doi:10.1016/j.jmb.2012.02.040 J. Mol. Biol. (2012) 418, 197–207

Contents lists available at www.sciencedirect.com

Journal of Molecular Biology

journal homepage: http://ees.elsevier.com.jmb

N-Terminal Domain of *Bombyx mori* Fibroin Mediates the Assembly of Silk in Response to pH Decrease

Yong-Xing He†, Nan-Nan Zhang†, Wei-Fang Li, Ning Jia, Bao-Yu Chen, Kang Zhou, Jiahai Zhang, Yuxing Chen and Cong-Zhao Zhou*

Hefei National Laboratory for Physical Sciences at Microscale and School of Life Sciences, University of Science and Technology of China, Hefei, Anhui 230027, People's Republic of China

Received 20 October 2011; received in revised form 18 February 2012; accepted 23 February 2012
Available online 1 March 2012

Edited by M. Moody

Keywords:
fibroin;
X-ray crystallography;
N-terminal domain;
structural transition;
micelle

Fibroins serve as the major building blocks of silk fiber. As the major component of fibroin, the fibroin heavy chain is a considerably large protein comprising N-terminal and C-terminal hydrophilic domains and 12 highly repetitive Gly-Ala-rich regions flanked by internal hydrophilic blocks. Here, we show the crystal structure of the fibroin N-terminal domain (FibNT) at pH 4.7, revealing a remarkable double-layered anti-parallel β-sheet with each layer comprising two FibNT molecules entangled together. We also show that FibNT undergoes a pH-responsive conformational transition from random coil to β-sheets at around pH 6.0. Dynamic light scattering demonstrates that FibNT tends to oligomerize as pH decreases to 6.0, and electron microscopy reveals micelle-like oligomers. Our results are consistent with the micelle assembly model of silk fibroin and, more importantly, show that the N-terminal domain in itself has the capacity to form micelle-like structures in response to pH decrease. Structural and mutagenesis analyses further reveal the important role of conserved acidic residues clustered in FibNT, such as Glu56 and Asp100, in preventing premature β-sheet formation at neutral pH. Collectively, we suggest that FibNT functions as a pH-responsive self-assembly module that could prevent premature β-sheet formation at neutral pH yet could initiate fibroin assembly as pH decreases along the lumen of the posterior silk gland to the anterior silk gland.

Introduction

Silk produced by the silkworm *Bombyx mori* is a high-strength natural protein fiber that has been used to make textiles since 3000 BC.[1] Each silk fiber is about 10–25 μm wide[2] and composed of two core filaments bundled together through a sericin coat enveloping the filaments. The core filament is assembled from fibroins, which consist of three components: a 390-kDa heavy chain [fibroin heavy chain (FibH)],[3,4] a 26-kDa light chain,[5] and a 25-kDa glycoprotein.[6] After having been synthesized in posterior silk glands, fibroins are secreted into the lumen as an aqueous solution (12–15% by weight) and extruded by peristaltic motion to the ducts of middle silk glands, where the concentration increases to 20–30%.[1] In the lumen of anterior silk glands, the pH of the processing solution falls to around 4.9, salt concentration increases, and fibroins are finally spun into water-insoluble fibers by mechanical shear and the stretching action of the spinneret.[7] Artificially reeling silk from immobilized silkworms at faster and steady speeds greatly

*Corresponding author. E-mail address: zcz@ustc.edu.cn.
† Y.-X.H. and N.-N.Z. contributed equally to this work.
Abbreviations used: FibNT, fibroin N-terminal domain; FibH, fibroin heavy chain; SAD, single-wavelength anomalous diffraction; SeMet, selenomethionine; DLS, dynamic light scattering.

improves silk strength and toughness, which can even be compared to those of spider dragline silk, suggesting that spinning conditions are crucial to the mechanical properties of *B. mori* silk.[8]

Silk fibroin has a crystalline portion of about 95%, featuring abundant highly repetitive structural motifs GAGAGS, and the rest is amorphous, mostly composed of hydrophilic residues. As the major component of fibroin complex, FibH is composed of 5263 residues and comprises a highly conserved N-terminal hydrophilic domain with a signal peptide, a C-terminal hydrophilic domain, and 12 hydrophilic linkers interspersed among 12 GA-rich repetitive regions that account for its crystalline nature[3,4] (Fig. S1). Interestingly, the 12 linkers share highly conserved sequences, yet their roles remain elusive. As early as the 1950s, it had been proposed, based on X-ray diffraction analysis, that polypeptides in the crystalline portion of fibroin were arranged as anti-parallel pleated β-sheets.[9] The adjacent sheets pack together at distances of about 3.5 and 5.7 Å, of which the longer intersheet distance is explained by the presence of residues with bulky side chains, such as tyrosines. The planes of the β-sheets lie parallel with the axis of the fiber, which consists of bundles of nanofibrils with lateral dimensions of approximately 20 Å × 60 Å.[10] These nanofibrillar structures, which represent well-ordered crystallites in silk, are suggested to be embedded in a matrix of less ordered noncrystalline regions.[11] The hydrogen-bonding network of β-structured crystallites plays a major role in determining the strength and rigidity of the material,[12] while the noncrystalline regions adopting much more flexible conformations are likely to contribute to elasticity.[13]

The folding and assembly mechanism of silk fibroins has long been investigated. We previously fitted the FibH sequence to the anti-parallel β-sheet model and proposed that ~10 repeats of GAGAGS in the crystalline region of FibH should form a long β-strand of ~200 Å^3, which is around the upper limit of the crystallite size.[14] It was found that reconstituted fibroins, when incubated in water under mechanical shear, can form a native-like fiber dominated by a so-called 'parallel β' structure, where the β-strand is arranged parallel with the fiber axis. However, when alcohol is used to trigger assembly (without shear), there emerge entangled amyloid-like nanofibers with a 'cross-β' structural feature.[15] This structural variation implies that fibroin has an incredibly flexible and moldable structure, which senses the exterior conditions and undergoes conformational transitions accordingly. One essential step in silk spinning is the reduction of pH from the posterior part of the silk gland to the anterior part of the silk gland, which triggers the gelation of the condensed fibroin.[16] By mimicking the spinning conditions *in vitro*, Jin and Kaplan proposed an intriguing model of silk assembly: the spinning process is initiated by the self-assembly of micelle-like particles (100–200 nm), propagates through the aggregation of particles into larger globules (0.8–15 μm) and gel-like states, and finally ends with elongation and alignment under physical shear, leading to the formation of silk fiber.[17]

We crystallized the fibroin N-terminal domain (FibNT) at pH 4.7, and the crystal structure reveals an entangled β-sheet dimer. We also show that FibNT undergoes a pH-responsive conformational transition from random coil to β-sheets at around pH 6.0 and demonstrate the capability of FibNT to form oligomers with micelle-like substructures in a pH-responsive fashion. Structural and mutagenesis analyses further reveal the important role of conserved acidic residues clustered in FibNT, such as Glu56 and Asp100, in preventing premature β-sheet formation at neutral pH. Our results provide the first structural insights into the assembly mechanism of *B. mori* silk at the atomic level, which could possibly apply to all hymenopteran silks because of the high conservation of their N-terminal domains.

