



Solution Structure of Tubuliform Spidroin N-Terminal Domain and Implications for pH Dependent Dimerization

Megija Šede^{1†}, Jēkabs Fridmanis^{1†}, Martins Otikovs¹, Jan Johansson², Anna Rising^{2,3}, Nina Kronqvist² and Kristaps Jaudzems^{1,4*}

¹Department of Physical Organic Chemistry, Latvian Institute of Organic Synthesis, Riga, Latvia, ²Department of Biosciences and Nutrition, Neo, Karolinska Institutet, Huddinge, Sweden, ³Department of Anatomy, Physiology and Biochemistry, Swedish University of Agricultural Sciences, Uppsala, Sweden, ⁴Faculty of Chemistry, University of Latvia, Riga, Latvia

OPEN ACCESS

Edited by:

Kyeong Kyu Kim,
Sungkyunkwan University, South
Korea

Reviewed by:

Keiji Numata,
RIKEN, Japan
Seung-Wook Chi,
Korea Research Institute of Bioscience
and Biotechnology (KRIBB), South
Korea
Jeong-Yong Suh,
Seoul National University, South Korea

*Correspondence:

Kristaps Jaudzems
kristaps.jaudzems@osi.lv

[†]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Structural Biology,
a section of the journal
Frontiers in Molecular Biosciences

Received: 05 May 2022

Accepted: 26 May 2022

Published: 14 June 2022

Citation:

Šede M, Fridmanis J, Otikovs M,
Johansson J, Rising A, Kronqvist N
and Jaudzems K (2022) Solution
Structure of Tubuliform Spidroin N-
Terminal Domain and Implications for
pH Dependent Dimerization.
Front. Mol. Biosci. 9:936887.
doi: 10.3389/fmolb.2022.936887

The spidroin N-terminal domain (NT) is responsible for high solubility and pH-dependent assembly of spider silk proteins during storage and fiber formation, respectively. It forms a monomeric five-helix bundle at neutral pH and dimerizes at lowered pH, thereby firmly interconnecting the spidroins. Mechanistic studies with the NTs from major ampullate, minor ampullate, and flagelliform spidroins (MaSp, MiSp, and FISP) have shown that the pH dependency is conserved between different silk types, although the residues that mediate this process can differ. Here we study the tubuliform spidroin (TuSp) NT from *Argiope argentata*, which lacks several well conserved residues involved in the dimerization of other NTs. We solve its structure at low pH revealing an antiparallel dimer of two five- α -helix bundles, which contrasts with a previously determined *Nephila antipodiana* TuSp NT monomer structure. Further, we study a set of mutants and find that the residues participating in the protonation events during dimerization are different from MaSp and MiSp NT. Charge reversal of one of these residues (R117 in TuSp) results in significantly altered electrostatic interactions between monomer subunits. Altogether, the structure and mutant studies suggest that TuSp NT monomers assemble by elimination of intramolecular repulsive charge interactions, which could lead to slight tilting of α -helices.

Keywords: spider silk, tubuliform spidroin, N-terminal domain, NMR structure, dimerization mechanism

INTRODUCTION

Orb weaving spiders stand out among silk producers found in nature as they can produce up to seven different types of silk, each tailored for a specific function. The main components of the spider silk threads are large proteins called spidroins. Major ampullate spidroin (MaSp) is the main component of dragline silk, the strongest and toughest of all spider silk types (Ayoub et al., 2007). Spider silk formed from minor ampullate spidroin (MiSp) is used for auxiliary spiral stabilization (Gosline et al., 1986; Colgin and Lewis, 1998). Aciniform spidroin (AcSp) is the main component of wrapping silk used to immobilize prey (Tremblay et al., 2015). Pyriform spidroin (PySp) makes up the attachment discs, which lash the joints of the web and attaches dragline silk to surfaces (Perry et al., 2010). Flagelliform spidroin (FISP) forms capture spiral silk, which can extend up to 500% of its length (Eisoldt et al., 2011). Aggregate spidroin (AgSp), coats the FISP threads and is used as a glue to capture prey on the web (Opell and Hendricks, 2010; Collin et al., 2016). Tubuliform spidroin (TuSp)

forms the outer layer of spider egg cases and protects the eggs from the external environment (Tian and Lewis, 2005; Eisoldt et al., 2011).

Each spidroin is produced in a specific gland. The major ampullate gland, which has been studied most extensively, has a long, winding, and narrow tail, a wider ampulla or sac, and an S-shaped narrowing duct connected to the sac via a funnel (Andersson et al., 2013). After secretion in the tail segment, spidroins are held in the ampulla at up to 50% (w/w) concentrations (Hijirida et al., 1996) and upon silk spinning, pulled through the tapered duct. Here, changes in the surrounding environment conditions (i.e., reduced pH, altered ion composition, increased partial CO₂ pressure) and increasing shear forces cause them to assemble and form solid threads (Knight and Vollrath, 2001; Rousseau et al., 2004; Andersson et al., 2014; Sparkes and Holland, 2017, 2019). The tubuliform gland, on the other hand, is long, noodle-shaped, and smooth without a distinguishable ampulla-shaped storage sac (Chaw and Hayashi, 2018). Also, the conditions in the tubuliform gland are not known.

The conditionally high solubility and regulation of spider silk formation is mediated by two conserved spidroin terminal domains—the N-terminal domain (NT) and the C-terminal domain (CT) (Askarieh et al., 2010; Hagn et al., 2010), whereas the highly variable central repetitive domain (Rep), is responsible for the silk properties (Gosline et al., 1999). MaSp NTs from different spider species were early on found to be pH sensitive (Gaines et al., 2010; Landreh et al., 2010). In the ampulla at pH > 7 the NT forms a monomeric five-helix bundle showing a distinct dipolar distribution of charged amino acids (Askarieh et al., 2010; Jaudzems et al., 2012). Under these conditions, it likely promotes solubility of the spidroin Rep domain by forming the shell of micelle-like particles with the aggregation-prone regions sequestered in their core (Kronqvist et al., 2017). As spidroins are pulled through the spider silk duct, the pH is reduced to <5.7 (Andersson et al., 2014), which causes sequential protonation of a cluster of glutamic acids on the NT's surface. This disrupts several charge interactions leading to movement of α -helices, relocation of a wedged W10 side chain from a buried to a surface exposed conformation, and subsequent dimerization of the domain. The NT dimerization results in firmly interconnected spidroins (CT is dimeric already during storage), which ensure propagation of pulling forces during the silk fiber formation (Kronqvist et al., 2014; Ries et al., 2014).

The pH dependent dimerization mechanism seems conserved among NTs from different silk types and species (Heiby et al., 2017); however, the exact structural details are divergent (Otikovs et al., 2015). Through studies of site-directed mutants bearing glutamate to glutamine substitutions, the residues E79, E84 and E119 were identified to be protonated in the *Euprostenops australis* MaSp NT dimer (Kronqvist et al., 2014). In *Araneus ventricosus* MiSp NT, E84 is substituted by a serine and the nearby E73 was instead found to be protonated (Otikovs et al., 2015). Similar mutants were investigated for *Nephila clavipes*

FlSp NT and *Latrodectus hesperus* MaSp NT, however, the exact carboxylates to be protonated could not be identified (Bauer et al., 2016; Sarr et al., 2022). AcSp from *Nephila antipodiana* displays a very different charge distribution on the protein surface, nevertheless it still forms dimers at low pH and in presence of physiological salt concentrations (Chakraborty et al., 2020).

Three other MaSp NT mutants have been investigated to characterize its monomeric conformation. Residues D40 and K65 located at opposite tips of the NT molecular dipole form an intermolecular salt bridge in the dimer and were proposed to mediate initial monomer association (Schwarze et al., 2013; Kronqvist et al., 2014). The residue charge reversal in the mutant NT_{D40KK65D} resulted in disturbed dipolar interactions between NT monomers and abolished its pH sensitivity (Kronqvist et al., 2017). This variant shows excellent solubility-enhancing properties and has been applied as solubility tag for recombinant production of aggregation-prone proteins and peptides (Kronqvist et al., 2017, 2022; Sarr et al., 2018; Abelein et al., 2020). Another pH independent NT monomer was prepared by replacing an alanine residue in the middle of the dimer subunit interface with an arginine, thereby preventing self-association through charge repulsion (Jaudzems et al., 2012). In the third mutant, all six methionine residues in the protein core were replaced by leucines, which significantly reduced protein plasticity and abolished the movement of α -helices necessary for NT dimerization (Heiby et al., 2019).

The structures of all TuSp domains from *Nephila antipodiana* were the first published silk protein structures (Lin et al., 2009). The TuSp NT structure was determined by solution NMR in presence of 100 mM dodecylphosphocholine (DPC) to avoid protein aggregation and using a protein construct, which lacked the first 36 amino acids. The structure shows a four-helix bundle, which differs from the more recent structures of other NTs comprising five helices (**Figure 1**). The amino acid sequence of TuSp NT also displays some notable differences in comparison to the consensus sequence of other NTs. The E119 of MaSp NT involved in the protonation during dimerization is replaced by an arginine. Furthermore, the sequence contains no methionines, which were found to enable the helical reorganization during MaSp NT dimerization. Additionally, in TuSp NT (similarly to FlSp NT), W10 is replaced by a phenylalanine.

Considering the additional mechanistic insights gained from the recent studies of MaSp, MiSp and FlSp NT and the contrasting structural data of TuSp NT, we decided to re-investigate its structure and possible dimerization mechanism using a protein construct that comprises the complete conserved domain sequence from *Argiope argentata*. We use nuclear magnetic resonance (NMR) to characterize the TuSp NT conformation at pH > 7 and to determine its structure at pH 5.5. Next, several methods including NMR, size exclusion chromatography (SEC), and circular dichroism (CD) are used on the wild type (wt) protein as well as site-directed mutants to investigate the dimerization mechanism of TuSp NT.

