

SCIENTIFIC REPORTS



OPEN

Conservation of folding and association within a family of spidroin N-terminal domains

Julia C. Heiby¹, Suhaila Rajab¹, Charlotte Rat¹, Christopher M. Johnson² & Hannes Neuweiler¹

Web spiders synthesize silk fibres, nature's toughest biomaterial, through the controlled assembly of fibroin proteins, so-called spidroins. The highly conserved spidroin N-terminal domain (NTD) is a pH-driven self-assembly device that connects spidroins to super-molecules in fibres. The degree to which forces of self-assembly is conserved across spider glands and species is currently unknown because quantitative measures are missing. Here, we report the comparative investigation of spidroin NTDs originating from the major ampullate glands of the spider species *Euprosthenoops australis*, *Nephila clavipes*, *Latrodectus hesperus*, and *Latrodectus geometricus*. We characterized equilibrium thermodynamics and kinetics of folding and self-association using dynamic light scattering, stopped-flow fluorescence and circular dichroism spectroscopy in combination with thermal and chemical denaturation experiments. We found cooperative two-state folding on a sub-millisecond time scale through a late transition state of all four domains. Stability was compromised by repulsive electrostatic forces originating from clustering of point charges on the NTD surface required for function. pH-driven dimerization proceeded with characteristic fast kinetics yielding high affinities. Results showed that energetics and kinetics of NTD self-assembly are highly conserved across spider species despite the different silk mechanical properties and web geometries they produce.

The question of how a linear chain of amino acids spontaneously folds into a highly ordered three-dimensional structure continues to be a central topic in molecular biology¹. Important insights come from comparative studies of homologous proteins, where differences in folding mechanisms can be traced back to minor sequence changes^{2–4}. Such studies allow the dissection of the roles of sequence and topology in folding. A class of highly evolved proteins that exhibit rather unusual amino acid composition are the silk proteins produced by web spiders. Spider silk is produced by the controlled assembly of spider fibroins, so-called spidroins, within the spinning gland of the animal yielding threads of outstanding mechanical properties tailored for distinct functionalities. Material scientists are trying to decrypt the assembly process and to reproduce it in the laboratory^{5–7}. The bulk of a spidroin sequence consists of repetitive poly-alanine and glycine-rich peptide motifs of simple amino acid composition, which are unstructured under storage conditions in the gland and form mainly β -sheet secondary structure in solid fibres. The repetitive central segments are terminated by the globular folded N- and C-terminal domains (NTD and CTD), which provide water-solubility on the one hand and connectivity in response to mechanical and chemical stimuli on the other^{8,9}. NTD and CTD are five-helix bundles that form homo-dimers. While the CTD is a covalent homo-dimer stabilized by a disulfide linkage, the NTD is monomeric under storage conditions in the gland and undergoes self-association in the spider's spinning duct in response to changes of solution pH and salt composition, thus polymerizing spidroins to form super-molecules^{7,10}. pH-triggered NTD self-association is ultrafast and involves site-specific protonation events and conformational change^{11–15}.

NTD and CTD represent the most conserved sequence areas of spidroins, with no structural homologues identified so far, underscoring their importance in the process of silk formation^{8,10}. Interestingly, the unusual amino acid composition of the central spidroin segments extends into the terminal domains: in the NTD, alanine is the most frequently found amino acid followed by serine; alanine and serine together take up ~30% of an NTD sequence. By contrast, the number of charged side chains is rather low. This is surprising considering the high water-solubility of the domain, which is commonly provided by side chain charges.

¹Department of Biotechnology & Biophysics, Julius-Maximilians-University Würzburg, Am Hubland, 97074, Würzburg, Germany. ²Medical Research Council Laboratory of Molecular Biology, Francis Crick Avenue, Cambridge, CB2 0QH, United Kingdom. Correspondence and requests for materials should be addressed to H.N. (email: hannes.neuweiler@uni-wuerzburg.de)

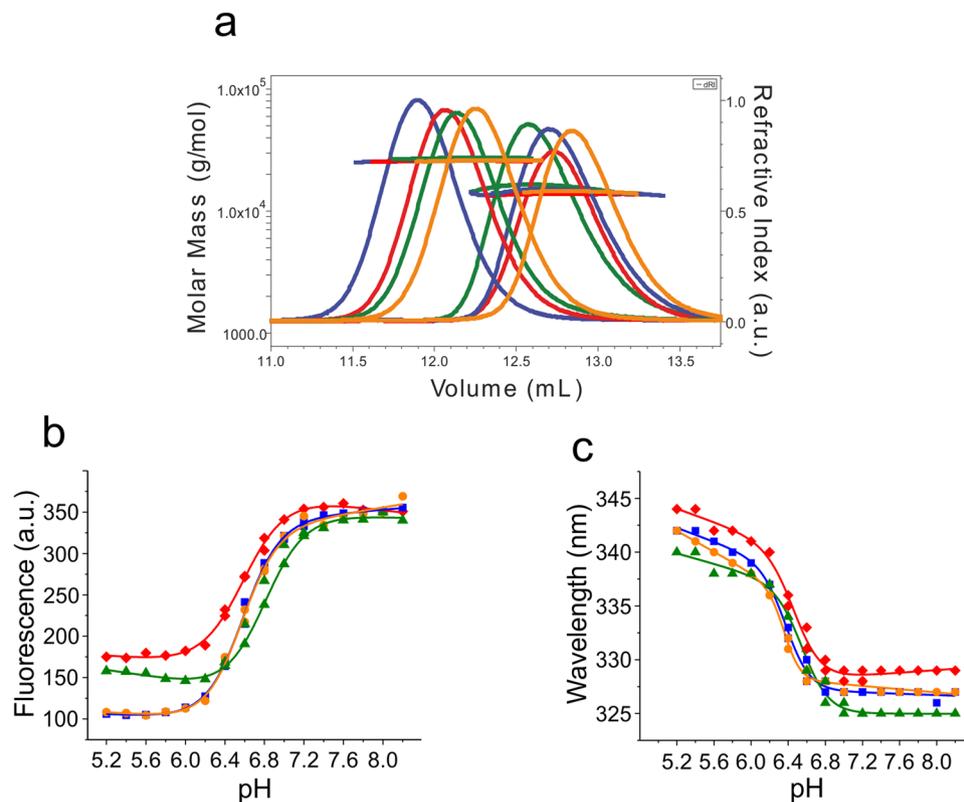


Figure 1. Functionality of spidroin NTDs. (a) SEC-MALS chromatograms of NTDs under monomer and dimer solution conditions showing refractive index against elution volume. The evaluated mass from light scattering analysis is indicated by the horizontal lines. Lh (blue), Lg (orange), Nc (green), and Ea (red). (b) Steady-state tryptophan fluorescence emission intensities of Lh (blue squares), Lg (orange circles), Nc (green triangles), and Ea (red diamonds) NTD recorded under conditions of varying solution pH. Trp fluorescence emission was recorded at 330 nm and normalized to the maximal emission intensity for reasons of clarity. (c) Wavelength of maximal fluorescence emission of the four homologues recorded under conditions of varying solution pH (color code from above applies). Solid lines are guides to the eye.