Results

Overall structure of FibNT

The N-terminal hydrophilic domain devoid of the signal peptide (Ile23-Ser126, referred to as FibNT) was overexpressed in *Escherichia coli*. Through exhaustive crystal screening trials, we eventually crystallized FibNT in space group $P6_522$ at pH 4.7. To determine the structure, we collected a single-wavelength anomalous diffraction (SAD) data set at 50.0–3.0 Å resolution from a single crystal of selenomethionine (SeMet)-substituted protein. Initial electron density map calculated from the experimental phases lacked the features of side chains. After *B*-factor sharpening had been applied, the quality of electron density improved, and unambiguous assignments of the side chains were allowed (Fig. S2). The protein sequence registry was confirmed by a difference Fourier analysis of SeMet.

The asymmetric unit of FibNT crystals contains a homodimer of a twisted eight-stranded β-sheet with a 2-fold symmetric axis perpendicular to the β-sheet (Fig. 1a). This is completely distinct from the α-helical structures of the spidroin N-terminal domain,[18,19] although the mechanisms of fibroin and spidroin assemblies were thought to be similar.[20] The two chains A and B are structurally similar, with a root-mean-square deviation (RMSD) of 2.1 Å over 64 C^α atoms, but they differ profoundly in the N-terminal segments Phe26-Val35, which form a short helix packing at the β-sheet in chain A but with a loop protruding to the solvent in chain B. Each chain is

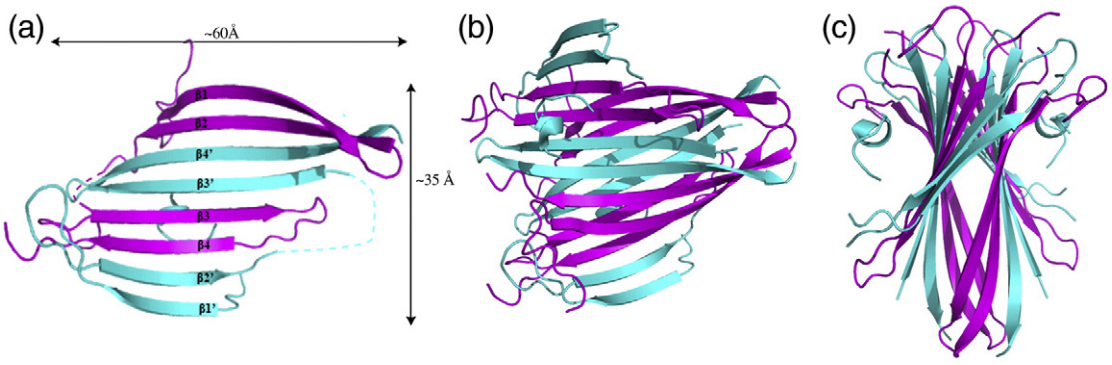

Fig. 1. Overall structure of FibNT. (a) Entangled dimer structure, subunit A (cyan), and subunit B (pink) are shown in cartoon representation. (b) Front view of the tetramer. (c) Side view of the tetramer.

composed of two β-hairpins and part of the loop connecting β2 and β3, and the last 18 residues (Gly109-Ser126) at the C-terminus are disordered. The dimer has a maximum length of ~60 Å along the β-strands and a width of ~35 Å across the strands. Two FibNT dimers further assemble into a tetramer (dimer of dimers) through a crystallographic 2-fold rotation axis (Fig. 1b and c). This crystallographic dimer–dimer interface involves a broad range of hydrophobic and hydrogen bonding interactions. The distance between the centers of the two sheets is an average of 10 Å.

Dimer entanglement

During model building, we found that disruptions of electron density between strands β2 and β3 result in ambiguous chain tracing, resulting in β-sheet topologies of $\beta1_A$–$\beta2_A$–$\beta4_A$–$\beta3_A$–$\beta3_B$–$\beta4_B$–$\beta2_B$–$\beta1_B$ and $\beta1_A$–$\beta2_A$–$\beta4_B$–$\beta3_B$–$\beta3_A$–$\beta4_A$–$\beta2_B$–$\beta1_B$ within the dimer (Fig. 2). However, given that the distances between the disrupted ends of $\beta2_A$ and $\beta3_A$ and between the disrupted ends of $\beta2_B$ and $\beta3_B$ are ~38 and 30 Å, respectively, the missing residues cannot fill in the loop connecting the two β-hairpins in the topology $\beta1_A$–$\beta2_A$–$\beta4_A$–$\beta3_A$–$\beta3_B$–$\beta4_B$–$\beta2_B$–$\beta1_B$, even when assuming that all these residues adopt fully extended conformations. Therefore, we established that the topology of the dimer is $\beta1_A$–$\beta2_A$–$\beta4_B$–$\beta3_B$–$\beta3_A$–$\beta4_A$–$\beta2_B$–$\beta1_B$. This entangled topology increases (approximately twice) the number of FibNT interchain hydrogen bonds by exchanging the β-hairpins of each chain within the dimer, compared to the presumable dimer without entanglement. Residues Thr36-Asn65 form the first β-hairpin structure with a type 1 β-turn in residues Asp49-Gly52, and residues Glu78-Ser107 form the second β-hairpin with a type 1 β-turn in residues Asp89-Gly92. These two β-hairpins have four available β-edges: strands β2 and β4 reciprocally form intermolecular hydrogen bonds, and strand β3 forms hydrogen bonds with its counterpart from the other chain (Fig. 2). The hydrogen bond registries in both chains are well ordered and strictly the same, suggesting that the β-sheet alignment is well stabilized and represents the native registry of the hydrogen bonds.