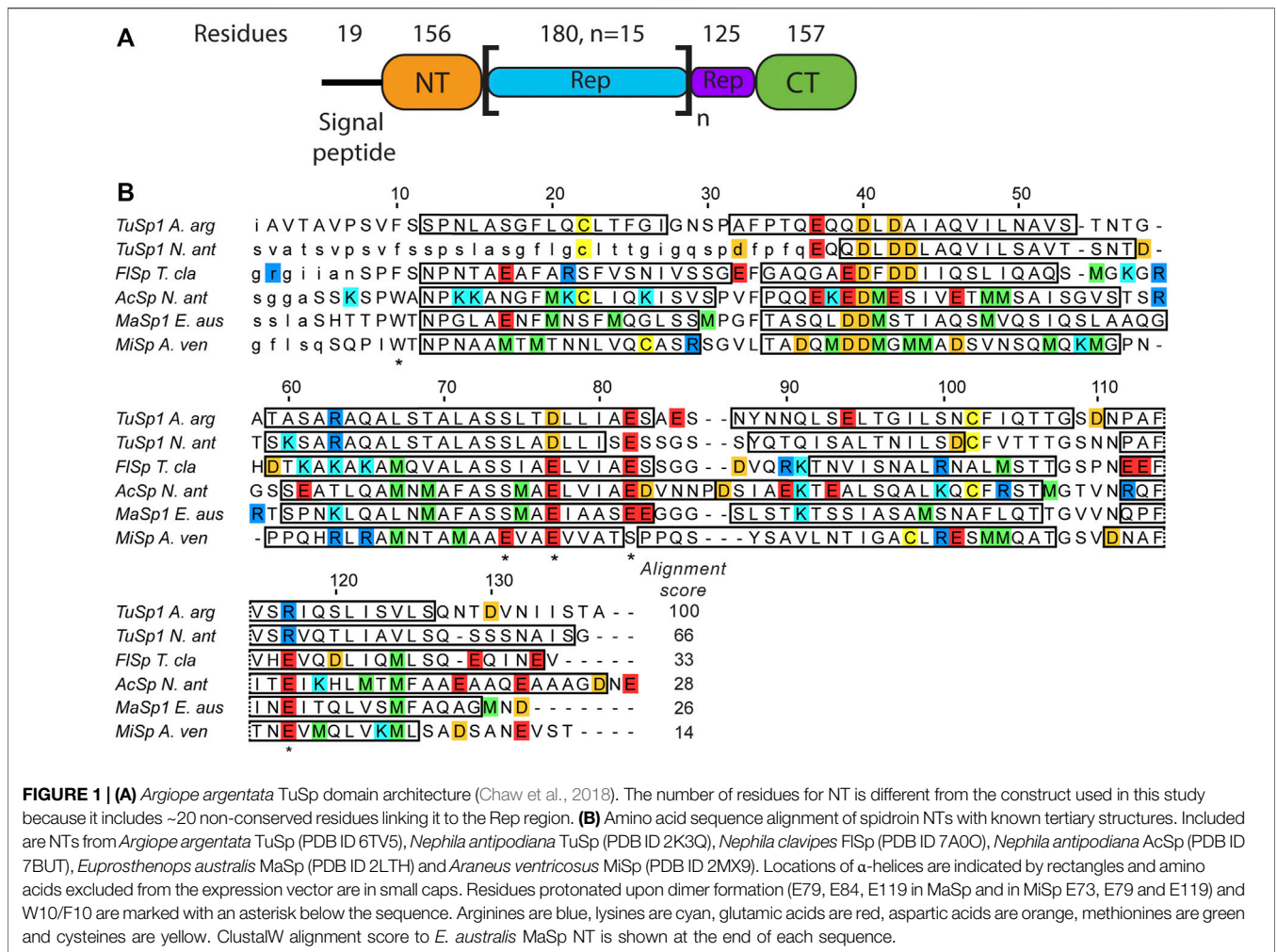


FIGURE 1 | (A) *Argiope argentata* TuSp domain architecture (Chaw et al., 2018). The number of residues for NT is different from the construct used in this study because it includes ~20 non-conserved residues linking it to the Rep region. **(B)** Amino acid sequence alignment of spidroin NTs with known tertiary structures. Included are NTs from *Argiope argentata* TuSp (PDB ID 6TV5), *Nephila antipodiana* TuSp (PDB ID 2K3Q), *Nephila clavipes* FISp (PDB ID 7A0O), *Nephila antipodiana* AcSp (PDB ID 7BUT), *Euprosthenoops australis* MaSp (PDB ID 2LTH) and *Araneus ventricosus* MiSp (PDB ID 2MX9). Locations of α -helices are indicated by rectangles and amino acids excluded from the expression vector are in small caps. Residues protonated upon dimer formation (E79, E84, E119 in MaSp and in MiSp E73, E79 and E119) and W10/F10 are marked with an asterisk below the sequence. Arginines are blue, lysines are cyan, glutamic acids are red, aspartic acids are orange, methionines are green and cysteines are yellow. ClustalW alignment score to *E. australis* MaSp NT is shown at the end of each sequence.

METHODS

Site-Directed Mutagenesis of *A. argentata* TuSp NT

pET-32a-HisTrxHisNT plasmid containing wt *A. argentata* TuSp 1 NT sequence and pET-28a-HisNT containing TuSp NT_{E37QE82Q} and TuSp NT_{E37QE82QE85Q} were ordered from BioCat GmbH (Heidelberg, Germany). For expression of TuSp NT, TuSp NT_{A70E}, and TuSp NT_{A70R} the pET-32a vector was used. TuSp NT_{D40RR63D}, TuSp NT_{E37QE82Q}, TuSp NT_{E37QE82QE85Q}, and TuSp NT_{E37QE82QE85QD130N} were expressed by using the pET-28a vector (Supplementary Figure S1 for the amino acid sequences). To obtain TuSp NT_{A70E}, TuSp NT_{A70R} and TuSp NT_{E37QE82QE85QD130N}, pET-32a-HisTrxHisNT plasmid containing TuSp NT sequence or pET-28a-HisNT containing TuSp NT_{E37QE82QE85Q} were subjected to point mutagenesis using Phusion site-directed mutagenesis kit (ThermoFisher Scientific, United States). At first, two PCR reactions were carried out in parallel with forward and reverse primers in separate tubes. After PCR, the two products were combined in a 1:1 ratio and another

thermocycle was run. PCR products were digested with DpnI, and purity of plasmids was checked on 1% agarose gel. The plasmids were used to transform chemically competent *E. coli* XL1-Blue cells by heat shock, which was followed by plasmid extraction and sequence verification.

Protein Expression and Purification

The plasmids were used for heat shock transformation of chemically competent *E. coli* BL21 (DE3) cells. Overnight cultures were inoculated at a ratio of 1:100 into LB medium containing 50 μ g/ml kanamycin (pET-28a plasmids) or 100 μ g/ml ampicillin (pET-32a plasmids) and cells were further grown at 37°C and 180 rpm to an OD₆₀₀ of 0.6. Expression was induced by addition of isopropyl β -D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.5 mM and cultures were further incubated overnight at 25°C. The cells were harvested by centrifugation at 7000 g for 10 min and stored at -20°C.

Cells were thawed and resuspended in the immobilized metal-affinity chromatography (IMAC) loading buffer (20 mM Tris-HCl, 8 M urea, 300 mM NaCl, 15 mM imidazole, 1 mM DTT, pH 8.0). The cells were sonicated, and the lysate was cleared by

centrifugation at 20000 g for 40 min. The supernatant was then loaded on a 2×5 ml HisTrap HP column (Cytiva). Unbound proteins were washed off with 5 column volumes of loading buffer. Bound proteins were eluted in 2 ml fractions using 500 mM imidazole, 20 mM Tris-HCl, 8 M urea, 300 mM NaCl, 1 mM DTT, pH 8.0. The absorbance at 280 nm was measured for each fraction and protein-containing fractions were pooled and dialyzed against 20 mM Tris-HCl, 300 mM NaCl, 1 mM EDTA, 1 mM DTT, pH 8.0 at 4°C. Complete refolding of the protein was verified by NMR. Fusion proteins were cleaved with tobacco etch virus (TEV) protease (ratio of 1:10–1:20, enzyme to substrate, w/w) during dialysis against TEV reaction buffer (50 mM Tris-HCl, 1 mM DTT, 0.5 mM EDTA, pH 8.0) at 4°C. After the cleavage, sample was loaded on the HisTrap HP column to remove the cleaved fusion tag and unbound target protein was collected. Protein was further purified by gel filtration with a HiLoad 16/600 Superdex 75 pg size exclusion column (Cytiva). The concentration of protein was determined by using the Pierce Bicinchoninic Acid (BCA) protein assay kit (Thermo Fisher Scientific, United States) according to the manufacturer's recommendations. The correct size of protein was confirmed by SDS-PAGE using 4–12% Bis-Tris polyacrylamide gel (ThermoFisher Scientific, United States), stained with Coomassie brilliant blue dye.

For production of ^{15}N - and ^{13}C , ^{15}N -labeled samples the same procedure was used except that M9 minimal medium containing $^{15}\text{NH}_4\text{Cl}$ and ^{13}C -glucose as the sole sources of nitrogen and glucose, respectively, was used.