Unusual amino acid composition and high degree of sequence conservation make spidroin NTDs an interesting system both from the viewpoint of fundamental folding research and material science. Open questions are: does the unusual amino acid composition of NTDs translate into an unusual mechanism of folding? Are energetics and kinetics of folding and self-association conserved? This question is in particular interesting in light of species-dependent differences in strengths and structures of silks¹⁶.

Here, we report the comparative investigation of folding and association of NTDs originating from major ampullate spidroin 1 (MaSp1) of four different spider species, namely the black widow (*Latrodectus hesperus*, Lh) the brown widow (*Latrodectus geometricus*, Lg), the golden orb spider (*Nephila clavipes*, Nc), and the nursery web spider (*Euprosthenoops australis*, Ea). The Ma gland forms the toughest fiber used to build the web frame and a life-line and is a focus of current material science. We found that all four domains folded on a similar sub-millisecond time scale via a conventional two-state mechanism. The Nc homologue, however, was significantly less stable and folded more slowly compared with the other domains. Rate constants of pH-triggered domain self-association and dissociation were similar. Results showed that, despite species-dependent differences in web geometries and silk mechanical properties, the energetics and kinetics of NTD self-assembly are conserved.

Results

Self-association of spidroin NTDs. We produced NTDs of spidroin 1 from the major ampullate gland (MaSp1) of the species Ea, Lh, Lg, and Nc using heterologous overexpression in *E. coli* bacterial cells followed by chromatographic purification. We characterized self-association using analytical size-exclusion chromatography in combination with multi-angle light scattering (MALS) spectroscopy (SEC-MALS). NTDs were investigated under conditions that favor either the monomeric state, i.e. in 50 mM aqueous phosphate buffer pH 7.0 with the solution ionic strength (*I*) adjusted to 200 mM using potassium chloride, or the dimeric state, i.e. in 20 mM (N-morpholino)ethansulfonic acid (MES) pH 6.0 with the *I* adjusted to 60 mM. Molecular mass moments of SEC-MALS experiments confirmed that the domains were either monomers or dimers under these conditions and that they were able to undergo functional dimerization (Fig. 1a).

MaSp1 NTDs contain a single tryptophan (Trp) that changes conformation from buried to solvent-exposed upon dimerization of the native protein^{8,13–15}. Trp is more solvent exposed in the dimer so the fluorescence of this state is lower. We recorded Trp fluorescence emission intensities of the four homologues under solution

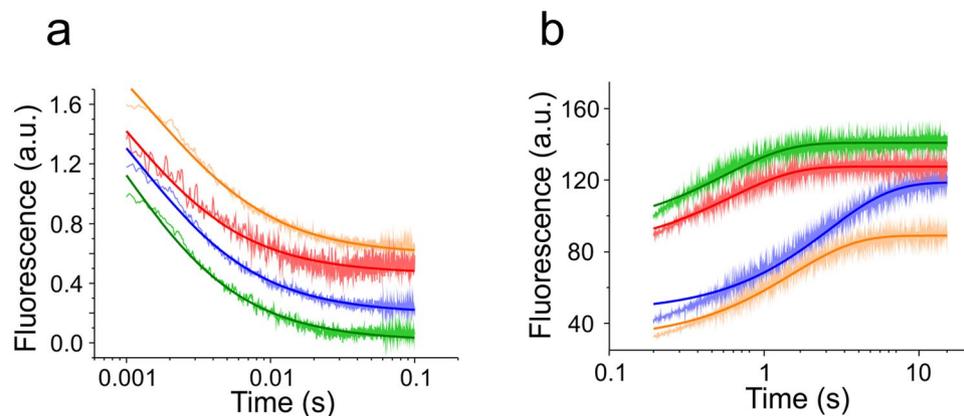


Figure 2. Kinetics of association and dissociation. **(a)** Representative transients of association recorded using stopped-flow Trp fluorescence spectroscopy (Lh: blue, Lg: orange, Nc: green, Ea: red). 500 nM NTD was rapidly mixed from pH 7.0 into pH 6.0 buffered solution. Each trace is an average over 24 single shots. The traces are offset by 0.2 a.u. along the y-axis for reasons of clarity. Dark coloured lines represent data fits using a kinetic model of dimerization. **(b)** Representative transients of NTD dissociation measured by stopped-flow AttoOxa11 fluorescence spectroscopy (Lh: blue, Lg: orange, Nc: green, Ea: red). Transients were recorded by chasing 100 nM fluorescently modified NTD prepared at pH 6.0 with tenfold excess of non-labelled NTD prepared at same solution conditions. Traces are offset along the y-axis for reasons of clarity. Dark coloured lines are mono-exponential data fits.

Construct	k_{ass} ($10^9 \text{ M}^{-1} \cdot \text{s}^{-1}$)	k_{diss} (s^{-1})	K_{d} (10^{-9} M)
<i>L. hesperus</i>	2.19 ± 0.30	0.31 ± 0.02	0.14 ± 0.01
<i>L. geometricus</i>	1.39 ± 0.12	0.64 ± 0.02	0.46 ± 0.03
<i>N. clavipes</i>	2.58 ± 0.20	1.88 ± 0.05	0.73 ± 0.04
<i>E. australis</i>	1.58 ± 0.24	1.68 ± 0.43	1.07 ± 0.11

Table 1. Rate constants of association and dissociation of NTDs.

conditions of varying pH (pH between 5 and 8). We observed strong fluorescence quenching and bathochromic shifts of fluorescence emission maxima at low pH, in agreement with increased solvent exposure of the Trp side chain upon dimerization (Fig. 1b,c). Minor differences in fluorescence amplitudes and wavelengths of maximal emission intensity measured between low and high pH may be explained by minor differences in tertiary packing and degrees of solvent-exposure of the Trp side chains of the different homologues.