Characterization of pH-responsive conformational transition and size distribution of FibNT

Due to the essential role of pH in fibroin assembly, we probed the secondary structures of FibNT in the pH interval 7.0–3.0 using circular dichroism (CD) spectroscopy (Fig. 3a). At pH 7.0, CD spectra demonstrate a strong negative band centered around 197 nm and a positive band around 212 nm,

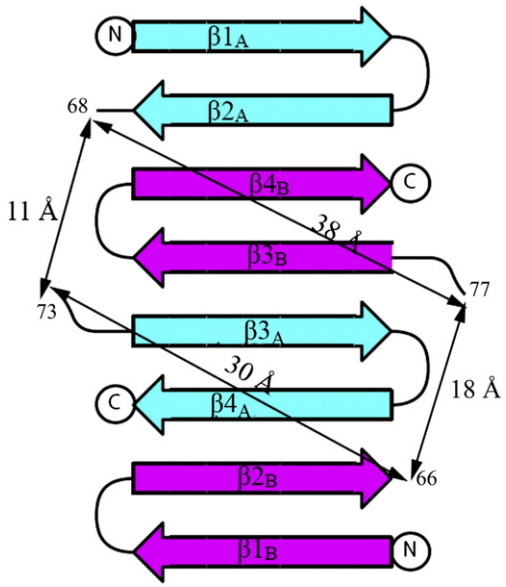

Fig. 2. Topology of FibNT dimer. The residue numbering of the disrupted ends of β-hairpins is marked in the figure. Structural analysis unambiguously established the topology of FibNT as $\beta1_A$–$\beta2_A$–$\beta4_B$–$\beta3_B$–$\beta3_A$–$\beta4_A$–$\beta2_B$–$\beta1_B$.

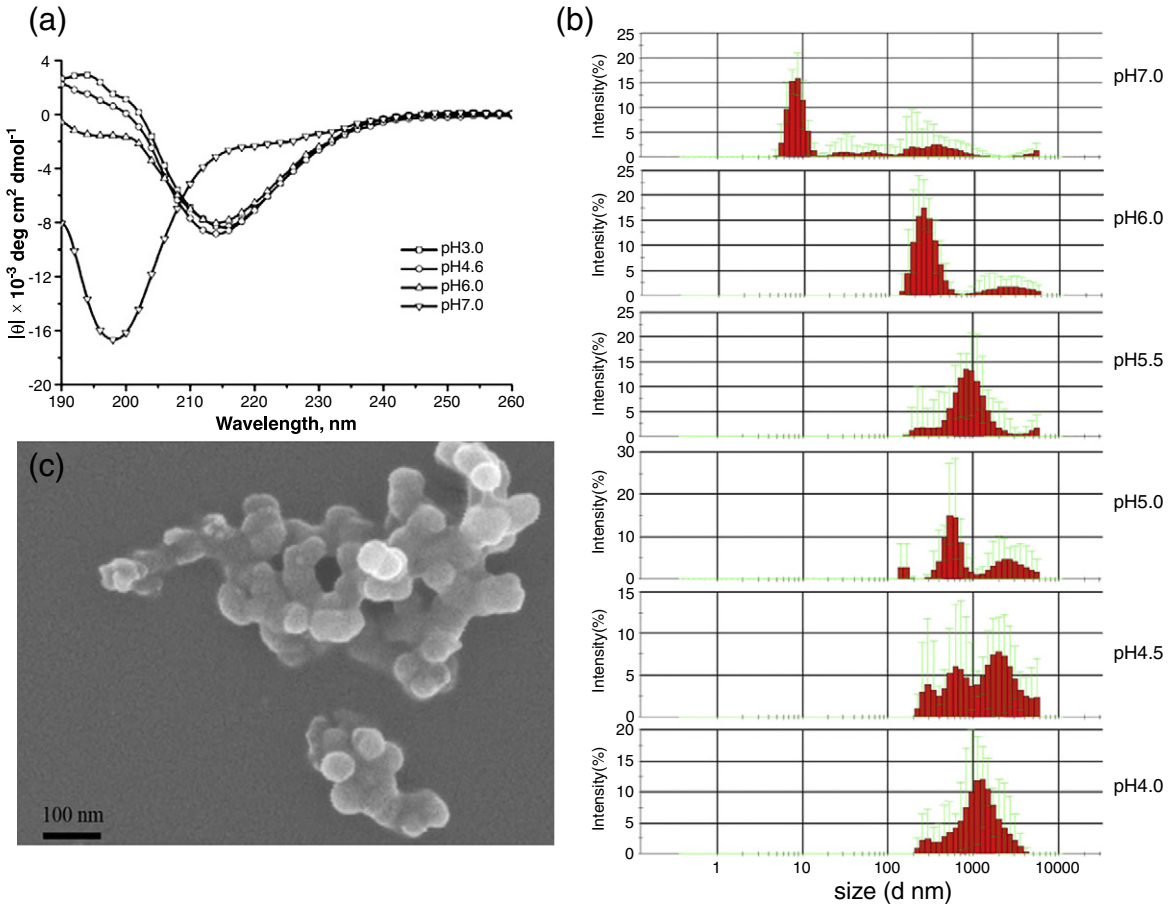

Fig. 3. (a) CD spectra of FibNT at pH 7.0 (down-pointing triangles), pH 6.0 (up-pointing triangles), pH 4.6 (circles), and pH 3.0 (squares). (b) Analysis of the oligomer size distributions of FibNT. Representative DLS spectra are shown at pH 7.0, 6.0, 5.5, 5.0, 4.5, and 4.0. (c) SEM image of FibNT oligomers formed at pH 5.0.

indicating the dominance of random-coil structure. However, when the pH decreases to 6.0, the two bands drastically invert to a positive band at ~195 nm and to a negative band at ~215 nm, showing the emergence of β-sheet conformation. As the pH drops to 3.0, the fraction of the β-sheet structure increases to ~40%. Titration from pH 3.0 to pH 7.0 turns FibNT back to the random-coil conformation (data not shown).

We also conducted dynamic light scattering (DLS) studies to probe the size distribution of FibNT in the pH interval 7.0–4.0 (Fig. 3b). At pH 7.0, the size distribution observed for FibNT reveals mainly particles with a hydrodynamic radius (R_h) of ~9 nm, representing low-molecular-weight species. The measurement also reveals particles with large R_h values ranging from ~10 nm to 1 μm. At pH 6.0, however, the peak of the size distribution curve shifts to ~200 nm, indicating that most of the proteins undergo higher oligomerization. With a further decrease in pH values, the size of the oligomers increases; at pH 4.0, most of the FibNT molecules have R_h values at around 1 μm. It is important to note that a pH-responsive pattern of oligomerization coincides with that of conformational transition, with turning points between pH 7.0 and pH 6.0. The close association of these two events suggests that oligomerization of FibNT is likely to rely on the emergence of a well-ordered β-structure.

Morphology of FibNT oligomers

To determine morphological manifestations of FibNT oligomers, we examined the preparations of FibNT with a scanning electron microscope (Fig. 3c). At pH 5.0 and a protein concentration of 1 mg/ml, FibNT oligomers adopt clustered micelle-like structures, with diameters of the micelles in the range of 50–100 nm. A previously postulated model of silk assembly suggests that internal hydrophobic blocks in fibroin are likely to be the driving force of micelle formation via extensive hydrophobic interactions, while the N-terminal and C-terminal hydrophilic

blocks undergo either homo-oligomerization or hetero-oligomerization to seal the hydrophobic core and to form the outer edges of the micelles.[17] The micelle-like morphology of FibNT oligomers that we show here not only is consistent with the micelle-forming silk assembly model but also identifies FibNT as a self-assembly unit that could possibly initiate fibroin assembly in response to pH decrease.