NMR Spectroscopy

The protein sample used for NMR structure determination was concentrated to approximately 1 mM in 20 mM sodium acetate- d_3 (NaOAc- d_3), pH 5.5, 20 mM NaCl, 0.03% (w/v) NaN_3 , 5% D_2O (v/v) buffer. NMR spectra were acquired at 298 K on a Bruker Avance Neo 600 MHz spectrometer equipped with a TCI z-gradient cryoprobe. For backbone assignment, 3D HNCA, 3D HNC(O), 3D CBCA(CO)HN, and 3D HN(CA)CO spectra were acquired. For side chain assignment and structure determination, 3D [$^1\text{H}, ^1\text{H}$]-nuclear Overhauser spectroscopy (NOESY)- ^{15}N -heteronuclear single-quantum correlation (HSQC), 3D [$^1\text{H}, ^1\text{H}$]-NOESY- ^{13}C (aliphatic)-HSQC, 3D [$^1\text{H}, ^1\text{H}$]-NOESY- ^{13}C (aromatic)-HSQC spectra were acquired with a mixing time of 80 ms. Chemical shifts were referenced internally to the residual water signal at 4.77 ppm relative to 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS).

For recording of 2D ^{15}N - ^1H HSQC NMR spectra, ^{15}N -labeled samples were prepared in either 20 mM NaOAc- d_3 , 20 mM NaCl, pH 5.5 or 20 mM sodium phosphate (NaPi), 300 mM NaCl, pH 7.2 buffers. The spectra were acquired using either Bruker Avance III HD 800 MHz spectrometer equipped with an TXI z-gradient room temperature probe or Bruker Avance Neo 600 MHz spectrometer equipped with a TCI z-gradient cryoprobe.

The molecular weight of TuSp NT at pH 5.5 was estimated using 2D ^1H -detected ^{15}N relaxation experiments. The relaxation data were recorded on the Bruker Avance III HD 800 MHz spectrometer equipped with an TXI z-gradient room temperature probe. Spectra were processed in TopSpin 4.1.1

and signal intensities were measured with CARA 1.8.4.2 (Keller, 2004). Residue-specific protein backbone amide ^{15}N NMR R_1 and R_2 relaxation times were calculated from the peak intensities using two-parameter exponential fit model in Relax 4.0.3 (d'Auvergne and Gooley, 2008). Seven T_1 (20–2000 ms) and seven T_2 (10–250 ms) delays were used for the recording of the decay curves. Protein rotational correlation time (τ_c) was calculated using the equation

$$\tau_c \approx \frac{1}{4\pi\nu_N} \sqrt{6 \frac{T_1}{T_2} - 7}$$

(Kay et al., 1989), where ν_N is the ^{15}N resonance frequency (in Hz), T_1 is average longitudinal and T_2 average transverse relaxation time value of rigid protein regions.

NMR Structure Calculation

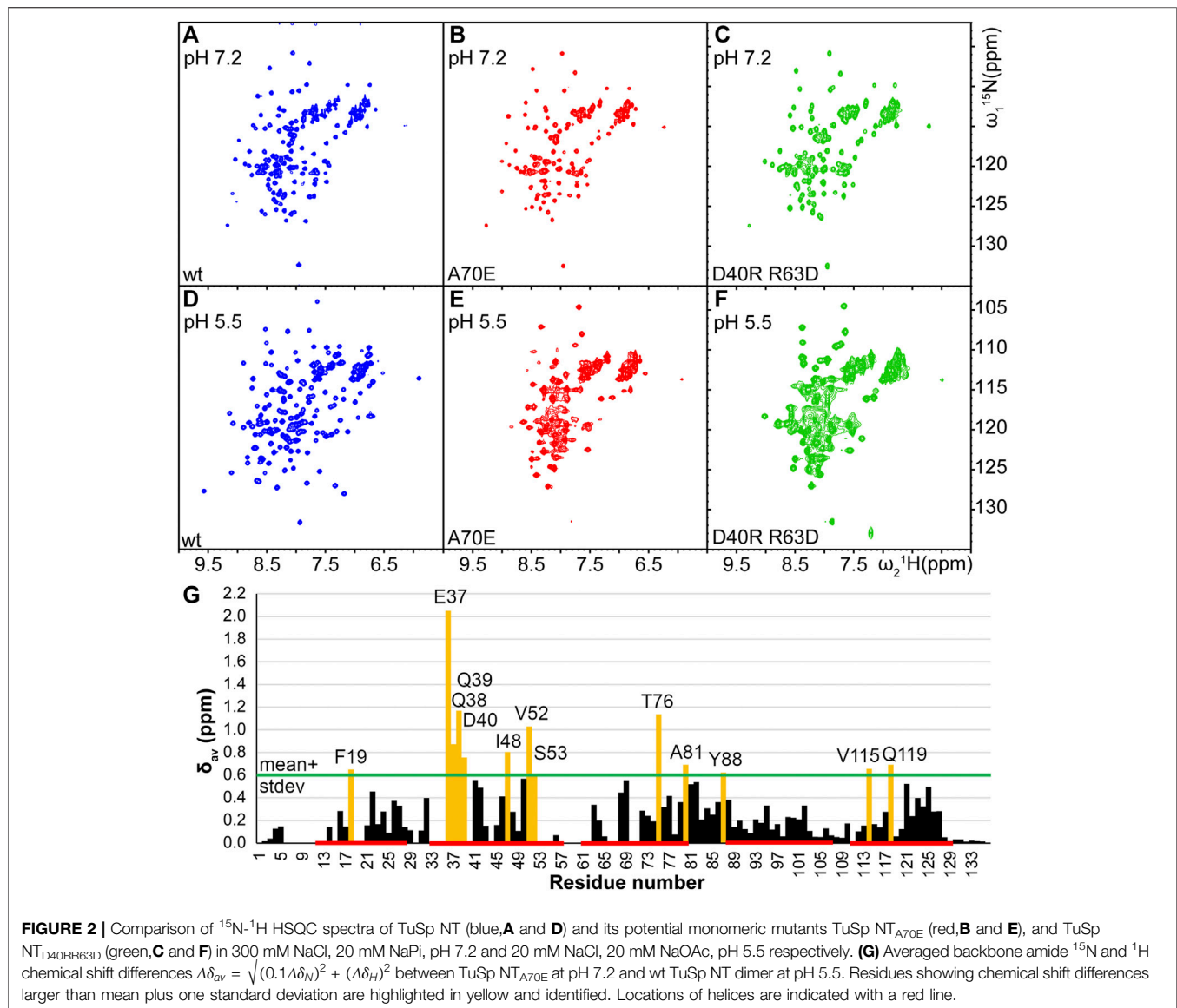
All acquired spectra were processed using TopSpin 4.1.1 software. The assignment of protein backbone and side chain chemical shifts was performed with CARA 1.8.4.2 (Keller, 2004). The peak lists for dimeric structure calculation were acquired from an initial monomeric structure calculation using UNIO-ATNOS/CANDID 2.0.2 (Herrmann et al., 2002a, 2002b; Serrano et al., 2012) in conjunction with CYANA 2.1 (Güntert et al., 1997). Subsequent manual inspection of the NOESY spectra allowed the identification of 11 unambiguous unique intermolecular contacts that were used as distance constraints with an upper limit of 5 Å in the following structure calculations. Notably, these contacts are only a small fraction of all intermolecular contacts found during automated NOE assignment and structure calculation. The dimeric structure calculation was done using seven iterations of CYANA protocol, starting with 100 random conformers that were subjected to simulated-annealing with 10,000 steps of torsion-angle molecular dynamics. The structural statistics were further improved by applying TALOS+ angle restraints. For the final structure calculation, a homology model obtained from SwissProt was used as a starting conformer in the first cycle of the calculation to improve convergence. The 20 conformers with the lowest residual CYANA target-function values after CYANA cycle 7 were energy minimized in explicit water using CNS (Brunger, 2007).

Size Exclusion Chromatography

Size exclusion chromatography on the purified TuSp NT and its variants was performed using an Äkta Purifier 10 system. 50 μl of purified protein at a concentration of 3 mg/ml were loaded onto a Superdex 75 Increase 10/300 GL size exclusion column (Cytiva) previously equilibrated in the relevant buffer (20 mM Tris-HCl, 300 mM NaCl, pH 8.0 or 20 mM NaOAc, 20 mM NaCl, pH 5.5). The molecular weight calculation was performed according to the calibration found in the column manual, where conalbumin (75 kDa) was measured to elute at 8.7 ml, ovalbumin (44 kDa) at 9.65 ml, carbonic anhydrase (29 kDa) at 11.2 ml, ribonuclease (13.7 kDa) at 13.3 ml and aprotinin (6.5 kDa) at 16.1 ml.

Circular Dichroism Spectroscopy

CD experiments were performed on a Jasco J-1500 CD spectrometer using 300 μl cuvettes with 1 mm path length. All



measurements were acquired in 20 mM sodium phosphate buffer at pH 8.0 or pH 5.5 at 10 μM protein concentration. Spectra were recorded from 260 to 185 nm, first at 25°C, then after heating for 15 min at 95°C and finally after cooling down to 25°C. 5 scans were recorded for each temperature to calculate an average spectrum. For temperature melting curves, the CD intensity at 222 nm was monitored between 25 and 95°C with 1°C/min increase.

RESULTS

NMR Sample Preparation

A. argentata TuSp NT with N-terminal His-Trx fusion tag was expressed as insoluble inclusion bodies in *E. coli* BL21 (DE3). Following solubilization in 8 M urea, the protein was purified using IMAC and refolded by dialysis against Tris buffer at pH 8

and 4°C. The His-Trx solubility tag was subsequently cleaved with TEV protease. The cleaved protein was once again passed through an IMAC column to remove the tag and further purified by gel filtration in presence of 300 mM NaCl to isolate monomeric TuSp NT.

In order to determine optimal solution conditions for the NMR structure determination, the protein was dialyzed against several different buffers with pH between 5.5 and 8, and variable amount of salt. The spectral quality was evaluated by ^{15}N - ^1H HSQC NMR spectroscopy. A high-quality spectrum was obtained only at low pH conditions corresponding to the spider's spinning duct (20 mM NaOAc at pH 5.5, 20 mM NaCl), as indicated by ~130 well separated peaks with similar intensities (**Figure 2**). At high pH conditions as in the ampulla (20 mM phosphate at pH 7.2 or 20 mM Tris-HCl at pH 8) and in presence of 0.3–1 M NaCl the number of well separated observed peaks was reduced to ~110 suggesting that the protein is in conformational exchange.