We measured kinetics of NTD self-association using rapid mixing experiments. A stopped-flow apparatus was used to rapidly change solution pH and I from monomer to dimer conditions. Kinetics of self-association was followed by decaying Trp fluorescence intensities, which signalled dimer formation (Fig. 2a). We fitted a kinetic model of dimerization to the measured fluorescence intensity transients in order to obtain the microscopic rate constants of self-association, k_{ass} . Association of each homologue was similarly fast, i.e. k_{ass} was on the order of $10^9 \text{ M}^{-1} \text{ s}^{-1}$ (Table 1). Rate constants of dissociation, k_{diss} , were measured using chasing experiments. NTD dimerization can be detected by fluorescence self-quenching of an extrinsic label at sequence position 50, which is located at the rim of the association interface¹¹. We labelled the oxazine fluorophore AttoOxa11 to sequence position 50 of all four homologues using thiol modification of single-point cysteine (Cys) mutants. One labelled subunit of fluorescence self-quenched homo-dimers prepared in pH 6 buffer was chased off by rapidly mixing in an excess of non-labelled protein. Dissociation was thus signalled by an increase of fluorescence emission intensity of the AttoOxa11 label. The fluorescence transients obtained from stopped-flow experiments were fitted using an exponential function, yielding k_{diss} (Fig. 2b). Minor deviations from the exponential fit at fast time scale are likely explained by the presence of small populations of oligomeric species formed under solution conditions that favour fibre formation. Rate constants of dissociation of the four homologues were similar, i.e. k_{diss} was on the order of 1 s^{-1} (Table 1). Using these kinetic results we calculated the equilibrium dissociation constants of dimers, $K_{\text{d}} = k_{\text{diss}}/k_{\text{ass}}$, which were all in the low nanomolar range (Table 1).

Equilibrium denaturation of NTD monomers. To investigate the folding equilibrium of NTDs we performed thermal and chemical denaturation experiments using circular dichroism (CD) and Trp fluorescence spectroscopy. In order to avoid complications in data analysis that would arise from an additional monomer/dimer equilibrium, denaturation was performed in pH 7 buffered solutions where I was adjusted to 200 mM - conditions where the NTDs were monomeric. The CD signal at 222 nm was used as a probe for α -helical secondary structure. Tertiary structure was probed by native Trp fluorescence emission. Thermal denaturation data

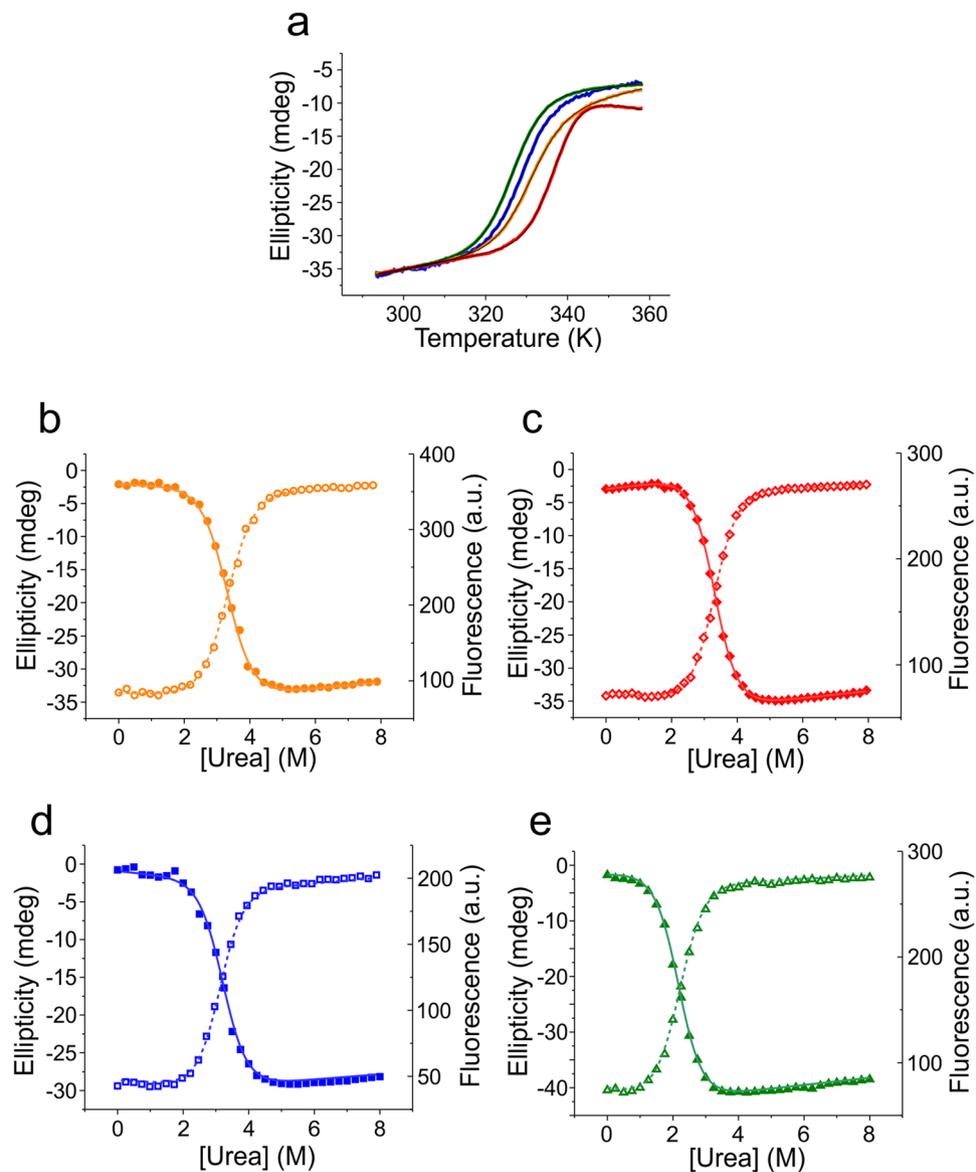


Figure 3. Equilibrium denaturation of NTD monomers. (a) Normalized melting curves of NTDs recorded at pH 7.0 abbreviation for ionic strength: I adjusted to 200 mM using far-UV CD spectroscopy at 222 nm (Ea: red, Lh: blue, Lg: orange, Nc: green). Black lines are data fits using the thermodynamic model for a two-state equilibrium. (b–e) Equilibrium chemical denaturation. Data were recorded by manual titration using urea as denaturant in pH 7.0 buffered solution with I adjusted to 200 mM ((b) Lg, orange circles; (c) Ea, red diamonds; (d) Lh, blue squares; (e) Nc, green triangles). Denaturation was measured using far-UV CD spectroscopy at 222 nm (closed symbols) and Trp fluorescence spectroscopy at 330 nm (open symbols). Solid lines are data fits using the thermodynamic model for a two-state equilibrium.