Structural basis for pH-responsive conformational transition

Despite the very distinct structures of FibNT and the spidroin N-terminal domain, both proteins can be self-assembled in response to pH decrease. It was demonstrated that the α-helical spidroin N-terminal domain readily self-assembles at pH 6.3 and displays delayed assembly when pH is above 7.0, and this pH-responsive assembly involves conserved acidic residues such as Asp40 and Glu8418. In FibNT, there are 19 acidic residues distributing on both faces of the dimeric β-sheet. Some of these acidic residues may have their pK_a values up-shifted near the transition point and therefore ionize at neutral pH to impede protein folding and assembly through charge repulsion (Fig. 4a). As clustering of acidic residues can increase side-chain pK_a values, we looked into the FibNT structure to identify clusters of conserved acidic residues. Notably, the

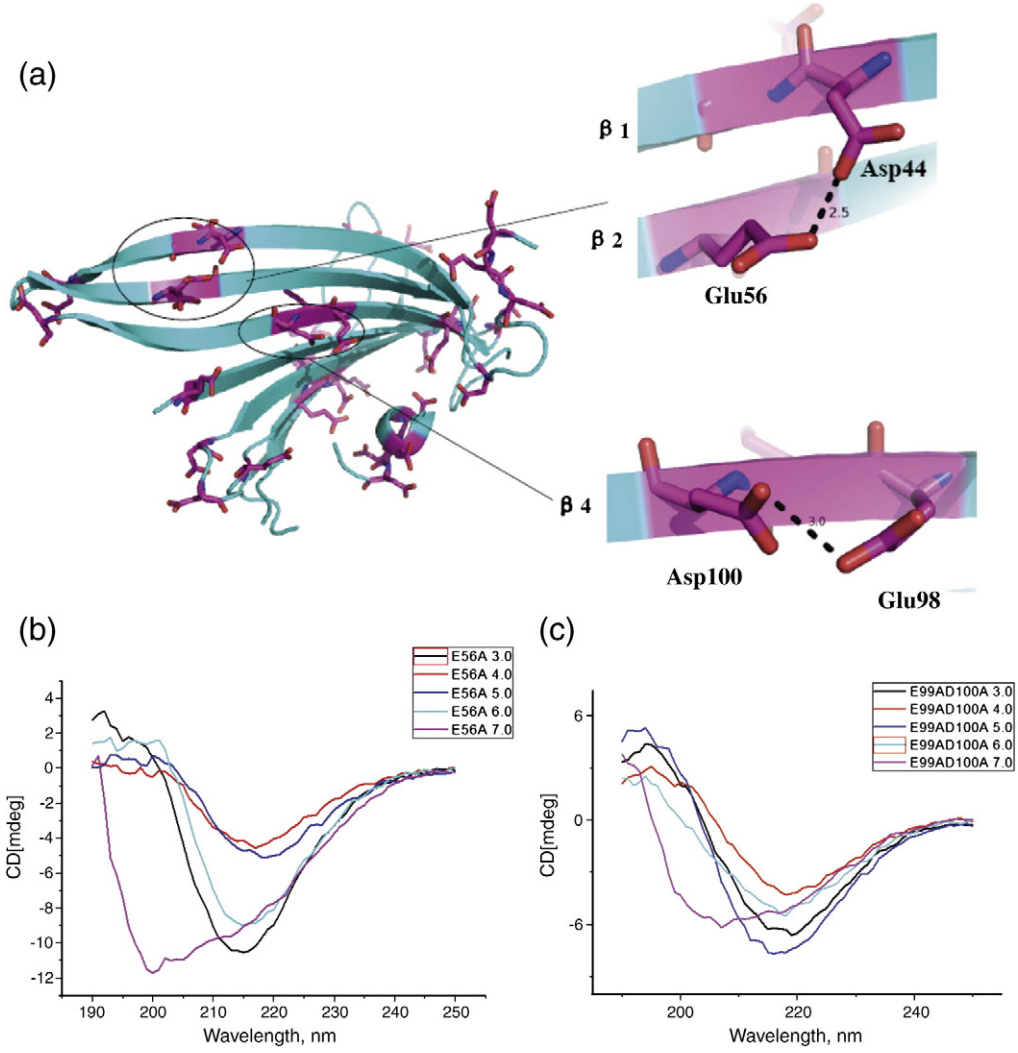

Fig. 4. (a) Acidic residue distribution in the structure of FibNT. The hydrogen bonds formed between two pairs of acidic residues (Asp44-Glu56 and Glu98-Asp100) are highlighted on the right. (b and c) CD spectra of FibNT mutants E56A and E99A/D100A in the pH interval 7.0–3.0. The broad negative peaks between 200 and 220 nm for both proteins at pH 7.0 indicate that they have a fraction of β-sheet conformation and are more folded than the wild-type FibNT at neutral pH.

strictly conserved Glu56 and Asp100 cluster with Asp44 and Glu98, respectively, by forming hydrogen bonds (Fig. 4a). Ionization of these residues can lead to charge repulsion of the side chains and can possibly reduce structure stability or even impair the well-ordered registry of the β-sheets. To test this hypothesis, we disrupted charge repulsion by constructing the Glu56Ala mutant and the Glu99Ala/Asp100Ala (Glu99 is also strictly conserved) double mutant and characterized the secondary structures of the mutant proteins by CD as a function of pH (Fig. 4b and c). Both mutants still show pH-responsive conformational transition but have β-sheet conformation emerging already at pH 7.0. It indicates that mutation of these residues can promote β-sheet formation at neutral pH rather than impair it, suggesting that charge repulsion between closely located acidic residues plays an important role in preventing premature β-sheet formation at neutral pH. It seems that the same mechanism is also used in the spidroin N-terminal domain, since Asp40 and Glu84 form a hydrogen bond and since the single mutant Asp40 or the double mutant Asp40-Glu84 was reported to promote self-assembly even at pH 8.0. The fibroin and spidroin N-terminal domains give us a nice example of convergent evolution at the molecular level—two structurally unrelated proteins evolving independently to attain similar functions by using similar chemical mechanisms.