NMR Analysis of TuSp NT Monomer at High pH

As described above, at pH > 7 and in presence of high salt concentrations (i.e. conditions, where MaSp NT is monomeric) TuSp NT showed a reduced number of HSQC NMR signals. To stabilize its monomeric conformation, three mutants were prepared by introducing single or double-site mutations, which prevent dimerization of MaSp NT (Jaudzems et al., 2012; Kronqvist et al., 2017). Firstly, A70 located in the middle of the dimerization interface was replaced with glutamic acid or arginine residues that could prevent dimerization through electrostatic and steric repulsion (mutants TuSp NT_{A70E} and TuSp NT_{A70R}). Secondly, the molecular dipole of each subunit important for the initial association of monomers was disrupted by charge reversal of the residues D40 and R63 (mutant TuSp NT_{D40RR63D}). TuSp NT_{A70E} and TuSp NT_{D40RR63D} variants could be produced in sufficient quantities for NMR experiments using the same protocol as for the wt protein, whereas TuSp NT_{A70R} could not be properly refolded. The inability to refold TuSp NT_{A70R} is likely associated with formation of a new electrostatic interaction that is incompatible with the correct fold of the protein.

The HSQC spectra of TuSp NT_{A70E} and TuSp NT_{D40RR63D} at both, pH 7.2 and pH 5.5, were similar to the spectrum of wt TuSp NT at pH 7.2 suggesting that dimerization was prevented. However, for TuSp NT_{A70E} at pH 5.5 and for TuSp NT_{D40RR63D} at both pH values the spectra also showed peak broadening implying that the monomeric conformation was not fully stabilized (Figure 2). The HSQC spectrum of TuSp NT_{A70E} showed slight improvement in spectral quality at pH 7.2 and 300 mM NaCl compared to wt protein as manifested by a more uniform peak intensity distribution. Therefore, we recorded and analyzed 3D triple-resonance spectra for chemical shift assignment of its backbone resonances. The assignment was obtained for 104 out of 136 residues and analysis of secondary chemical shifts indicated presence of five α -helices (Supplementary Figure S2). The residues 7–16, 20–21, 34–36, 54–63, 67–68, 71–72 corresponding largely to the N-terminal loop region, N-terminal part of first helix, loop between second and third helix as well as first half of third alpha helix could not be assigned. Mapping of these regions on the MaSp NT monomer structure showed that they cluster in the proximity of W10 (Supplementary Figure S2), suggesting that F10 in TuSp NT could be involved in a conformational exchange process, which results in disappearance of ~20 backbone amide signals at pH 7.2.

NMR Structure of TuSp NT Dimer at Low pH

To verify that TuSp NT is in a dimeric conformation at pH 5.5, ¹⁵N NMR relaxation measurements were performed. The bulk T₁ and T₂ relaxation times were determined to be 605 and 17 ms affording an estimated molecular tumbling correlation time of about 17 ns (assuming molecular isotropic motion), which matches well with the expected molecular mass of the TuSp NT dimer (28 kDa). The backbone assignment was performed using standard triple-resonance 3D experiments (Supplementary Figure S3) and distance restraints were obtained from three 3D¹⁵

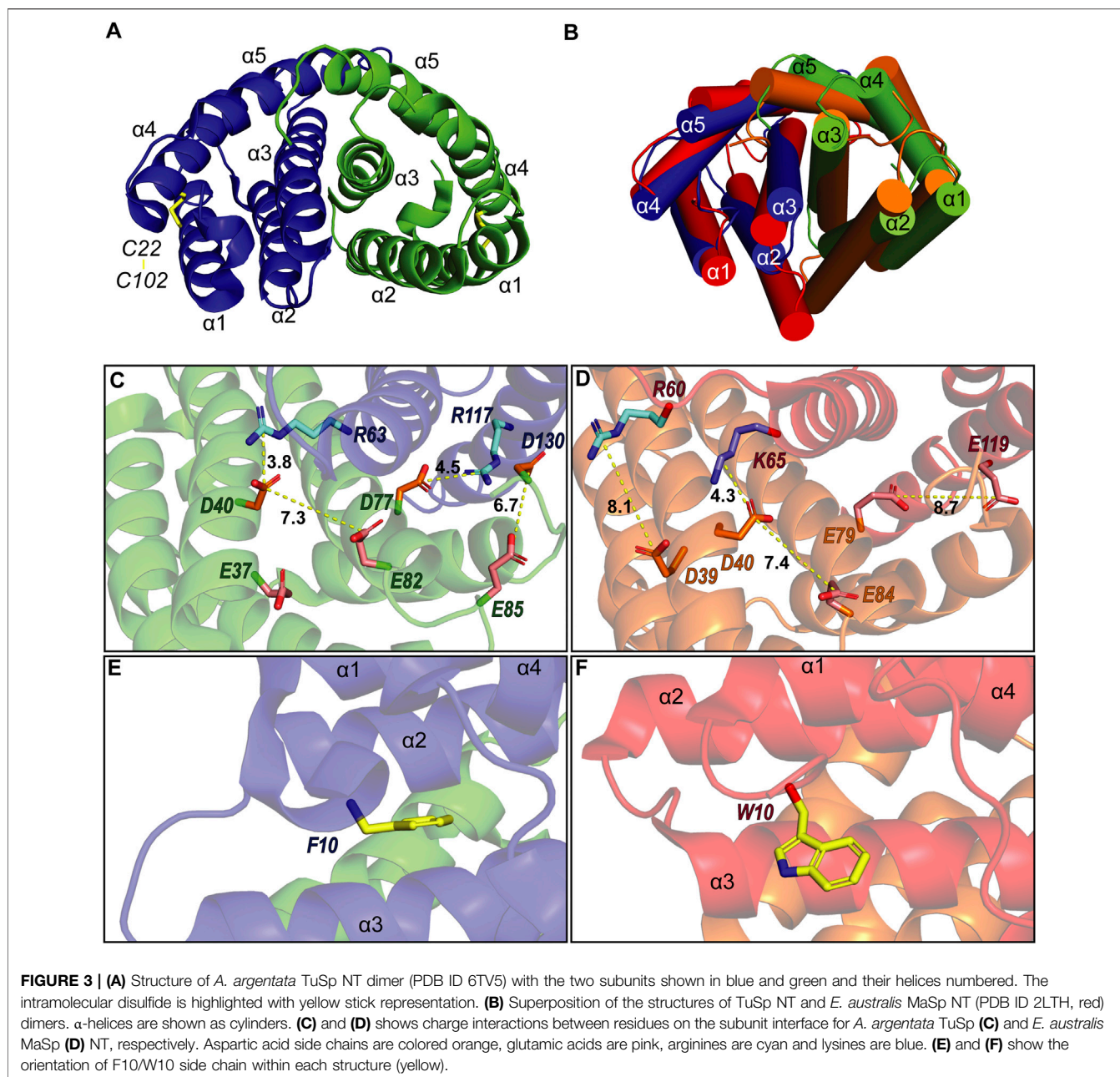
N/¹³C-resolved NOESY-HSQC spectra. The dimeric structure of TuSp NT was calculated from 3663 NOEs including 84 intermolecular restraints (Figure 3, Table 1 for structural statistics). The folded part of each subunit is composed by residues 7–127 and consists of five α -helices (Figure 3A). The first and fourth α -helices are covalently linked through a disulfide bond between C22 and C102, which is confirmed by the cysteine ¹³C β chemical shift values of ~37 ppm corresponding to the oxidized state. The dimer interface is formed by the second, third and fifth α -helices. The overall structure is an antiparallel homodimer and strongly resembles the dimeric structures of other spidroin NTs (Figure 3B) displaying a backbone RMSD of 2.11 Å to MaSp NT (over 121 residues). However, the helical orientations are slightly different, in particular, the C-terminus of α 2 is tilted towards α 1 and α 3.

As described in the introduction, charged residues play an important role in the dimerization mechanism of the NTs. Of the three glutamates (E79, E84, and E119) that become protonated during MaSp NT dimerization, only E84 is conserved in TuSp NT (as E82). This residue shows a similar side chain orientation in the TuSp NT dimer, pointing towards D40 of the same subunit, with which it forms a characteristic handshake interaction in the MaSp NT dimer structure (Figures 3C,D). However, E79 and E119 in MaSp NT are replaced by D77 and R117 in TuSp NT. Since these residues also show similar side chain orientations, an intermolecular salt bridge is formed in the TuSp NT dimer. This finding implies a different mechanism for pH dependent dimerization of TuSp NT, not involving protonation of D77 (and R117). In the vicinity of TuSp NT E82 is another, non-conserved glutamate residue E37, which may be important for elevating the pK_a value of E82. Two other acidic residues, E85 and D130 from the same subunit, are close in space and pointing towards each other in some of the NMR conformers (Figure 3C). Thus, they could potentially be involved in the pH sensitivity via stabilization of a dimerization-capable conformation upon protonation. The TuSp NT dimer is further stabilized by the conserved D40-R63 salt bridge (D40-K65 in MaSp NT). The F10 residue is buried in a pocket between the helices and in contrast to MaSp NT W10 is not solvent exposed in the dimer but assumes a conformation more like that in MaSp NT monomer (Figures 3E,F).

Comparison of the assigned chemical shifts of the TuSp NT dimer and TuSp NT_{A70E} (Figure 2G) showed that the largest differences are clustered in helix α 2, C-terminal part of helix α 3, the loop between α 3 and α 4 as well as helix α 5. As expected, all these regions are located close to the dimer subunit interface. The acidic amino acid residues that show largest chemical shift changes are E37, D40, D42, and E82.