NTD homologue	<i>L. hesperus</i>	<i>L. geometricus</i>	<i>N. clavipes</i>	<i>E. australis</i>
T_m (K)	329.9 ± 0.1	331.0 ± 0.1	326.8 ± 0.1	337.0 ± 0.1
ΔH_m (kcal/mol)	54.8 ± 0.7	50.1 ± 0.4	54.3 ± 0.2	67.5 ± 0.4

Table 2. Thermal denaturation data.

acquired using CD spectroscopy were well described by a thermodynamic two-state model of folding (Fig. 3a). Fitted enthalpies and melting temperatures are shown in Table 2. Similarity of enthalpies of unfolding indicated similarly strong, non-covalent intramolecular interaction networks that stabilize the fold. The slightly higher enthalpy obtained for Ea may be explained by the observed irreversibility of thermal unfolding of this homologue.

Chemical denaturation was performed using urea as denaturant and monitored using CD and Trp fluorescence spectroscopy. The denaturation transitions obtained by both methods were well described by a two-state model invoking a linear free energy relationship (Fig. 3b–e). For each homologue, the equilibrium m -values and

NTD homologue	Signal	m_{D-N}	$[\text{Urea}]_{50\%}$	ΔG_{D-N}
		($\text{kcal}\cdot\text{M}^{-1}\cdot\text{mol}^{-1}$)	(M)	($\text{kcal}\cdot\text{mol}^{-1}$)
<i>L. hesperus</i>	CD	1.64 ± 0.04	3.09 ± 0.01	5.07 ± 0.14
	Fluo	1.58 ± 0.08	3.20 ± 0.02	5.06 ± 0.29
<i>L. geometricus</i>	CD	1.46 ± 0.05	3.39 ± 0.02	4.95 ± 0.20
	Fluo	1.53 ± 0.06	3.34 ± 0.02	5.10 ± 0.23
<i>N. clavipes</i>	CD	1.59 ± 0.04	2.19 ± 0.02	3.48 ± 0.12
	Fluo	1.67 ± 0.04	2.18 ± 0.01	3.64 ± 0.10
<i>E. australis</i> *	CD	1.69 ± 0.03	3.30 ± 0.01	5.60 ± 0.12
	Fluo	1.69 ± 0.02	3.31 ± 0.01	5.60 ± 0.08

Table 3. Equilibrium chemical denaturation data. *Data from ref.²³.

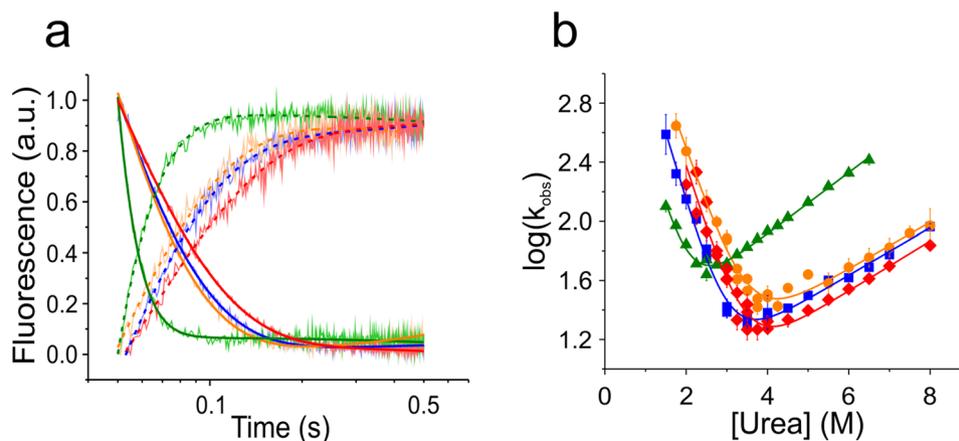


Figure 4. Folding kinetics. **(a)** Representative kinetic transients of folding (decay) and unfolding (rise) measured using stopped-flow Trp fluorescence spectroscopy (Lh: blue, Lg: orange, Nc: green, Ea: red). Each kinetic trace shown in light colour is a single shot. Solid lines and dashed lines are mono-exponential fits to unfolding and folding data, respectively. Data are normalized to 1 for reasons of clarity. Shown unfolding traces were measured at 5 M urea and shown folding traces were measured at 3 M urea. **(b)** Observed relaxation rate constants plotted versus urea concentration (chevron analysis). Solid lines are fits to the data using a kinetic model for a barrier-limited two-state transition. Same colour code as in **(a)** applies.

transition mid-points from CD and Trp fluorescence data were in good agreement (Table 3). Free energies of unfolding were similar, except for the Nc homologue, which had a significantly reduced stability that arose from a lower denaturation mid-point (Table 3).

Folding kinetics of NTD monomers. We measured kinetics of folding using chemical denaturation and stopped-flow Trp fluorescence spectroscopy. NTD samples were prepared in pH 7 buffer at either zero or six molar urea and were rapidly mixed to higher or lower denaturant concentrations to either unfold or refold the protein. Obtained Trp fluorescence intensity time traces were fitted using single-exponential functions containing a linear baseline drift (Fig. 4a). The linear baseline drift occurred on a slow time scale and can be explained by photo-bleaching or sample diffusion in the mixing zone of the stopped-flow machine. For each homologue we plotted the observed rate constants *versus* denaturant concentration and performed chevron analysis (Fig. 4b). Observed rate constants fitted well to a kinetic model for a barrier-limited two-state transition. The extrapolated microscopic rate constants of folding and unfolding of the homologues Ea, Lh, and Lg ranged between $k_f = 10,000\text{--}13,000\text{ s}^{-1}$ and $k_u = 3\text{--}5\text{ s}^{-1}$ (Table 4). The Nc homologue had a slower rate constant of folding and a faster rate constant of unfolding compared with its family members, in agreement with its lower stability found in equilibrium experiments (Tables 3 and 4). Thermodynamic quantities derived from kinetics were in good agreement with values derived from equilibrium data. The sum of kinetic folding and unfolding *m*-values ($m_f + m_u$) compared well with the respective equilibrium *m*-values, m_{eq} . Equilibrium free energies were in reasonable agreement with quantities calculated from k_f and k_u ($\Delta G = -RT \ln(k_u/k_f)$), supporting a two-state model of folding model¹⁷. Tanford β -values ($\beta_T = m_f/(m_f + m_u)$) were all around 0.85 (Table 4) showing that folding proceeded through a similarly compact, native-like transition state.