Discussion

The FibNT structure presented here reveals a remarkable two-layered entangled β-sheet structure. To our surprise, the spatial arrangement of the strands is analogous to the amyloidal cross-β structure, which is associated with many neurodegenerative diseases. A typical amyloid fiber is composed of repeating substructures that are made up of β-strands running perpendicular to the fiber axis. An X-ray fiber diffraction pattern of amyloids features a meridional reflection at ~4.7 Å (corresponding to inter-β strand spacing) and an equatorial reflection at ~6–11 Å (corresponding to the distance between stacked β-sheets).[21,22] The geometric parameters of FibNT β-sheets coincide well with those of amyloid, with an inter-β strand spacing of ~4.8 Å and an intersheet distance of ~10 Å. More importantly, we demonstrate that FibNT undergoes higher oligomerization when pH decreases to 6.0 and can be self-assembled into micelle-like structures with diameters around 50–100 nm. The dimer entanglement that we observed in the crystal structure of FibNT could be essential for FibNT oligomerization because, in solution, each FibNT monomer can donate one β-hairpin to reconstitute the eight-stranded β-sheet, resulting in oligomerization in a head-to-tail fashion (Fig. 5). Note that the loops connecting the β-hairpins are very flexible and do not restrain the reconstituted β-sheets in the same plane, so the oligomers display the morphology of micelle-like particles other than nanofibrils, which are usually associated with two-dimensional molecular assembly and propagation.

Another possible mechanism for the oligomerization of FibNT could be the abundantly exposed hydrophobic residues on both sides of the dimeric β-sheet, which serve as alternative assembly interfaces. One face of the dimeric β-sheet forms the dimer–dimer interface through extensive hydrophobic and polar interactions. This interface is very

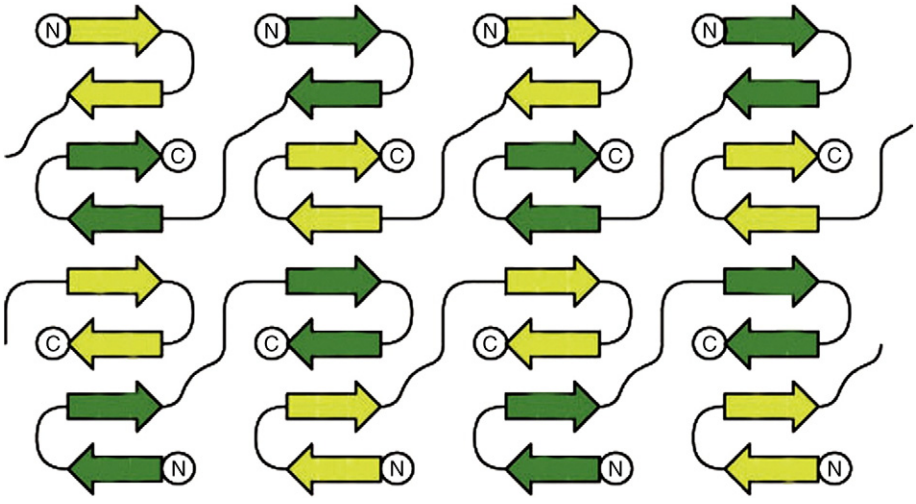

Fig. 5. Oligomerization model of FibNT via β-hairpin entanglements. In this model, each FibNT monomer donates one β-hairpin to reconstitute the eight-stranded β-sheet and subsequently forms oligomers in a head-to-tail fashion.

similar to the so-called 'steric zipper,' which refers to a compact dehydrated interface that makes two layers of β-sheet interdigitated during amyloid polymerization. A number of hydrophobic residues, such as Ile59, Ile81, Val85, Val97, and Val101, lie at the center of the dimer–dimer interface (Fig. 6a), and multiple sequence alignment (Fig. S3) indicates that these residues are well conserved in the FibNT of hymenopteran insects, underscoring the functional importance of these residues. A well-conserved hydrophobic patch composed of residues Leu102, Phe84, Ile86, and Ile96 is also found on the other face of the dimeric β-sheet (Fig. 6b). This hydrophobic patch is buried by the N-terminal short helix of one FibNT monomer, with the side chain of hydrophobic Phe26 pointing inwards and with hydrophilic residues Asp25 and Glu28 pointing outwards. However, in the other monomer, this N-terminal segment flips away from the hydrophobic patch, which could possibly bind to the hydrophobic patch of another molecule in solution.

Our crystal structure and mutagenesis analyses also highlight the important role of clustered acidic residues in preventing premature β-sheet formation at neutral pH. Note that the hydrophilic linkers also contain a number of acidic residues with a predicted p*I* of around 4.5; these interspersed regions may also play a role in preventing the premature assembly of fibroin at neutral pH. In light of our results and the micelle assembly model, we propose a model for how FibNT mediates silk assembly: In the lumen of posterior silk glands where the pH value is around 6.9, FibNT exists in random-coil conformation and has net negative charges due to the ionization of acidic residues (Fig. 7). Due to charge repulsion, the premature β-sheet formation and the self-assembly of fibroin are prevented at this point. With the procession of