Dimerization Analysis

Unlike MaSp NT, the TuSp NT sequence does not contain the W10 residue, which prevents monitoring of the monomer-to-dimer transition by tryptophan fluorescence (Kronqvist et al., 2014; Otikovs et al., 2015). Hence, we used HSQC NMR spectroscopy, size exclusion chromatography (SEC) and circular dichroism (CD) to evaluate the dimerization behavior of TuSp NT and three mutants, in which the potentially titratable glutamate/aspartate residues were replaced by the respective



amides. Based on the structure analysis of the TuSp NT dimer and chemical shift comparison between TuSp NT_{A70E} and TuSp NT dimer, a double, a triple and a quadruple mutant was designed, and the mutations were introduced by site-directed mutagenesis (mutants TuSp NT_{E37QE82Q}, TuSp NT_{E37QE82QE85Q} and TuSp NT_{E37QE82QE85QD130N}). The monomeric mutants TuSp NT_{A70E} and TuSp NT_{D40RR63D} were also included as controls.

Comparison of HSQC spectra of TuSp NT at pH 7.2 and 5.5 with the corresponding spectra of the three mutants indicated high structural similarity at pH 5.5 (corresponding to dimeric state) but less similar conformations at pH 7.2 (**Supplementary Figure S4**). However, all the mutants retained pH responsiveness as the spectra at pH 7.2 and 5.5 differed significantly. This result

suggests that some other amino acid residue than the mutated ones or some of the investigated residues but in different combinations are protonated in this pH interval.

SEC was performed in presence of 300 mM NaCl at pH 8 and 20 mM NaCl at pH 5.5. The experiments with TuSp NT showed that it migrates with an apparent molecular weight of 13.1 kDa at pH 8 and of 26.3 kDa at pH 5.5 (**Figure 4**). These values agree well with the calculated molecular weights of monomers (14.1 kDa) and dimers (28.1 kDa). The apparent molecular weight of TuSp NT_{A70E} at pH 8 was 11.0 kDa. As its HSQC spectrum showed no signs of partial unfolding, the reduced apparent molecular weight could suggest that this mutation indeed stabilizes the monomeric state. However, at low pH TuSp NT_{A70E} eluted with an apparent

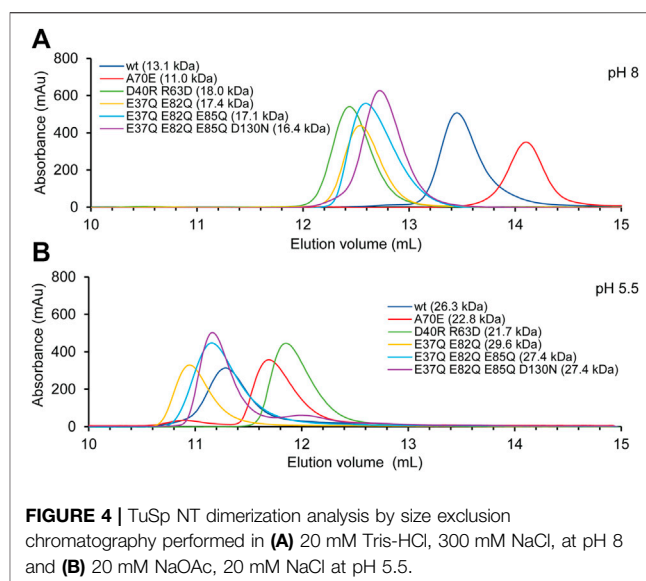
TABLE 1 | Input for the structure calculation and structural statistics for the energy-minimized NMR structure of TuSp NT at pH 5.5.

Distance constraints	
Total NOE	3663
Intra-residue ($ i-j = 0$)	1010
Inter-residue	2653
Sequential ($ i-j = 1$)	1077
Medium-range ($ i-j < 4$)	760
Long-range ($ i-j > 5$)	732
Intermolecular	84
Violations (mean and s.d.)	
Distance constraints (Å)	0.0103 ± 0.0010
Max. distance constraint violation (Å)	0.37 ± 0.17
Number ≥ 0.1 Å	31 ± 5
PARALLHDG force field energies (kcal/mol)	
Total	-9987 ± 96
van der Waals	-1106 ± 33
Electrostatic	-10597 ± 102
Average pairwise r.m.s.d.* (Å)	
Heavy atoms (residues 7–55, 60–127)	0.92 ± 0.06
Backbone (residues 7–55, 60–127)	0.64 ± 0.07
Ramachandran plot statistics (%)	
Most favored regions	89.0
Additional allowed regions	10.0
Generously allowed regions	0.5
Disallowed regions	0.5 ^a

^aThe residues found in the disallowed regions are A84 (in 6 conformers), E85 (10 conformers), D130 (2 conformers), N132 (1 conformer).

molecular weight of 22.8 kDa indicating that it can still dimerize. TuSp NT_{D40RR63D} eluted with apparent molecular weights of 18.0 (pH 8) and 21.7 kDa (pH 5.5) suggesting an intermediate state at both pHs, which is in line with presence of both species and with data from HSQC NMR experiments. The three acidic residue mutants TuSp NT_{E37QE82Q}, TuSp NT_{E37QE82QE85Q}, TuSp NT_{E37QE82QE85QD130N} displayed elution volumes consistent with an intermediate state at pH 8 and dimer state at pH 5.5. Thus, all three mutants kept responsiveness to pH suggesting that we failed to create a constitutive dimer.

CD experiments were carried out to evaluate the refolding capacity and thermal stability of the different TuSp NT variants at pH 8 and pH 5.5. At 25°C the CD spectra of all analyzed proteins showed a maximum at 190 nm and double minima at 208 and 222 nm irrespective of the sample pH, consistent with α -helical secondary structure. Upon heat denaturation at 95°C the protein structures, while retaining some helical structure, mostly converted to random coil as indicated by a broad minimum around 203 nm. After lowering the temperature back to 25°C all variants were able to refold to a similar, mostly α -helical structure as before heat denaturation (Supplementary Figure S5). Thermal stability was estimated from heat denaturation curves and the melting temperatures (T_m) were determined from the half-denaturation points between the native and unfolded states. Dimerization of MaSp NT was shown to increase protein stability through neutralization of intramolecular repelling charge clusters (Kronqvist et al., 2014), which either become protonated or form intermolecular salt bridges. Despite the same number of charged residues (11), the thermal stability of TuSp NT was higher (T_m 62°C and 75°C at pH 8 and pH 5.5,

**FIGURE 4 |** TuSp NT dimerization analysis by size exclusion chromatography performed in (A) 20 mM Tris-HCl, 300 mM NaCl, at pH 8 and (B) 20 mM NaOAc, 20 mM NaCl at pH 5.5.

respectively) than reported for wt *E. australis* MaSp NT [T_m 54°C and 65°C (Kronqvist et al., 2014)] (Figure 5), which is probably due to the presence of a disulfide bond in TuSp NT. TuSp NT_{A70E} showed reduced stability at pH 5.5 (T_m 61°C and 67°C at pH 8 and pH 5.5, respectively), in agreement with destabilization of the dimer conformation. TuSp NT_{D40RR63D} showed high stability at both pHs (T_m 73°C and 74°C at pH 8 and pH 5.5, respectively), in line with intramolecular salt bridge formation. TuSp NT_{E37QE82Q} showed increased stability at pH 8 (T_m 70°C and 74°C at pH 8 and pH 5.5), which confirms elimination of a destabilizing interaction. TuSp NT_{E37QE82QE85Q} and TuSp NT_{E37QE82QE85QD130N} variants showed similar stability at both pHs (T_m 69–70°C and T_m 70–72°C, respectively), and at the same time were more stable than TuSp NT at pH 8, but less stable at pH 5.5. The decreased stability at pH 5.5 suggests neutralization of residues not involved in the protonation.

DISCUSSION

We performed structural studies and dimerization analysis of TuSp NT from *A. argentata* using NMR, SEC, and CD experiments, which showed that the pH-dependent stable dimer formation is conserved in tubuliform silk. However, TuSp NT could not be stabilized in the monomeric conformation for NMR structure determination regardless of the environmental conditions. Furthermore, introduction of mutations in TuSp NT_{K40DD63R} and TuSp NT_{A70E}, corresponding to monomeric and pH insensitive MaSp and FlSp NT variants (Jaudzems et al., 2012; Kronqvist et al., 2017; Sarr et al., 2022), failed to fully stabilize TuSp NT in the monomeric conformation. In order to identify residues that could potentially participate in the pH dependent dimerization mechanism, we determined the solution structure of TuSp NT dimer at pH 5.5. Guided by the insights from the structure, we