Influence of solution ionic strength on folding of NTD monomers. To investigate the electrostatic contribution to stability, we studied the influence of *I* on the folding of NTDs. We recorded thermal denaturation data of each homologue using CD spectroscopy in pH 7 buffered solutions with *I* ranging between 0.041 M and 1.00 M. Results showed that the stability of all four homologues increased similarly with increasing *I*, which was

Homologue	k_f	m_f	k_u	m_u	ΔG_{D-N}	β_T
<i>L. hesperus</i>	$10\,123 \pm 2\,278$	1.29 ± 0.06	4.67 ± 0.70	0.22 ± 0.01	4.55 ± 0.16	0.85 ± 0.01
<i>L. geometricus</i>	$11\,899 \pm 3\,035$	1.09 ± 0.06	5.34 ± 1.16	0.22 ± 0.02	4.56 ± 0.20	0.83 ± 0.01
<i>N. clavipes</i>	$3\,652 \pm 1\,328$	1.39 ± 0.13	12.16 ± 1.02	0.28 ± 0.01	3.38 ± 0.22	0.83 ± 0.01
<i>E. australis</i>	$13\,257 \pm 6\,348$	1.20 ± 0.11	3.28 ± 1.14	0.23 ± 0.03	4.92 ± 0.35	0.84 ± 0.02

Table 4. Kinetics of folding. k_f and k_u (s^{-1}); m_f and m_u ($kcal \cdot M^{-1} \cdot mol^{-1}$); ΔG_{D-N} ($kcal/mol$).

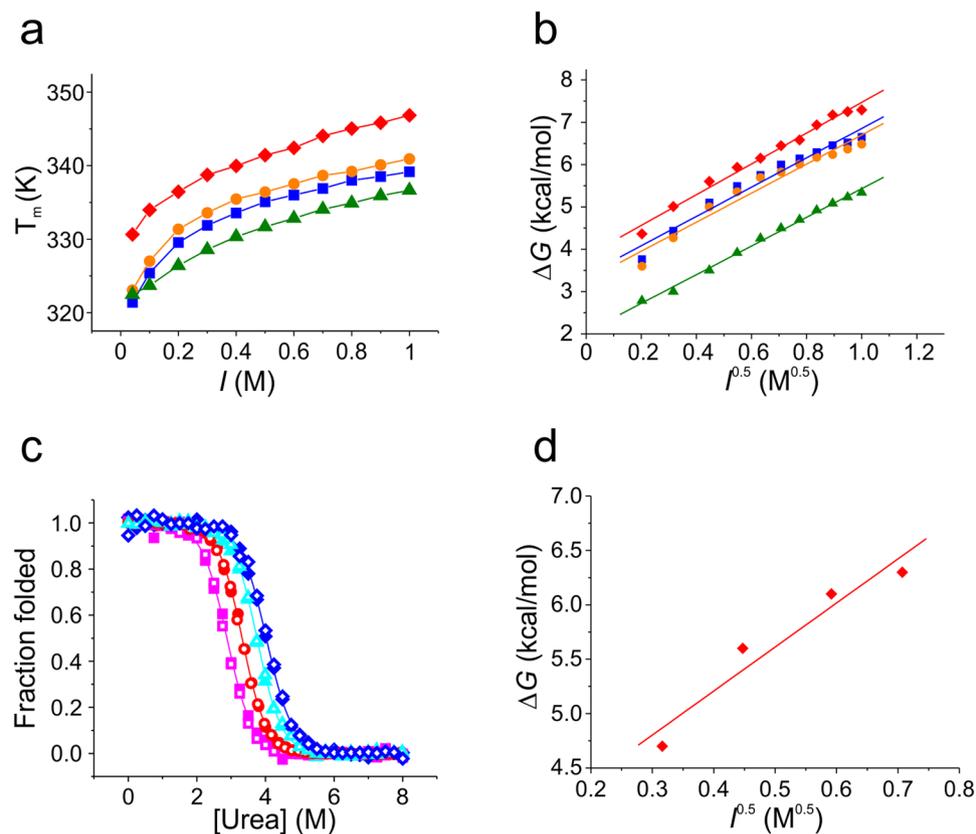


Figure 5. Influence of salt on NTD monomer folding and stability. (a) Dependence of thermal stability on I measured in pH 7.0 buffered solutions using far-UV CD spectroscopy at 222 nm. Thermal denaturation mid-points are plotted versus I (Lh: blue squares, Lg: orange circles, Nc: green triangles, Ea: red diamonds). Solid lines are a guide to the eye. (b) Changes of free energy plotted versus the square root of I . Solid lines are linear data fits. Same colour code as in (a) applies. (c) Chemical denaturation of the Ea NTD measured in pH 7.0 buffered solution using far-UV CD (closed symbols) and Trp fluorescence spectroscopy (open symbols). Data were normalized to the folded fraction. Solid lines are data fits using a thermodynamic model for a two-state equilibrium. Magenta squares: 100 mM I ; red circles: 200 mM I ; cyan triangles: 350 mM I ; blue diamonds: 500 mM I . (d) Plot of the change of free energy versus the square root of I measured using chemical denaturation. The solid line is a linear fit to the data.

evident from the observed increase of thermal denaturation midpoints (Fig. 5a). Findings were in agreement with previous observations made for the Lh homologue¹⁵. We found that the observed increase of stability was the result of Debye-Hückel screening of point charges¹⁸: plots of ΔG versus the square root of I ($I^{0.5}$) were linear for each protein (Fig. 5b). The slope of a linear fit to ΔG plotted versus $I^{0.5}$ can be interpreted as an equilibrium m -value that accounts for the effect if I on stability, m_{eq} ¹⁹, in analogy to the well-known m_{eq} in the classical linear free-energy relationship that applies to denaturants. From this analysis we found that the electrostatic contribution to stability was highly conserved within the NTD family, i.e. m_{eq} of Ea, Nc, Lh, and Lg was 3.6 ± 0.2 , 3.4 ± 0.1 , 3.5 ± 0.2 , and 3.4 ± 0.3 $kcal \cdot mol^{-1} \cdot M^{-0.5}$, respectively. The result was confirmed by chemical denaturation experiments performed under conditions of varying I on the Ea homologue. Denaturation data showed mid-points that increased with I while the classical equilibrium m -value, m_{eq} , remained unchanged (Fig. 5c). A linear fit to the plot of ΔG versus $I^{0.5}$ yielded $m_{eq} = 4.0 \pm 0.8$ $kcal \cdot mol^{-1} \cdot M^{-0.5}$ (Fig. 5d), a value that was within error of the one obtained from thermal denaturation data.

