

# Force distribution determines optimal length of $\beta$ -sheet crystals for mechanical robustness†

Senbo Xiao,<sup>abc</sup> Wolfram Stacklies,<sup>b</sup> Cedric Debes<sup>a</sup> and Frauke Gräter<sup>\*ab</sup>

Received 11th June 2010, Accepted 12th August 2010

DOI: 10.1039/c0sm00510j

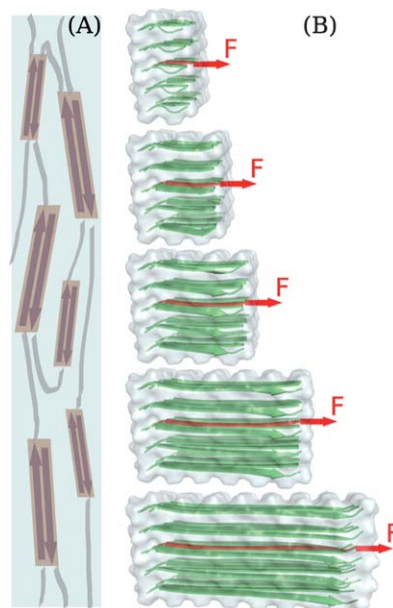
**Atomistic models of  $\beta$ -sheet crystals with varying number of alanines were mechanically tested in Molecular Dynamics simulations. The gain in mechanical robustness per residue is maximal for strands that are 8 residues long, because the external force is efficiently deflected into neighbouring strands over the distance they span, rendering 8 residues a mechanical optimum for a semi-crystalline material like silk.**

Among the secondary structure types a protein can adopt,  $\beta$ -sheets are generally those with the highest resistance to mechanical force. Consequently, evolution has selected  $\beta$ -sheets as the major building blocks of the most resilient proteins, such as titin immunoglobulin and fibronectin domains,<sup>1,2</sup> amyloid fibers,<sup>3</sup> or the crystalline phase in natural silk fibers.<sup>4,5</sup>  $\beta$ -Sheets are similarly found in the crystalline phases of some polymers like polyamides or polyurethanes. They feature hydrogen bonds periodically formed between adjacent  $\beta$ -strands in the sheet, the concurrent rupture of which can resist high forces.

A prime example of a natural  $\beta$ -sheet structure optimized for the ability of bearing high mechanical load is the  $\beta$ -sheets formed by silk proteins. Silk is a material with outstanding toughness not matched by any other synthetic material. The crystalline units are formed from stacked  $\beta$ -sheets, each of which comprises several strands of mostly 8 alanine residues (or alternating alanine-glycine, depending on individual type of silk). These crystals act as force-bearing units, cross-linking the amorphous phase they are embedded in (Fig. 1A). The mechanical stability of a molecular architecture like a  $\beta$ -sheet depends on a number of important factors, such as the chemistry, size, and relative arrangement of the chains. Given a certain composition and geometry of a  $\beta$ -sheet crystal with respect to the applied force, is there an optimal length of the strands in the crystal so that the material maximizes its macroscopic mechanical toughness? In the case of silk, is one of the structures shown in Fig. 1B optimally using the protein material for mechanical performance? Optimizing the resistance against rupture gained by each unit of the protein or polymer would

allow the efficient usage of the material, a competitive advantage in a world of limited resources.

We first determined the preferred length of strands in naturally occurring  $\beta$ -sheets, *i.e.* the strand length dominantly selected by evolution. We measured the distribution of  $\beta$ -strand length in the SCOP (Structural Classification of Proteins)<sup>6</sup> database by determining the secondary structure of the SCOP domains using STRIDE (STRuctural IDentification).<sup>7</sup> STRIDE is known to give a more accurate assignment for 70% of cases as compared to DSSP.<sup>8</sup> The length distribution for all SCOP domains features a maximum at 4–5 residues (Fig. 2), which reflects a positive selection for foldable and functional proteins in general and only partly an optimization for mechanical robustness. Interestingly, when restricting the analysis to a protein fold known for its mechanical function in the biological context and for its extra-ordinary stability against mechanical force, namely the immunoglobulin-like  $\beta$ -sandwich fold (<http://scop.mrc-lmb.cam.ac.uk/scop/data/scop.b.c.b.html>), the overall distribution shifts to longer strands, with a maximum at 6 residues per



