimental tools are available for determining the intermediate structures at atomic resolution since the intermediates may not be homogenous in size and conformation. With the relatively small size, capacity of silk fiber formation and availability of solution structure, AcSp1-RP_{tr} can serve as a model system for the development of methods for obtaining more detailed structural information of the heterogeneous intermediates, and it will be a good model for untangling the mechanism of ordered assembly of silk proteins.

Acknowledgments This work was supported by a grant from Ministry of Education, Singapore (R154000453112).

References

Askarieh G, Hedhammar M, Nordling K, Saenz A, Casals C, Rising A, Johansson J, Knight SD (2010) Self-assembly of spider silk proteins is controlled by a pH-sensitive relay. Nature 465(7295): 236–238

Cornilescu G, Delaglio F, Bax A (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology. J Biomol NMR 13(3):289–302

Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 6(3):277–293

Gatesy J, Hayashi C, Motriuk D, Woods J, Lewis R (2001) Extreme diversity, conservation, and convergence of spider silk fibroin sequences. Science 291(5513):2603–2605

Güntert P, Mumenthaler C, Wüthrich K (1997) Torsion angle dynamics for NMR structure calculation with the new program DYANA. J Mol Biol 273(1):283–298

Hagn F, Eisoldt L, Hardy JG, Vendrely C, Coles M, Scheibel T, Kessler H (2010) A conserved spider silk domain acts as a molecular switch that controls fibre assembly. Nature 465(7295):239–242

Hayashi CY, Blackledge TA, Lewis RV (2004) Molecular and mechanical characterization of aciniform silk: uniformity of iterated sequence modules in a novel member of the spider silk fibroin gene family. Mol Biol Evol 21(10):1950–1959

He YX, Zhang NN, Li WF, Jia N, Chen BY, Zhou K, Zhang J, Chen Y, Zhou CZ (2012) N-Terminal domain of *Bombyx mori* fibroin mediates the assembly of silk in response to pH decrease. J Mol Biol 418(3–4):197–207

Hijirida DH, Do KG, Michal C, Wong S, Zax D, Jelinski LW (1996) C-13 NMR of *Nephila clavipes* major ampullate silk gland. Biophy J 71(6):3442–3447

Holm L, Rosenstrom P (2010) Dali server: conservation mapping in 3D. Nucleic Acids Res (Web Server issue) 38:545–549

Huang W, Lin Z, Sin YM, Li D, Gong Z, Yang D (2006) Characterization and expression of a cDNA encoding a tubuliform silk protein of the golden web spider *Nephila antipodiana*. Biochimie 88(7):849–858

Koradi R, Billeter M, Wuthrich K (1996) MOLMOL: a program for display and analysis of macromolecular structures. J Mol Graph 14(1):51–55

Laskowski RA, Rullmann JAC, MacArthur MW, Kaptein R, Thornton JM (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. J Biomol NMR 8(4):477–486

Lefevre T, Boudreault S, Cloutier C, Pezolet M (2011) Diversity of molecular transformations involved in the formation of spider silks. J Mol Biol 405(1):238–253

Lewis RV (2006) Spider silk: ancient ideas for new biomaterials. Chem Rev 106(9):3762–3774

Lin Z, Huang W, Zhang J, Fan JS, Yang D (2009) Solution structure of eggcase silk protein and its implications for silk fiber formation. Proc Natl Acad Sci USA 106(22):8906–8911

Stark M, Grip S, Rising A, Hedhammar M, Engstrom W, Hjalm G, Johansson J (2007) Macroscopic fibers self-assembled from recombinant miniature spider silk proteins. Biomacromolecules 8(5):1695–1701

Vasanthavada K, Hu X, Falick AM, La Mattina C, Moore AM, Jones PR, Yee R, Reza R, Tuton T, Vierra C (2007) Aciniform spidroin, a constituent of egg case sacs and wrapping silk fibers from the black widow spider *Latrodectus hesperus*. J Biol Chem 282(48): 35088–35097

Xu YQ, Zheng Y, Fan JS, Yang DW (2006) A new strategy for structure determination of large proteins in solution without deuteration. Nat Methods 3(11):931–937

Xu L, Tremblay ML, Meng Q, Liu XQ, Rainey JK (2011) (1)H, (13)C and (15)N NMR assignments of the aciniform spidroin (AcSp1) repetitive domain of *Argiope trifasciata* wrapping silk. Biomol NMR Assign 6(2):147−151