Discussion

Web spiders use up to seven specialized glands to synthesize silk fibres for various tasks including prey capture, reproduction and shelter^{6,7}. The basic principles of synthesis are thought to be conserved across glands and species. Conservation of mechanism is reflected in the conserved, modular sequence architecture of spidroins, which are the building blocks of silk. Yet, sizes, geometries, and mechanical properties of webs built by different spider species vary strongly. These differences may arise from sequence variations in the central, repetitive spidroin segments that make up the bulk of interactions in silk and possibly from different processing conditions in various spinning ducts^{5,6,16}. Little is known about the contribution of sequence modulations in the terminal domains to modulation of silk. Structural studies show that MaSp NTD folds from Ea, Lh, and Nc, as well as from a homologue of the minor ampullate gland of *Araneus ventricosus*^{8,15,20,21} are conserved, although minor differences in helix packing are observed²⁰. Biophysical measurements show similar signatures of pH- and salt-dependent dimerization of the *Euprosthenoops*, *Latrodectus*, *Nephila* and *Araneus* NTDs^{13–15,21}. However, quantitative measures of folding and self-association were missing and are reported here.

Our comparative study of MaSp1 NTDs from Ea, Nc, Lh, and Lg shows similar intensity-loss and red-shift of Trp fluorescence emission upon dimerization (Fig. 1b). Stopped-flow fluorescence experiments showed that the high speed of pH-triggered self-association was similar in all four homologues and responsible for tight binding, with K_d values in the low nM range (Table 1). Structural studies report subtle rearrangement of helices and a consequently altered dimer interface of the Nc homologue²⁰. We found that these structural rearrangements did apparently not translate into a modulation of dimerization kinetics or strength of association. Our results suggest that the MaSp1 NTD evolved as a pH-driven module that connects spidroins in the distal part of the gland by same mechanism and energetics irrespective of species. Findings underscore the importance of species-dependent sequence modulations in the central, repetitive segments for modulation of silk mechanical properties.

The folding of all four homologues appeared cooperative and was well described by two-state transitions - a behaviour that is frequently observed for small, single-domain proteins²². Cooperativity of folding was evident from the good agreement of thermodynamic quantities measured using two different structural probes and from the overall good agreement of quantities derived from equilibrium and kinetic experiments. We found remarkably fast sub-millisecond kinetics of folding of all four homologues (Table 4). Chevron analysis of folding kinetics supported the two-state model with no indications of populated folding intermediates. The kinetic m -value and extrapolated rate constant of folding of the Ea NTD was within error of the values previously estimated from temperature-jump experiments²³. Fast kinetics of folding suggest MaSp1 NTDs as an interesting family for future combined experimental and computational studies that can access overlapping time scales²⁴. Such studies yield atomic-detailed insights into pathways of folding²⁵.

The fold of spidroin NTDs is stabilized by an extensive hydrophobic core (Fig. 6a). Observed similarity of stabilities may be explained by the high degree of conservation of residue side chains that form the core: 15 of 24 core side chains are identical and the remaining 9 are of high similarity (Fig. 6b). This raises the question as to the origin of reduced stability of the Nc homologue compared with its family members (Tables 3 and 4). The subtle rearrangement of helices found in structural studies²⁰ and the resulting change of the tertiary interaction network may explain the finding.

The number of ionisable side chains in NTDs is low compared with the average content of 29% found in proteins²⁶ (Fig. 6). The few side chain charges, however, fulfil critical roles in the mechanism of pH-triggered self-association. Basic and acidic residue side chains cluster on opposing poles of the domain generating a macromolecular dipole that steers anti-parallel orientation in the dimeric assembly⁸. Some specific acidic side chains are involved in the pH-relay mechanism that locks the dimer^{8,11,12,20}. Clustering of point charges, however, compromises stability of the domains through intramolecular, repulsive electrostatic forces. There is thus a trade-off of stability of the monomeric fold versus stability of the dimeric assembly, both modulated by electrostatics. Whilst high concentrations of salt stabilize the monomeric fold, they destabilize the dimeric assembly. Decreasing concentrations of sodium chloride found along the spinning duct towards the tapering end where the fibre is formed²⁷ thus stabilizes the dimer. Salt ions in solution can alleviate intramolecular electrostatic strain either through direct binding or by the effect of Debye-Hückel screening. The latter is explained by the formation of clouds of counter-ions around point charges that screen Coulombic forces. This screening is related to the thickness of the ion cloud, which depends on the square root of I ^{18,19}. We found a linear dependence of the free energy of folding on the square root of I for all NTDs and thus identified Debye-Hückel screening as the mechanism of salt action on this domain. Although the Nc homologue exhibits the highest number of side chain charges (Fig. 6b) additional electrostatic strain appears not to be the origin of the reduced stability compared with the other homologues: m_{eq} of all four domains was similar. The mean m_{eq} of $3.5 \pm 0.1 \text{ kcal mol}^{-1} \text{ M}^{-0.5}$ was higher than the value reported for the small protein FynSH3 ($m_{eq} = 2.90 \pm 0.07 \text{ kcal mol}^{-1} \text{ M}^{-0.5}$)¹⁹. The higher m_{eq} of NTDs likely arises from the highly localized clustering of point charges that steers anti-parallel orientation during dimerization at the expense of a reduced stability of the fold.

In conclusion, our comparative study of folding and association within a family of MaSp1 NTDs reveals kinetics and energetics of self-assembly that are conserved across species. Small size, unusual sequence properties and high speed of folding and binding suggests the NTD family as an interesting system for future combined experimental and computational studies that may elucidate mechanisms of conformational change and self-assembly at atomic detail. Moreover, it will be interesting to see if conservation of folding and association also holds for spidroin NTDs originating from glands other than those used to spin dragline silk.

Methods

Protein synthesis, mutagenesis and fluorescence modification. Synthetic genes of NTDs from MaSp1 of the spider species Lh, Lg, Nc and Ea (GeneArt, Thermo Fisher Scientific) were cloned into a modified

Far-UV CD Spectroscopy. Far-UV CD spectroscopy was performed using a Jasco J-815 spectropolarimeter and a 1 mm path-length cuvette (Hellma) containing 10 μM protein. The signal was recorded at 222 nm probing α -helix secondary structure. Chemical denaturation experiments were conducted under the same solution conditions as described above for steady-state fluorescence experiments. Sample temperature was controlled using a Peltier thermocouple set to 298 K. Thermal denaturation was carried out in 50 mM phosphate buffer, pH 7.0, with I adjusted to 200 mM using potassium chloride and applying a temperature ramp at a rate of 1 K/min. Salt-dependent thermal denaturation data were acquired using 20 mM phosphate buffer, pH 7.0, with varying I adjusted using potassium chloride.