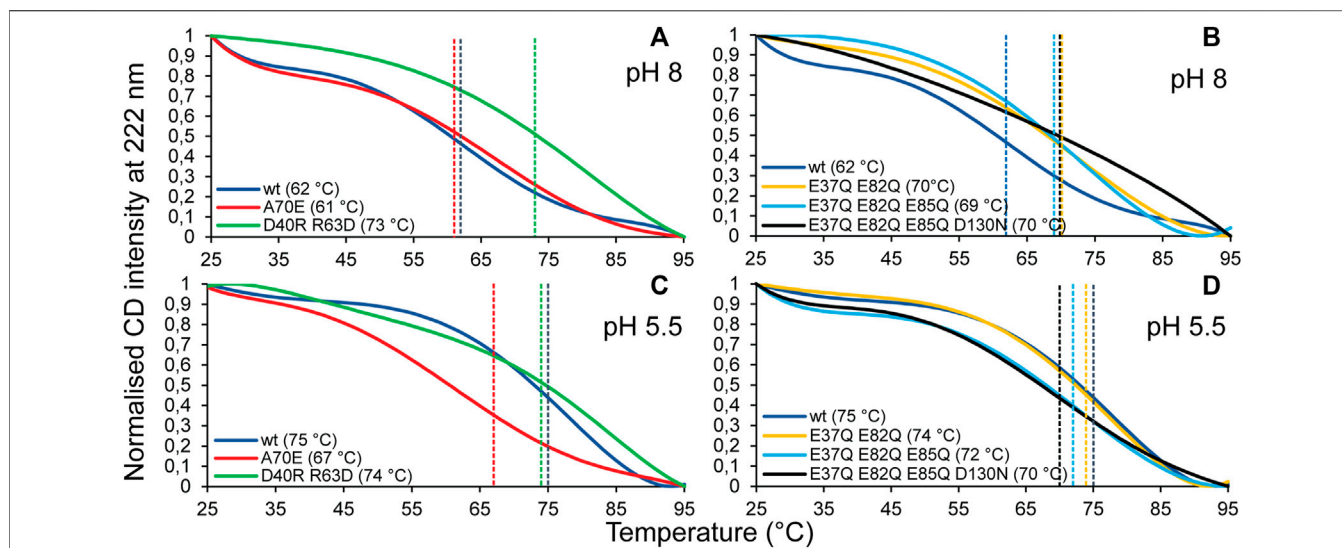


FIGURE 5 | CD melting temperature (T_m) scans of TuSp NT in 20 mM NaPi, pH 8 (A,B) and in 20 mM NaPi, pH 5.5 (C,D). The CD signal intensity at 222 nm was monitored between 25 and 95°C with 1°C/min increase.

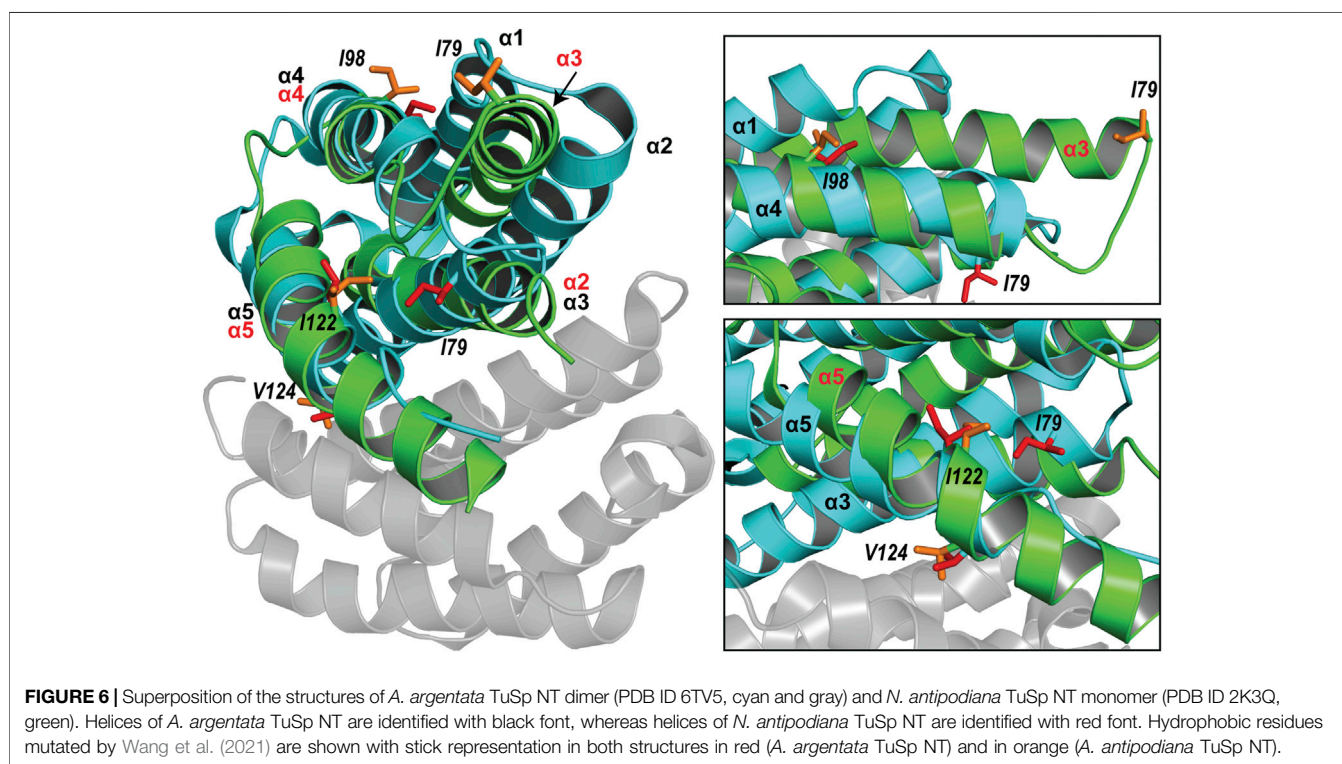


FIGURE 6 | Superposition of the structures of *A. argentata* TuSp NT dimer (PDB ID 6TV5, cyan and gray) and *N. antipodiana* TuSp NT monomer (PDB ID 2K3Q, green). Helices of *A. argentata* TuSp NT are identified with black font, whereas helices of *N. antipodiana* TuSp NT are identified with red font. Hydrophobic residues mutated by Wang et al. (2021) are shown with stick representation in both structures in red (*A. argentata* TuSp NT) and in orange (*A. antipodiana* TuSp NT).

designed the mutants TuSp NT_{E37QE82Q}, TuSp NT_{E37QE82QE85Q}, and TuSp_{E37QE82QE85QD130N}, in which several clustered and surface-exposed acidic residues were exchanged for their amide counterparts. However, none of them behaved as a constitutive dimer and retained pH sensitivity. SEC and HSQC NMR data showed that the variants assume an intermediate conformation (i.e., with at least one of the titrating

carboxylates neutralized) at high pH and are converted into dimeric conformation at pH 5.5. CD experiments showed that thermal stability of the mutants is higher at pH 8, however, not approaching the stability of TuSp NT at pH 5.5.

The spidroin NT domains from most spider silk types and species have been shown to adopt a monomeric or dimeric five-helix bundle structure depending on environmental conditions.

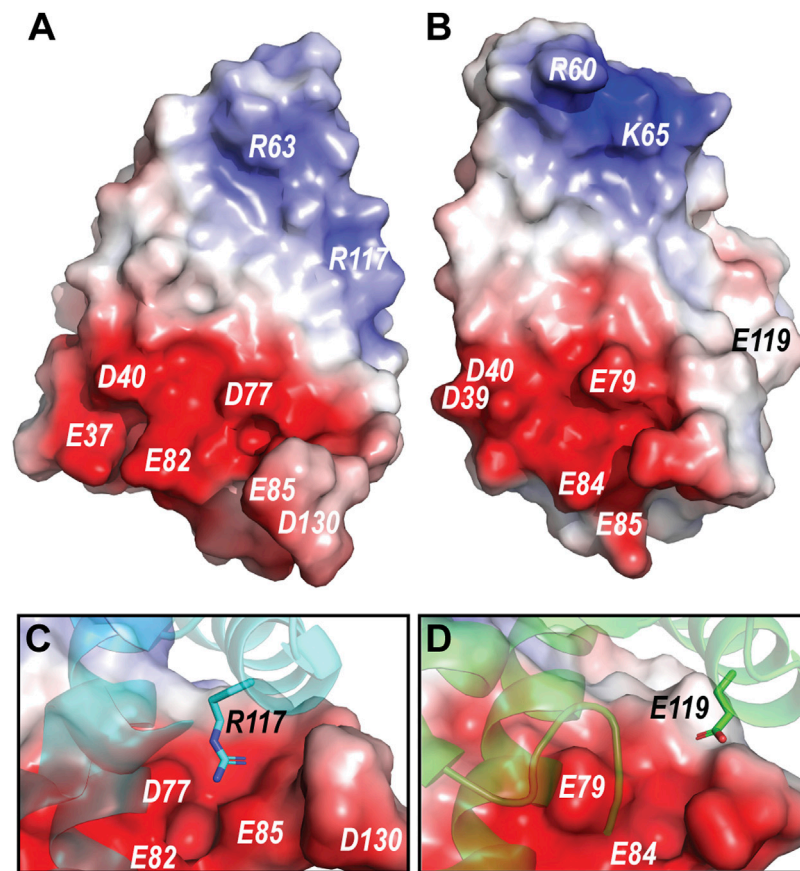


FIGURE 7 | (A) Electrostatic potential of *A. argentata* TuSp NT dimer subunit. **(B)** Electrostatic potential of *E. australis* MaSp NT (PDB ID 2LTH) dimer subunit. Red color indicates negative charges, and blue color indicates positive charges. The surface-exposed charged residues are labeled. **(C)** and **(D)** shows contact surfaces of the residues R117 and E119 from the opposite subunit at the TuSp NT and MaSp NT dimer interfaces, respectively.