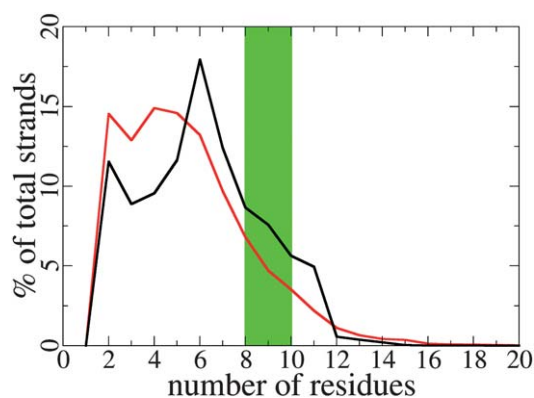
**Fig. 1** Molecular structure of silk. (A) Schematic drawing of a silk fiber comprising two different phases, namely the  $\beta$ -sheet crystals simplified as brown antiparallel arrows, cross-linking the amorphous peptides shown in grey. (B) Molecular models of silk protein crystals, varying in their strand length from 4 (top), 6, 8 (corresponding to natural dragline silk), and 12 to 16 residues (bottom). Central strands are shown in red with arrows indicating the external pulling force applied in FPMD simulations.

<sup>a</sup>Heidelberg Institute for Theoretical Studies, Schloss-Wolfsbrunnengasse 35, Heidelberg, Germany. E-mail: frauke.graeter@h-its.org; Fax: +49-6221-533298; Tel: +49-6221-533267

<sup>b</sup>Key Laboratory for Computational Biology, CAS-MPG Partner Institute for Computational Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, 320 Yueyang Road, Shanghai, 200031, China

<sup>c</sup>The Hartmut Hoffmann-Berling International Graduate School of Molecular and Cellular Biology, Heidelberg University, Heidelberg, Germany

† This paper is part of a *Soft Matter* themed issue dedicated to the International Soft Matter Conference 2010.



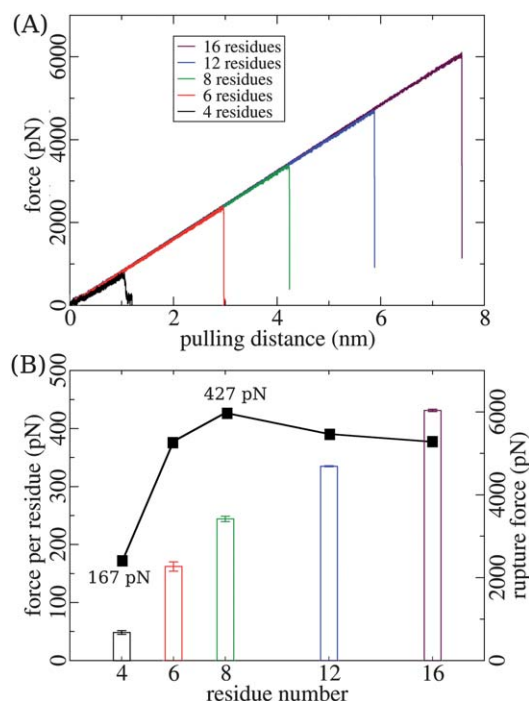
**Fig. 2**  $\beta$ -Strand lengths distribution in proteins. The distribution for all protein structures in SCOP shows a peak at 4–5 residues (red), and for proteins of the SCOP immunoglobulin-like fold a peak at 6 residues (black).  $\beta$ -strand lengths found in silk protein sequences, shown as green bar, vary between 8 and 10 residues.

b-strand. A similar distribution with a peak for 6 residues is found when only considering domains of titin, tenascin, and fibronectin from this family, for which clear evidence for their force-carrying function is at hand (data not shown). Spider silk shows strand lengths of 8–10 residues,<sup>9</sup> and thus another shift in the distribution towards longer  $\beta$ -strands as compared to other protein domains (Fig. 2, green bar). Since the primary evolutionary constraint for silk crystalline regions is their mechanical stability, we conclude that  $\beta$ -strand lengths of 8–10 residues are an optimal choice from a material design point of view. This length is roughly two times the length generally dominating proteins under multiple evolutionary constraints.