Time-resolved fluorescence experiments. Association and dissociation kinetics of NTDs were measured using stopped-flow fluorescence spectroscopy. Kinetics of association were measured by recording the native Trp fluorescence signals of NTDs on an Applied Photophysics (SX 18MV) machine equipped with a xenon-lamp as excitation source. Monomeric protein samples were prepared at varying concentrations in 10 mM MES pH 7.0 with I adjusted to 150 mM using potassium chloride and rapidly mixed into 20 mM MES buffer, pH 6.0, using a volumetric mixing ratio of 1:1.1. Each homolog was measured at three different concentrations between 0.1 μM and 1 μM . Each measurement was an average of 20 individual shots. Kinetics of dissociation were measured using chasing experiments following the AttoOxa11 fluorescence of modified NTDs on a SFM-2000 BioLogic stopped-flow machine equipped with a 639 nm diode laser as excitation source. In chasing experiments, 0.1 μM dimeric AttoOxa11-modified NTD sample in 20 mM MES pH 6.0 buffer with I adjusted to 60 mM using potassium chloride was chased with 9 μM non-modified NTD prepared in the same buffer. 0.3 mg/mL bovine serum albumin (BSA) and 0.05% Tween-20 were applied as solution additives to suppress glass surface interactions of fluorescently modified protein samples. Six to nine kinetic transients were measured for each homolog and averaged. Kinetics of folding and unfolding of homologs were measured under conditions of chemical denaturation by recording the native Trp fluorescence signals of NTDs using a SFM-2000 BioLogic stopped-flow machine equipped with a 280-nm diode as excitation source. For folding and unfolding kinetics 100 μM protein samples were prepared in buffered solutions containing either zero or six molar urea in 50 mM MOPS, pH 7.0, with I adjusted to 200 mM using potassium chloride. Samples were rapidly mixed into urea solutions of varying concentrations applying a volumetric mixing ratio of 1:10 using the stopped-flow machine. Samples were filtered through 0.2 mm syringe filters before measurement. All stopped-flow measurements were recorded at 298 K. Temperature was adjusted using a circulating water bath.

Data analysis. Equilibrium denaturation data were fitted using the thermodynamic model for a two-state transition between native and denatured states. The spectroscopic signal S can be expressed as a function of denaturant concentration or temperature, here shown as a function of the perturbation P ²⁸:

$$S(P) = \frac{\alpha_N + \beta_N \cdot P + (\alpha_D + \beta_D \cdot P) \cdot \exp(-\Delta G_{D-N}(P)/RT)}{1 + \exp(-\Delta G_{D-N}(P)/RT)} \quad (1)$$

where α_N , β_N , α_D , and β_D are the linearly sloping baselines of native and denatured states, R is the gas constant, T the temperature, and ΔG_{D-N} the difference in free energy between native and denatured state.

In chemical denaturation experiments the folding equilibrium was perturbed by the denaturant urea ($P = [\text{urea}]$). ΔG_{D-N} as a function of chemical denaturant is described by the linear-free energy relationship²⁹:

$$\Delta G_{D-N}([\text{urea}]) = \Delta G_{D-N} - m_{D-N}[\text{urea}] \quad (2)$$

where $[\text{urea}]$ is the concentration of urea and m_{D-N} is the equilibrium m -value that describes the sensitivity of the folding equilibrium to denaturant. Experimental errors of ΔG_{D-N} were determined from propagated errors of fitted values of m_{D-N} and mid-point concentrations of urea ($[\text{urea}]_{50\%}$).

In thermal denaturation experiments, the folding equilibrium is perturbed by heat ($P = T$) and the free energy of unfolding ΔG_{D-N} is described as followed:

$$\Delta G_{D-N}(T) = \Delta H_m \cdot \left(1 - \frac{T}{T_m} \right) - \Delta C_p \left[T_m - T + T \ln \left(\frac{T}{T_m} \right) \right] \quad (3)$$

where ΔH_m is the enthalpy of unfolding at the transition midpoint, T_m the mid-point temperature and ΔC_p is the difference in heat capacity between native and denatured state. For ΔC_p we applied the empirical value of 14 cal K^{-1} mol^{-1} per residue of a generic polypeptide chain³⁰. Changes of stability with varying I was estimated using the Schellman formalism assuming conservation of mid-point entropies³¹:

$$\Delta \Delta G = \frac{\Delta H_m}{T_m} \cdot \Delta T_m \quad (4)$$

where $\Delta \Delta G$ is the change of free energy of unfolding due to the change of I , ΔH is the enthalpy of unfolding, and ΔT_m is the difference in the apparent mid-point temperatures of the denaturation curves.

Kinetic transients of folding/unfolding from stopped flow experiments were fitted to a single exponential function containing a linear baseline drift:

$$S(t) = a \exp(-k_{obs}t) + bt + c \quad (5)$$

$S(t)$ is the fluorescence signal as function of time, a is the amplitude and k_{obs} the observed rate constant of the transient. The parameters b and c describe the linear drift of the baseline. k_{obs} is the sum of the microscopic rate constants for folding and unfolding (k_f and k_u). The change of k_{obs} as a function of denaturant concentration was analyzed by fitting the data to the chevron model for a barrier-limited two-state transition that follows the linear-free-energy relationship¹⁷:

$$\log k_{obs}([urea]) = \log[k_f \exp(-m_{TS-D}[urea]/RT) + k_u \exp(m_{TS-N}[urea]/RT)] \quad (6)$$

m_{TS-D} and m_{TS-N} are the kinetic m -values of folding and unfolding, respectively, where TS denotes the transition state separating denatured and native free energy wells. k_f and k_u are the microscopic rate constants of folding and unfolding, respectively, under standard solution conditions in the absence of denaturant.

Rate constants of self-association of NTDs were obtained from fitting kinetic transients to a reaction model of protein dimerization³²:



where N is the folded, monomeric NTD, N_2 is the dimeric NTD, and k_{ass} is the bimolecular rate constant of association. The differential equation can be solved to give³²:

$$S(t) = S_{t=0} + S \frac{(k_{app}t)}{(1 + k_{app}t)} \quad (8)$$

where $S(t)$ is the time-dependent signal, $S_{t=0}$ is the signal at time $t=0$, S is the signal amplitude change, and k_{app} is the apparent rate constant. The apparent rate constant is related to the association rate constant k_{ass} ³²:

$$k_{app} = c_N \cdot k_{ass} \quad (9)$$

where c_N is the protein concentration in terms of monomer.