However, the previously determined TuSp NT structure from *N. antipodiana* shows an atypical four-helix arrangement that is distinct from the five-helix NT structures, raising questions about the fold conservation. Despite high sequence similarity (66% identity), our herein determined TuSp NT structure from *A. argentata* reveals a five-helix composition, which is highly similar to the structures of MaSp, MiSp, AcSp, and FlSp NTs. Superposition of the two TuSp NT structures shows a different organization of the first three helices (**Figure 6**). The helix $\alpha 1$ in *A. argentata* TuSp NT is replaced by $\alpha 3$ in the *N. antipodiana* structure, $\alpha 3$ is substituted by $\alpha 2$, while helix $\alpha 2$ of the *A. argentata* protein is missing in the *N. antipodiana* structure. Knowing that the first 36 residues of the *N. antipodiana* protein were truncated for the structure determination, the differences are apparently due to a helical reorganization aimed to preserve its hydrophobic core, which is mainly formed by the helices $\alpha 1$ and $\alpha 3$ - $\alpha 5$. Besides, the solution conditions used for each structure determination are vastly different—the *N. antipodiana* TuSp NT was studied at neutral pH (50 mM Tris-HCl, pH 7) in presence of 100 mM DPC, whereas we employed

acidic conditions (20 mM sodium acetate, 20 mM NaCl, pH 5.5). Accordingly, the *N. antipodiana* structure at pH 7 shows a monomer, whereas the *A. argentata* TuSp NT at pH 5.5 is a dimer. The use of a shorter construct and presence of 100 mM DPC, which according to the authors was necessary to avoid aggregation, raises concerns about the relevance of the *N. antipodiana* structure. Our study shows that the structured part of TuSp NT begins already at residue 10, both for the wt protein at pH 5.5 (**Figure 3**) as well as the A70E mutant at pH 8.0 (**Supplementary Figure S2**). Hence, the complete first α -helix (residues S12-I27) and beginning of the second α -helix (residues P31-Q36) is lacking in the *N. antipodiana* TuSp NT structure. Several hydrophobic residues as I80 and I98 showing surface localization in the *N. antipodiana* structure, which have been suggested to play a role in the intermolecular association of TuSp NT (Wang et al., 2021) are buried in the hydrophobic core in our structure (**Figure 6**). Additionally, replacement of the $\alpha 1$ - $\alpha 4$ helical interface by $\alpha 3$ - $\alpha 4$ as in the *N. antipodiana* structure would not be possible in a full-length construct, because C22 from $\alpha 1$ forms a disulfide bond with C102 from $\alpha 4$, placing these

helices next to each other. Altogether, the newly obtained results indicate that, under physiological conditions, full-length TuSp NT adopts a five-helix bundle structure as the other NTs.

The pH dependent dimerization of MaSp and MiSp NT was found to be mediated by sequential protonation of three specific glutamic acid residues (E79, E84, E119 in MaSp and E73, E76, E115 in MiSp NT). In the *A. argentata* TuSp NT, only E84 (E82 in TuSp NT) is structurally preserved, whereas E79 and E119 are replaced by D77 and R117, respectively, and form an intermolecular salt bridge in the dimer structure. These substitutions alter the electrostatic surface potential of each monomer subunit, making the positively charged pole more extensive (i.e. spread out) than in MaSp NT (Figure 7), which could facilitate intermolecular association. Additionally, the charge repulsion between two glutamate residues at the dimer interface is abolished potentially allowing the subunits to associate before the protonation events. This could explain why we were unable to stabilize the monomeric conformation of TuSp NT for NMR structure determination. Besides these charge interactions, pre-arrangement of α -helices has been suggested to be important for the NT dimerization. In MaSp NT the helical reorganization is facilitated by swinging out of the wedged W10 residue and subsequent repacking of the hydrophobic core. In TuSp NT W10 is replaced by F10, which is buried in the hydrophobic core of the dimer structure. This is similar to FlSp NT, for which the monomer and dimer structures showed the same buried F10 side chain orientation (Sarr et al., 2022). Thus, the structural reorganization during TuSp NT dimerization does not seem to require relocation of F10 side chain. Heiby et al. (2019) reported that the unusually high content of methionines within the core region of *E. australis* MaSp NT plays an important role in the monomer-dimer structural transition. Exchange of the core methionines with the bulkier leucines made the monomeric MaSp NT more rigid, which abolished the movement of its helices and ability to form a dimer. The lack of the tryptophan and methionine residues in TuSp NT sequence (Figure 1) could abolish its hydrophobic core plasticity, locking the core in a dimer-like conformation also at neutral pH, albeit with different dynamics. Instead, the monomer-dimer structural rearrangement likely involves slight movement of helices due to neutralization (protonation) of repulsing charges across intramolecular helical interfaces.

Since E82 shows a similar side chain orientation to E84 in MaSp NT and additionally has E37 located in close proximity, we hypothesized that neutralization of these two residues may be important for establishing the characteristic handshake interaction between E82 and D40 as seen in the MaSp NT dimer. Although E37 is not conserved in most other spidroins, it is highly conserved between TuSp NTs from different species (Supplementary Figure S6). However, analysis of the double mutant TuSp NT_{E37QE82Q} showed that some other residue is additionally protonated at low pH. The only other well conserved glutamate residue among TuSp NTs is E85, which in some conformers of our structure is spatially close to D130 of the same subunit and the charge clash could prevent it from

assuming a dimer-compatible conformation. However, preparation and characterization of the triple and quadruple mutants TuSp NT_{E37QE82QE85Q} and TuSp NT_{E37QE82QE85QD130N} gave similar results as for TuSp NT_{E37QE82Q}. This result does not rule out the involvement of at least some of the mutated residues in the observed pH sensitivity, because the studied combinations may bear disruptive mutations affecting the pK_a of other residues that participate in the protonation events of the wt protein.

In summary, our obtained results clearly indicate that the mechanism of pH-dependent dimerization is different for TuSp NT than it is for MaSp and MiSp NT. We show that in contrast to previous findings TuSp NT has the same five-helix fold of other NTs and forms a stable dimer at low pH. However, its unique amino acid sequence does not allow full stabilization of the monomer conformation at the conditions, where MaSp NT forms stable monomers (pH 7.2 in presence of 300 mM NaCl). This may be linked to the lack of a well-defined spidroin storage sac as well as different physiological conditions (especially, pH gradient) in the tubuliform glands, which remain to be characterized. Furthermore, the pH dependent stable dimer formation that takes place in the silk duct upon fiber formation involves a very different set of amino acid residues as compared to MaSp and MiSp NT. Further research is needed to identify the exact amino acid residues that become protonated in the TuSp NT dimer at low pH.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: <http://www.wwpdb.org/>, 6TV5 <https://bmr.io/>, 34473.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

This work was funded by the European Regional Development Fund (agreement no 1.1.1.1/18/A/004). JF was supported by a doctoral scholarship from the Latvian Institute of Organic Synthesis.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmolb.2022.936887/full#supplementary-material>