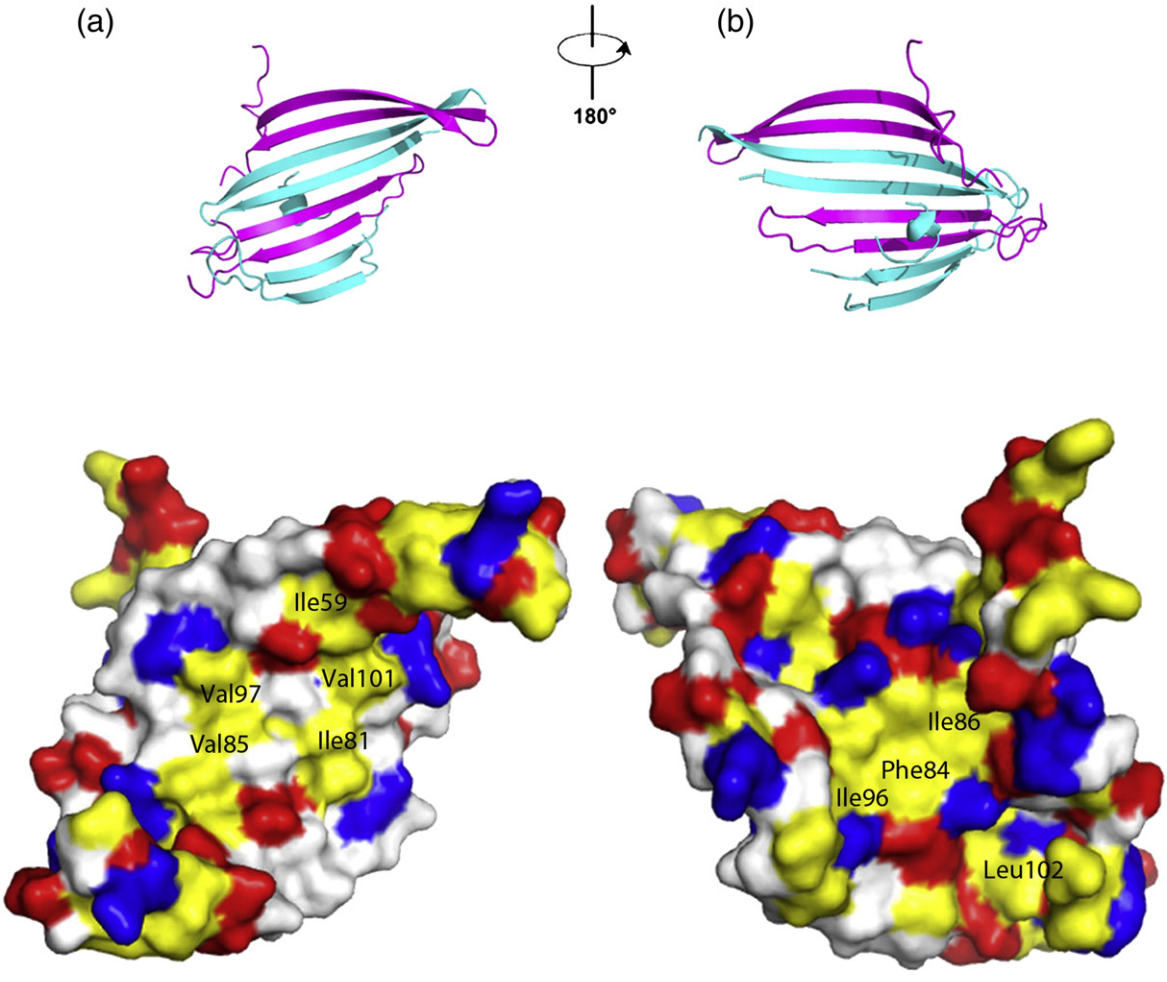

Fig. 6. Surface hydrophobic patches of FibNT. Hydrophobic residues on the surface are shown in yellow and marked in the figure, basic residues are shown in blue, acidic residues are shown in red, and the rest is shown in white. (a) Front view of FibNT (facing the dimer–dimer interface). (b) Back view of FibNT.

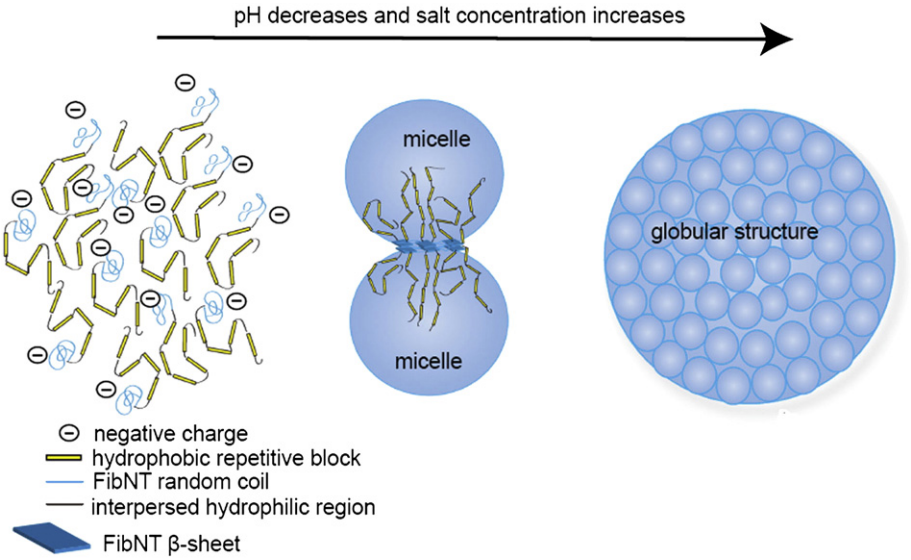

Fig. 7. Assembly model of *B. mori* fibroin. At a pH of around 6.9, FibNT exists in random-coil conformation and has net negative charges due to the ionization of acidic residues, thus preventing premature β-sheet formation. When local pH decreases, the acidic side chains of FibNT begin to protonate, resulting in less negative charges and weaker electrostatic repulsion, and drives FibNT to fold into β-sheets and to oligomerize via β-hairpin entanglements. Increasing ionic strength in silk glands could also contribute to the oligomerization of FibNT by shielding electrostatic repulsion. The hydrophobic patches could possibly serve as micelle–micelle docking sites.

fibroin solution along the silk gland lumen, local pH decreases, and the acidic side chains of FibNT (and perhaps the linker regions) begin to protonate; this results in less negative charges and weaker charge repulsion and drives FibNT to fold into β-sheets and to oligomerize via β-hairpin entanglements. Subsequently, micelle formation is promoted, with FibNT sequestering the hydrophobic regions and forming 'water-tight' micelle outer edges. Increasing ionic strength in silk glands could also promote the oligomerization of FibNT by shielding charge repulsion. The solvent-exposed hydrophobic patches at one face of the FibNT β-sheet could interact with hydrophobic repetitive regions, while those at the dimer–dimer interface could possibly serve as micelle–micelle docking sites, leading to much larger globular structures that can be finally spun into silk fiber under the physical shear imposed at the spinneret.

In summary, our results provide the first structural insights into the assembly mechanism of *B. mori* silk at the atomic level. Revealing FibNT as an essential module in mediating pH-responsive assembly, we find that the functions of N-terminal domains in *B. mori* and spiders are strikingly similar, despite the very distinct sequences and three-dimensional structures. The studies on the assembly mechanism of *B. mori* silk will not only complement our understanding of how the remarkable silk fiber forms but also provide structural insights for designing artificial silk materials.

Materials and Methods

Construction, expression, and purification of FibNT

The cDNA encoding the N-terminal domain of FibH (FibNT, residues 20–126) was cloned into a modified pET28a expression vector (Novgan) with an additional six-histidine coding sequence at the 5′ end of the genes. The recombinant plasmid was transformed into competent *E. coil* Rosetta (DE3) cells, which were cultured in 400 ml of 2×YT containing 0.01 mg/ml kanamycin. The culture was grown at 310 K to an A_{600} of 0.6 and induced with 0.2 mM IPTG for 4 h. After being harvested by centrifugation, the cells were resuspended in 20 mM cold Tris–HCl (pH 8.0) and 100 mM NaCl. After three steps of freeze thawing and subsequent sonication, the lysed cells were centrifuged at 16,000*g* for 20 min. The His-tagged proteins were purified with an Ni-NTA affinity chromatography column (Amersham Biosciences), followed by a gel-filtration column (HiLoad 16/60 Superdex 75 prep grade; Amersham Biosciences) equilibrated with 20 mM Tris–HCl (pH 8.0) and 100 mM NaCl, and then eluted in the same buffer. The resulting fractions were analyzed by 15% SDS-PAGE. Site-directed mutagenesis was performed using the QuickChange site-directed mutagenesis kit (Stratagene, La Jolla, CA, USA), with the plasmid encoding wild-type FibNT as template. The mutant proteins were expressed, purified, and stored in the same manner as the wild-type protein.