Rate constants of dissociation were obtained from fitting kinetic transients of chasing experiments to a mono-exponential rise function.

The equilibrium dissociation constant, K_D , was calculated as:

$$K_D = \frac{k_{diss}}{k_{ass}} \quad (10)$$

Data availability. The datasets generated and analysed during the current study are available from the corresponding author on reasonable request.

References

- Dill, K. A. & MacCallum, J. L. The protein-folding problem, 50 years on. *Science* **338**, 1042–1046 (2012).
- Nickson, A. A., Wensley, B. G. & Clarke, J. Take home lessons from studies of related proteins. *Curr Opin Struct Biol* **23**, 66–74 (2013).
- Nickson, A. A. & Clarke, J. What lessons can be learned from studying the folding of homologous proteins? *Methods* **52**, 38–50 (2010).
- Zarrine-Afsar, A., Larson, S. M. & Davidson, A. R. The family feud: do proteins with similar structures fold via the same pathway? *Curr Opin Struct Biol* **15**, 42–49 (2005).
- Vollrath, F. & Knight, D. P. Liquid crystalline spinning of spider silk. *Nature* **410**, 541–548 (2001).
- Heim, M., Keerl, D. & Scheibel, T. Spider silk: from soluble protein to extraordinary fiber. *Angew Chem Int Ed Engl* **48**, 3584–3596 (2009).
- Rising, A. & Johansson, J. Toward spinning artificial spider silk. *Nat Chem Biol* **11**, 309–315 (2015).
- Askarieh, G. *et al.* Self-assembly of spider silk proteins is controlled by a pH-sensitive relay. *Nature* **465**, 236–238 (2010).
- Hagn, F. *et al.* A conserved spider silk domain acts as a molecular switch that controls fibre assembly. *Nature* **465**, 239–242 (2010).
- Eisoldt, L., Thamm, C. & Scheibel, T. The role of terminal domains during storage and assembly of spider silk proteins. *Biopolymers* **97**, 355–361 (2011).
- Schwarze, S., Zwettler, F. U., Johnson, C. M. & Neuweiler, H. The N-terminal domains of spider silk proteins assemble ultrafast and protected from charge screening. *Nat Commun* **4**, 2815 (2013).
- Kronqvist, N. *et al.* Sequential pH-driven dimerization and stabilization of the N-terminal domain enables rapid spider silk formation. *Nat Commun* **5**, 3254 (2014).
- Jaudzems, K. *et al.* pH-dependent dimerization of spider silk N-terminal domain requires relocation of a wedged tryptophan side chain. *J Mol Biol* **422**, 477–487 (2012).
- Gaines, W. A., Sehorn, M. G. & Marcotte, W. R. Jr. Spidroin N-terminal domain promotes a pH-dependent association of silk proteins during self-assembly. *J Biol Chem* **285**, 40745–40753 (2010).
- Hagn, F., Thamm, C., Scheibel, T. & Kessler, H. pH-dependent dimerization and salt-dependent stabilization of the N-terminal domain of spider dragline silk—implications for fiber formation. *Angew Chem Int Ed Engl* **50**, 310–313 (2011).
- Vollrath, F. Strength and structure of spiders' silks. *Journal of biotechnology* **74**, 67–83 (2000).
- Jackson, S. E. & Fersht, A. R. Folding of chymotrypsin inhibitor 2.1. Evidence for a two-state transition. *Biochemistry* **30**, 10428–10435 (1991).
- Debye, P. & Hückel, E. The interionic attraction theory and deviations from ideal behavior in solution. *Phys. Z.* **24**, 185–206 (1923).
- de Los Rios, M. A. & Plaxco, K. W. Apparent Debye-Huckel electrostatic effects in the folding of a simple, single domain protein. *Biochemistry* **44**, 1243–1250 (2005).

20. Atkison, J. H., Parnham, S., Marcotte, W. R. Jr. & Olsen, S. K. Crystal Structure of the Nephila clavipes Major Ampullate Spidroin 1A N-terminal Domain Reveals Plasticity at the Dimer Interface. *J Biol Chem* **291**, 19006–19017 (2016).
21. Otkovs, M. *et al.* Diversified Structural Basis of a Conserved Molecular Mechanism for pH-Dependent Dimerization in Spider Silk N-Terminal Domains. *Chembiochem: a European journal of chemical biology* **16**, 1720–1724 (2015).
22. Jackson, S. E. How do small single-domain proteins fold? *Fold Des* **3**, R81–91 (1998).
23. Ries, J., Schwarze, S., Johnson, C. M. & Neuweiler, H. Microsecond folding and domain motions of a spider silk protein structural switch. *J Am Chem Soc* **136**, 17136–17144 (2014).
24. Gelman, H. & Gruebele, M. Fast protein folding kinetics. *Q Rev Biophys* **47**, 95–142 (2014).
25. Daggett, V. & Fersht, A. The present view of the mechanism of protein folding. *Nat Rev Mol Cell Biol* **4**, 497–502 (2003).
26. Pace, C. N., Grimsley, G. R. & Scholtz, J. M. Protein ionizable groups: pK values and their contribution to protein stability and solubility. *J Biol Chem* **284**, 13285–13289 (2009).
27. Knight, D. P. & Vollrath, F. Changes in element composition along the spinning duct in a Nephila spider. *Naturwissenschaften* **88**, 179–182 (2001).
28. Santoro, M. M. & Bolen, D. W. Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl alpha-chymotrypsin using different denaturants. *Biochemistry* **27**, 8063–8068 (1988).
29. Tanford, C. Protein denaturation. *Adv Protein Chem* **23**, 121–282 (1968).
30. Myers, J. K., Pace, C. N. & Scholtz, J. M. Denaturant m values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. *Protein Sci.* **4**, 2138–2148 (1995).
31. Becktel, W. J. & Schellman, J. A. Protein stability curves. *Biopolymers* **26**, 1859–1877 (1987).
32. Mallam, A. L. & Jackson, S. E. The dimerization of an alpha/beta-knotted protein is essential for structure and function. *Structure* **15**, 111–122 (2007).

Author Contributions

J.C.H. and H.N. designed experiments. J.C.H., S.R., C.R. and C.M.J. performed experiments. J.C.H., S.R., C.R., C.M.J., and H.N. analysed data. J.C.H., C.M.J., and H.N. interpreted results. J.C.H. and H.N. wrote the paper. All authors reviewed the manuscript.

Additional Information

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017