REFERENCES

- Abelein, A., Chen, G., Kitoka, K., Aleksis, R., Oleskovs, F., Sarr, M., et al. (2020). High-yield Production of Amyloid- β Peptide Enabled by a Customized Spider Silk Domain. *Sci. Rep.* 10, 235. doi:10.1038/s41598-019-57143-x
- Andersson, M., Chen, G., Otkovs, M., Landreh, M., Nordling, K., Kronqvist, N., et al. (2014). Carbonic Anhydrase Generates CO₂ and H⁺ that Drive Spider Silk Formation via Opposite Effects on the Terminal Domains. *PLoS Biol.* 12, e1001921. doi:10.1371/journal.pbio.1001921
- Andersson, M., Holm, L., Ridderstråle, Y., Johansson, J., and Rising, A. (2013). Morphology and Composition of the Spider Major Ampullate Gland and Dragline Silk. *Biomacromolecules* 14, 2945–2952. doi:10.1021/bm400898t
- Askarieh, G., Hedhammar, M., Nordling, K., Saenz, A., Casals, C., Rising, A., et al. (2010). Self-assembly of Spider Silk Proteins Is Controlled by a pH-Sensitive Relay. *Nature* 465, 236–238. doi:10.1038/nature08962
- Ayoub, N. A., Garb, J. E., Tinghitella, R. M., Collin, M. A., and Hayashi, C. Y. (2007). Blueprint for a High-Performance Biomaterial: Full-Length Spider Dragline Silk Genes. *PLoS One* 2, e514. doi:10.1371/journal.pone.0000514
- Bauer, J., Schaal, D., Eisoldt, L., Schweimer, K., Schwarzinger, S., and Scheibel, T. (2016). Acidic Residues Control the Dimerization of the N-Terminal Domain of Black Widow Spiders' Major Ampullate Spidroin 1. *Sci. Rep.* 6, 34442. doi:10.1038/srep34442
- Brunger, A. T. (2007). Version 1.2 of the Crystallography and NMR System. *Nat. Protoc.* 2, 2728–2733. doi:10.1038/nprot.2007.406
- Chakraborty, R., Fan, J.-s., Lai, C. C., Raghuvamsi, P. V., Chee, P. X., Anand, G. S., et al. (2020). Structural Basis of Oligomerization of N-Terminal Domain of Spider Aciniform Silk Protein. *Ijms* 21, 4466. doi:10.3390/ijms21124466
- Chaw, R. C., Collin, M., Wimmer, M., Helmrick, K.-L., and Hayashi, C. Y. (2018). Egg Case Silk Gene Sequences from Argiope Spiders: Evidence for Multiple Loci and a Loss of Function between Paralogs. *G3 Bethesda* 8, 231–238. doi:10.1534/g3.117.300283
- Chaw, R. C., and Hayashi, C. Y. (2018). Dissection of Silk Glands in the Western Black widow *Latrodectus Hesperus*. *J. Arachnology* 46, 159–161. doi:10.1636/JoA-16-S-063.1
- Colgin, M. A., and Lewis, R. V. (1998). Spider Minor Ampullate Silk Proteins Contain New Repetitive Sequences and Highly Conserved Non-silk-like "spacer Regions". *Protein Sci.* 7, 667–672. doi:10.1002/pro.5560070315
- Collin, M. A., Clarke, T. H., Ayoub, N. A., and Hayashi, C. Y. (2016). Evidence from Multiple Species that Spider Silk Glue Component ASG2 Is a Spidroin. *Sci. Rep.* 6, 21589. doi:10.1038/srep21589
- d'Auvergne, E. J., and Gooley, P. R. (2008). Optimisation of NMR Dynamic Models II: A New Methodology for the Dual Optimisation of the Model-free Parameters and the Brownian Rotational Diffusion Tensor. *J. Biomol. NMR* 40, 121–133. doi:10.1007/s10858-007-9213-3
- Eisoldt, L., Smith, A., and Scheibel, T. (2011). Decoding the Secrets of Spider Silk. *Mater. Today* 14, 80–86. doi:10.1016/S1369-7021(11)70057-8
- Gaines, W. A., Sehorn, M. G., and Marcotte, W. R. (2010). Spidroin N-Terminal Domain Promotes a pH-dependent Association of Silk Proteins during Self-Assembly. *J. Biol. Chem.* 285, 40745–40753. doi:10.1074/jbc.M110.163121
- Gosline, J. M., DeMont, M. E., and Denny, M. W. (1986). The Structure and Properties of Spider Silk. *Endeavour* 10, 37–43. doi:10.1016/0160-9327(86)90049-9
- Gosline, J. M., Guerette, P. A., Ortlepp, C. S., and Savage, K. N. (1999). The Mechanical Design of Spider Silks: from Fibroin Sequence to Mechanical Function. *J. Exp. Biol.* 202, 3295–3303. doi:10.1242/jeb.202.23.3295
- Güntert, P., Mumenthaler, C., and Wüthrich, K. (1997). Torsion Angle Dynamics for NMR Structure Calculation with the New Program DYANA. *J. Mol. Biol.* 273, 283–298. doi:10.1006/jmbi.1997.1284
- Hagn, F., Eisoldt, L., Hardy, J. G., Vendrely, C., Coles, M., Scheibel, T., et al. (2010). A Conserved Spider Silk Domain Acts as a Molecular Switch that Controls Fibre Assembly. *Nature* 465, 239–242. doi:10.1038/nature08936
- Heiby, J. C., Goretzki, B., Johnson, C. M., Hellmich, U. A., and Neuweiler, H. (2019). Methionine in a Protein Hydrophobic Core Drives Tight Interactions Required for Assembly of Spider Silk. *Nat. Commun.* 10, 4378. doi:10.1038/s41467-019-12365-5
- Heiby, J. C., Rajab, S., Rat, C., Johnson, C. M., and Neuweiler, H. (2017). Conservation of Folding and Association within a Family of Spidroin N-Terminal Domains. *Sci. Rep.* 7, 16789. doi:10.1038/s41598-017-16881-6
- Herrmann, T., Güntert, P., and Wüthrich, K. (2002b). Protein NMR Structure Determination with Automated NOE-Identification in the NOESY Spectra Using the New Software ATNOS. *J. Biomol. NMR* 24, 171–189. doi:10.1023/a:1021614115432
- Herrmann, T., Güntert, P., and Wüthrich, K. (2002a). Protein NMR Structure Determination with Automated NOE Assignment Using the New Software CANDID and the Torsion Angle Dynamics Algorithm DYANA. *J. Mol. Biol.* 319, 209–227. doi:10.1016/s0022-2836(02)00241-3
- Hijirida, D. H., Do, K. G., Michal, C., Wong, S., Zax, D., and Jelinski, L. W. (1996). ¹³C NMR of *Nephila clavipes* Major Ampullate Silk Gland. *Biophysical J.* 71, 3442–3447. doi:10.1016/S0006-3495(96)79539-5
- Jaudzems, K., Askarieh, G., Landreh, M., Nordling, K., Hedhammar, M., Jörnvall, H., et al. (2012). pH-Dependent Dimerization of Spider Silk N-Terminal Domain Requires Relocation of a Wedged Tryptophan Side Chain. *J. Mol. Biol.* 422, 477–487. doi:10.1016/j.jmb.2012.06.004
- Kay, L. E., Torchia, D. A., and Bax, A. (1989). Backbone Dynamics of Proteins as Studied by Nitrogen-15 Inverse Detected Heteronuclear NMR Spectroscopy: Application to Staphylococcal Nuclease. *Biochemistry* 28, 8972–8979. doi:10.1021/bi00449a003
- Keller, R. (2004). *The Computer Aided Resonance Assignment Tutorial*. Goldau: Cantina Verlag.
- Knight, D., and Vollrath, F. (2001). Changes in Element Composition along the Spinning Duct in a *Nephila* Spider. *Naturwissenschaften* 88, 179–182. doi:10.1007/s001140100220
- Kronqvist, N., Otkovs, M., Chmyrov, V., Chen, G., Andersson, M., Nordling, K., et al. (2014). Sequential pH-Driven Dimerization and Stabilization of the N-Terminal Domain Enables Rapid Spider Silk Formation. *Nat. Commun.* 5, 3254. doi:10.1038/ncomms4254
- Kronqvist, N., Rising, A., and Johansson, J. (2022). A Novel Approach for the Production of Aggregation-Prone Proteins Using the Spidroin-Derived NT* Tag. *Methods Mol. Biol. Clifton N. J.* 2406, 113–130. doi:10.1007/978-1-0716-1859-2_6
- Kronqvist, N., Sarr, M., Lindqvist, A., Nordling, K., Otkovs, M., Venturi, L., et al. (2017). Efficient Protein Production Inspired by How Spiders Make Silk. *Nat. Commun.* 8, 15504. doi:10.1038/ncomms15504
- Landreh, M., Askarieh, G., Nordling, K., Hedhammar, M., Rising, A., Casals, C., et al. (2010). A pH-dependent Dimer Lock in Spider Silk Protein. *J. Mol. Biol.* 404, 328–336. doi:10.1016/j.jmb.2010.09.054
- Lin, Z., Huang, W., Zhang, J., Fan, J.-S., and Yang, D. (2009). Solution Structure of Eggcase Silk Protein and its Implications for Silk Fiber Formation. *Proc. Natl. Acad. Sci. U.S.A.* 106, 8906–8911. doi:10.1073/pnas.0813255106
- Opell, B. D., and Hendricks, M. L. (2010). The Role of Granules within Viscous Capture Threads of Orb-Weaving Spiders. *J. Exp. Biol.* 213, 339–346. doi:10.1242/jeb.036947
- Otkovs, M., Chen, G., Nordling, K., Landreh, M., Meng, Q., Jörnvall, H., et al. (2015). Diversified Structural Basis of a Conserved Molecular Mechanism for pH-dependent Dimerization in Spider Silk N-Terminal Domains. *ChemBioChem* 16, 1720–1724. doi:10.1002/cbic.201500263
- Perry, D. J., Bittencourt, D., Siltberg-Liberles, J., Rech, E. L., and Lewis, R. V. (2010). Piriform Spider Silk Sequences Reveal Unique Repetitive Elements. *Biomacromolecules* 11, 3000–3006. doi:10.1021/bm1007585
- Ries, J., Schwarze, S., Johnson, C. M., and Neuweiler, H. (2014). Microsecond Folding and Domain Motions of a Spider Silk Protein Structural Switch. *J. Am. Chem. Soc.* 136, 17136–17144. doi:10.1021/ja508760a
- Rousseau, M.-E., Lefèvre, T., Beaulieu, L., Asakura, T., and Pézolet, M. (2004). Study of Protein Conformation and Orientation in Silkworm and Spider Silk Fibers Using Raman Microspectroscopy. *Biomacromolecules* 5, 2247–2257. doi:10.1021/bm049717v
- Sarr, M., Kitoka, K., Walsh-White, K.-A., Kaldmäe, M., Metlans, R., Tärs, K., et al. (2022). The Dimerization Mechanism of the N-Terminal Domain of Spider Silk Proteins Is Conserved Despite Extensive Sequence Divergence. *J. Biol. Chem.* 298, 101913. doi:10.1016/j.jbc.2022.101913
- Sarr, M., Kronqvist, N., Chen, G., Aleksis, R., Purhonen, P., Hebert, H., et al. (2018). A Spidroin-derived Solubility Tag Enables Controlled Aggregation of a Designed Amyloid Protein. *FEBS J.* 285, 1873–1885. doi:10.1111/febs.14451
- Schwarze, S., Zwettler, F. U., Johnson, C. M., and Neuweiler, H. (2013). The N-Terminal Domains of Spider Silk Proteins Assemble Ultrafast and Protected from Charge Screening. *Nat. Commun.* 4, 2815. doi:10.1038/ncomms3815
- Serrano, P., Pedrini, B., Mohanty, B., Geralt, M., Herrmann, T., and Wüthrich, K. (2012). The J-UNIO Protocol for Automated Protein Structure Determination by NMR in Solution. *J. Biomol. NMR* 53, 341–354. doi:10.1007/s10858-012-9645-2

- Sparkes, J., and Holland, C. (2017). Analysis of the Pressure Requirements for Silk Spinning Reveals a Pultrusion Dominated Process. *Nat. Commun.* 8, 594. doi:10.1038/s41467-017-00409-7
- Sparkes, J., and Holland, C. (2019). The Energy Requirements for Flow-Induced Solidification of Silk. *Macromol. Biosci.* 19, 1800229. doi:10.1002/mabi.201800229
- Tian, M., and Lewis, R. V. (2005). Molecular Characterization and Evolutionary Study of Spider Tubuliform (Eggcase) Silk Protein. *Biochemistry* 44, 8006–8012. doi:10.1021/bi050366u
- Tremblay, M.-L., Xu, L., Lefèvre, T., Sarker, M., Orrell, K. E., Leclerc, J., et al. (2015). Spider Wrapping Silk Fibre Architecture Arising from its Modular Soluble Protein Precursor. *Sci. Rep.* 5, 11502. doi:10.1038/srep11502
- Wang, J., Yuan, W., Qin, R., Fan, T., Fan, J.-s., Huang, W., et al. (2021). Self-assembly of Tubuliform Spidroins Driven by Hydrophobic Interactions Among Terminal Domains. *Int. J. Biol. Macromol.* 166, 1141–1148. doi:10.1016/j.ijbiomac.2020.10.269

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Šede, Fridmanis, Otikovs, Johansson, Rising, Kronqvist and Jaudzems. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.