Crystallization and data collection

FibNT was crystallized at 289 K used hanging-drop vapor diffusion. The crystals of FibNT were grown in a

drop of 5 mg/ml protein in 20 mM Tris–HCl (pH 8.0) and 100 mM NaCl with an equal volume of reservoir solution [1.2 M ammonia sulfate and 0.1 M sodium acetate (pH 4.6); the actual pH in the drop is 4.7] within 7 days. The SeMet derivative crystals were grown under the same conditions. The crystals were transferred to a cryoprotectant (reservoir solution supplemented with 25% glycerol) and flash-frozen in liquid nitrogen. Multiwavelength anomalous dispersion data for a SeMet derivative crystal were collected with an MX225 CCD (MARresearch, Germany) at the Shanghai Synchrotron Radiation Facility at a radiation wavelength of 0.9794 Å using 17 U at 100 K. All diffraction data were indexed, integrated, and scaled with HKL2000.[23]

Structure determination and refinement

The crystal structure of FibNT from a SeMet-substituted protein crystal was determined to a maximum resolution of 3.0 Å using the SAD phasing method. The SHELX program suite[24] was used to locate heavy atoms, and six selenium atoms were identified. The phase was calculated and further improved with the program SOLVE/RESOLVE‡.[25] Electron density maps calculated from solvent-flattened experimental phases lack side-chain features and cannot be traced, possibly due to a strong anisotropic diffraction effect along the *a* and *b* axes. When a *B*-factor of −50 was applied to the data,[26] the electron density was greatly sharpened, and an initial model was built automatically with the BACCANEER program implemented in the CCP4 program suite.[27] A complete model for FibNT was finally built manually from the initial model using the Coot program,[28] and side-chain assignments were confirmed by the positions of SeMet. The model was then refined with the Refmac5 program, and the final model was evaluated with the programs MolProbity[29] and Procheck.[30] The data collection and structure refinement statistics are listed in Table 1. Sequence alignment was performed using the programs CLUSTAL W[31] and ESPript.[32] Structural comparison was carried out using the DALI server§. All structure figures were created using the program PyMOL.[33]

CD spectroscopic analysis

The CD spectra of FibNT and mutants were recorded at 18 °C with a spectropolarimeter (Jasco, 810) using a quartz cuvette with a path length of 0.1 cm. The protein samples were dissolved to 0.1 mg/ml in the buffers of 20 mM potassium phosphate and 100 mM ammonium bisulfate at pH 3.0, 4.6, 6.0, and 7.0, respectively. All samples were centrifuged at 16,000 rpm for 10 min at 18 °C. The protein was omitted from the control sample, and each spectrum was corrected by subtracting the baseline (control). All experiments were performed in triplicate.

DLS and scanning electron microscopy

The effect of pH on the hydrodynamic diameter (R_h) of the N-terminal domain was measured at 25 ± 0.1 °C with an ALV/DLS/SLS-5022F spectrometer with a multi-τ digital time correlation (ALV5000) and a cylindrical 22-mW UNIPHASE He–Ne laser (λ = 632 nm). The intensity–intensity time correlation function $G^{(2)}(t,q)$ was measured to determine the linewidth distribution $G(\Gamma)$. For diffusive relaxation, Γ is related to the translational diffusion coefficient D of the scattering object (macromolecules or colloid particles) in dilute solution or dispersion by $D=(\langle\Gamma\rangle/q^2)_{C\to 0,q=0}$ and to the hydrodynamic radius from the Stockes–Einstern equation $R_h=k_BT/6\pi\eta D$, where η, k_B, and T are the solvent viscosity, the Boltzmann constant, and the absolute temperature, respectively. Hydrodynamic radius distribution was calculated from the Laplace inversion of a corresponding measured $G^{(2)}(t,q)$ using the CONTIN program. All DLS measurements were conducted at a scattering angle θ of 45° at 37 °C. Protein samples (at approximate concentrations of 1 mg/ml) were prepared in Tris–HCl buffer (20 mM Tris–HCl and 150 mM NaCl) in the pH interval 7.0–4.0.

For scanning electron microscopy imaging, the FibNT sample (1 mg/ml) at pH 5.0 was absorbed onto a carbon-coated copper grid by floating the grid on a drop of FibNT solution for 30 s. The excess solution was removed by filter paper blotting. The samples were dried and covered with

Table 1. Data collection, phasing, and refinement statistics

	SAD peak
Data collection statistics	
Wavelength (Å)	0.9794
Space group	$P6_522$
Unit cell parameters	
a, *b*, *c* (Å)	74.52, 74.52, 208.23
α,β, γ (°)	90, 90, 120
Resolution limit (Å)	50.00–3.00[a] (3.05–3.00)
Unique reflections	7427 (328)
Completeness (%)	99.4 (94.5)
R_{merge} (%)[b]	11.2 (65.4)
$I/\sigma I$	40.2 (3.3)
Redundancy	31.9 (19.1)
SAD phasing	
Heavy-atom sites	6
Mean factor of merit after SOLVE	0.35
Mean factor of merit after RESOLVE	0.68
Refinement statistics	
Resolution limit (Å)	37.26–3.00
R-factor (%)[c]	31.1
R-free (%)[d]	33.6
RMSD bond length (Å)[e]	0.008
RMSD bond angles (°)	1.143
Average of *B*-factors (Å^2)	101.8
Ramachandran plot[f]	
Most favored region (%)	89.9
Additionally allowed region (%)	7.9
Protein Data Bank entry	3UA0

[a] The values in parentheses refer to statistics in the highest-resolution bin.
[b] $R_{merge}=\sum_{hkl}\sum_i|I_i(hkl)-\langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of an observation, and $\langle I(hkl)\rangle$ is the mean value for its unique reflection. Summations are over all reflections.
[c] R-factor $=\sum_h|F_o(h)-F_c(h)|/\sum_h F_o(h)$, where F_o and F_c are the observed and calculated structure factor amplitudes, respectively.
[d] *R*-free was calculated with 5% of the data excluded from the refinement.
[e] RMSD from ideal values.
[f] Categories were defined by MolProbity.

‡ http://www.solve.lanl.gov
§ http://www2.ebi.ac.uk/dali/

a gold layer using a Hummer gold sputtering system and imaged with a Sirion200 electron microscope (FEI Co., Hillsboro, OR, USA) operated at 20 kV.

Accession number

The final coordinates and structure factors were deposited in the Protein Data Bank‖ under accession code 3UA0.

Acknowledgement

This work was sponsored by the 973 Project (no. 2012CB114601) from the Ministry of Science and Technology of China. We thank the staff at the Shanghai Synchrotron Radiation Facility and the Beijing Synchrotron Radiation Facility for data collection.

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2012.02.040

References

1. Matsumoto, A., Kim, H., Tsai, I., Wang, X., Cebe, P. & Kaplan, D. (2006). *Handbook of Fiber Chemistry*, 3rd edit. Taylor and Francis, Boca Raton, FL.
2. Zhao, H. P., Feng, X. Q. & Shi, H. J. (2007). Variability in mechanical properties of *Bombyx mori* silk. *Mater. Sci. Eng. C*, **27**, 675–683.
3. Zhou, C. Z., Confalonieri, F., Jacquet, M., Perasso, R., Li, Z. G. & Janin, J. (2001). Silk fibroin: structural implications of a remarkable amino acid sequence. *Proteins*, **44**, 119–122.
4. Zhou, C. Z., Confalonieri, F., Medina, N., Zivanovic, Y., Esnault, C., Yang, T. *et al.* (2000). Fine organization of *Bombyx mori* fibroin heavy chain gene. *Nucleic Acids Res.* **28**, 2413–2419.
5. Yamaguchi, K., Kikuchi, Y., Takagi, T., Kikuchi, A., Oyama, F., Shimura, K. & Mizuno, S. (1989). Primary structure of the silk fibroin light chain determined by cDNA sequencing and peptide analysis. *J. Mol. Biol.* **210**, 127–139.
6. Chevillard, M., Couble, P. & Prudhomme, J. C. (1986). Complete nucleotide sequence of the gene encoding the *Bombyx mori* silk protein P25 and predicted amino acid sequence of the protein. *Nucleic Acids Res.* **14**, 6341–6342.
7. Magoshi, J., Magoshi, Y. & Nakamura, S. (1994). Mechanism of fiber formation of silkworm. *Silk Polym.* **544**, 292–310.
8. Shao, Z. Z. & Vollrath, F. (2002). Materials: surprising strength of silkworm silk. *Nature*, **418**, 741.
9. Marsh, R. E., Corey, R. B. & Pauling, L. (1955). An investigation of the structure of silk fibroin. *Biochim. Biophys. Acta*, **16**, 1–34.
10. Dobb, M. G., Fraser, R. D. B. & Macrae, T. P. (1967). Fine structure of silk fibroin. *J. Cell Biol.* **32**, 289.
11. Chen, X., Knight, D. P., Shao, Z. & Vollrath, F. (2002). Conformation transition in silk protein films monitored by time-resolved Fourier transform infrared spectroscopy: effect of potassium ions on *Nephila* spidroin films. *Biochemistry*, **41**, 14944–14950.
12. Knowles, T. P., Fitzpatrick, A. W., Meehan, S., Mott, H. R., Vendruscolo, M., Dobson, C. M. & Welland, M. E. (2007). Role of intermolecular forces in defining material properties of protein nanofibrils. *Science*, **318**, 1900–1903.
13. Dong, Z. Y., Lewis, R. V. & Middaugh, C. R. (1991). Molecular mechanism of spider silk elasticity. *Arch. Biochem. Biophys.* **284**, 53–57.
14. Drummy, L. F., Farmer, B. L. & Naik, R. R. (2007). Correlation of the beta-sheet crystal size in silk fibers with the protein amino acid sequence. *Soft Matter*, **3**, 877–882.
15. Gong, Z., Huang, L., Yang, Y., Chen, X. & Shao, Z. (2009). Two distinct beta-sheet fibrils from silk protein. *Chem. Commun. (Cambridge)*, 7506–7508.
16. Terry, A. E., Knight, D. P., Porter, D. & Vollrath, F. (2004). pH induced changes in the rheology of silk fibroin solution from the middle division of *Bombyx mori* silkworm. *Biomacromolecules*, **5**, 768–772.
17. Jin, H. J. & Kaplan, D. L. (2003). Mechanism of silk processing in insects and spiders. *Nature*, **424**, 1057–1061.
18. Knight, S. D., Askarieh, G., Hedhammar, M., Nordling, K., Saenz, A., Casals, C. *et al.* (2010). Self-assembly of spider silk proteins is controlled by a pH-sensitive relay. *Nature*, **465**, 236–238; U125.
19. Lin, Z., Huang, W., Zhang, J., Fan, J. S. & Yang, D. (2009). Solution structure of eggcase silk protein and its implications for silk fiber formation. *Proc. Natl Acad. Sci. USA*, **106**, 8906–8911.
20. Kaplan, D. L. & Omenetto, F. G. (2010). New opportunities for an ancient material. *Science*, **329**, 528–531.
21. Eisenberg, D., Nelson, R., Sawaya, M. R., Balbirnie, M., Madsen, A. O., Riekel, C. & Grothe, R. (2005). Structure of the cross-beta spine of amyloid-like fibrils. *Nature*, **435**, 773–778.
22. Mitraki, A., Papanikolopoulou, K., Schoehn, G., Forge, V., Forsyth, V. T., Riekel, C. *et al.* (2005). Amyloid fibril formation from sequences of a natural beta-structured fibrous protein, the adenovirus fiber. *J. Biol. Chem.* **280**, 2481–2490.
23. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Macromol. Crystallogr. Part A*, **276**, 307–326.
24. Uson, I. & Sheldrick, G. M. (1999). Advances in direct methods for protein crystallography. *Curr. Opin. Struct. Biol.* **9**, 643–648.
25. Adams, P. D., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W. *et al.* (2002). PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **58**, 1948–1954.

‖ http://www.rcsb.org/pdb/

26. Strong, M., Sawaya, M. R., Wang, S., Phillips, M., Cascio, D. & Eisenberg, D. (2006). Toward the structural genomics of complexes: crystal structure of a PE/PPE protein complex from *Mycobacterium tuberculosis*. *Proc. Natl Acad. Sci. USA*, **103**, 8060–8065.
27. Collaborative Computational Project, Number 4. (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **50**, 760–763.
28. Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **60**, 2126–2132.
29. Davis, I. W., Leaver-Fay, A., Chen, V. B., Block, J. N., Kapral, G. J., Wang, X. *et al.* (2007). MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* **35**, W375–W383.
30. Laskowski, R. A., Macarthur, M. W., Moss, D. S. & Thornton, J. M. (1993). Procheck—a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291.
31. Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–4680.
32. Gouet, P., Robert, X. & Courcelle, E. (2003). ESPript/ENDscript: extracting and rendering sequence and 3D information from atomic structures of proteins. *Nucleic Acids Res.* **31**, 3320–3323.
33. DeLano, W. L. The PyMOL Molecular Graphics System. DeLano Scientific LLC, San Carlos, CA. http://www.pymol.org.