In contrast, in a recent computational study, 2D models of  $\beta$ -sheets consisting of three  $\beta$ -strands were mechanically analyzed and found to reach a mechanical optimum in terms of the pull-out resistance against force for a strand length of 4 residues.<sup>10</sup> However, the validity of this number for the naturally occurring 3D structures, which are stabilized by both intra-sheet hydrogen bonding and inter-sheet sidechain interactions, remains to be investigated.

To reconcile these observations, we here ask how the rupture force of a  $\beta$ -sheet structure varies with strand length to determine the mechanically optimal number of residues in a  $\beta$ -strand. We have recently built and mechanically characterized atomistic models of spider dragline silk protein crystals.<sup>11</sup> Crystals differing in their protein sequence and strand arrangement (parallel *versus* antiparallel) were compared and validated by experiments.<sup>12</sup> Their extra-ordinary rupture forces and stiffness explained their function of reinforcing silk fibers. Here, we built silk  $\beta$ -sheet protein crystals which differed in their number of alanines in each strand and therefore in the number of inter-strand hydrogen bonds from 4 to 16 (Fig. 1B), comprising those lengths most commonly found in nature. All of them consisted of five strands per  $\beta$ -sheet layer and five of these layers. We chose an antiparallel arrangement of strands in a  $\beta$ -sheet.<sup>11</sup> Crystals of parallel sheets are softer, but can be expected to show a similar scaling. We fully equilibrated the five all-atom structures in explicit water for 10 ns, before subjecting them to force-probe molecular dynamics (FPMD) simulations,<sup>13</sup> in which mechanical force was applied to the C-terminus of the central strand as indicated in Fig. 1B.

We used the simulation package GROMACS 4.0.5<sup>14</sup> to perform all simulations. In FPMD simulations, we moved a virtual harmonic



**Fig. 3** Force for rupturing  $\beta$ -sheet crystals of varying strand length. (A) Force profiles for models with strand lengths varying from 4 to 16 residues as indicated. We observe a linear increase in force prior to failure in all cases. (B) Rupture forces, *i.e.* maximal forces observed in FPMD simulations as shown in (A), again for models with varying strand lengths (color bar). The longer the crystal, the higher the rupture force required to pull out the central strand. By dividing the rupture force by the number of residues in the strand, we obtained a normalized rupture force per residue (black curve), which shows a maximum of 427 pN for strands of 8 residues.

spring with a force constant of 500 kJ/mol/nm<sup>2</sup> away from the pulled group, the terminal C- $\alpha$  atom of the central strand, with a pulling rate of 0.2 nm/ns. More details have been published previously.<sup>11</sup> The resulting force profiles are shown in Fig. 3A. A sudden drop of the force reflects failure of the crystalline structures. In all cases, prior to failure, we observed the same linear increase in force, indicating a high constant elastic modulus steadily increasing with the length of the  $\beta$ -strands. However, rupture forces strongly varied with strand length, with larger rupture forces for longer strands, as expected (Fig. 3A).

The relevant quantity for optimizing a material is the rupture force gained per residue in the crystal. Thus, we calculated the force per residue, *i.e.* per hydrogen bond and sidechain interaction, by normalizing the rupture force by the number of alanine groups in the strand (shown as black squares in Fig. 3B). We found a maximal force load of 427 pN/residue for  $\beta$ -sheets of 8 residues. The rupture force per residue in 4-residue-long crystals is 169 pN/residue, and thus less than half of the mechanical optimum. For longer  $\beta$ -sheets comprising 12 or 16 residues, the load carried per residue decreases again. Apparently, adding additional residues to an 8-residue long crystal does not lead to a significant gain in mechanical resistance. At even longer scales, we expect the total rupture force to reach a plateau.

What is a microscopic interpretation of the maximum force per residue for (Ala)<sub>8</sub>  $\beta$ -sheet crystals? Apparently, the external force, propagating along the pulled strand, is deflected vertically into

adjacent strands *via* hydrogen bonds and sidechain interactions such that the residues at a distance of 8 residues and more from the pulled terminus do not take part in carrying any significant strain. To test this interpretation, we performed force distribution analysis (FDA) to detect the internal propagation of stress in the structure prior to rupture.<sup>15</sup>

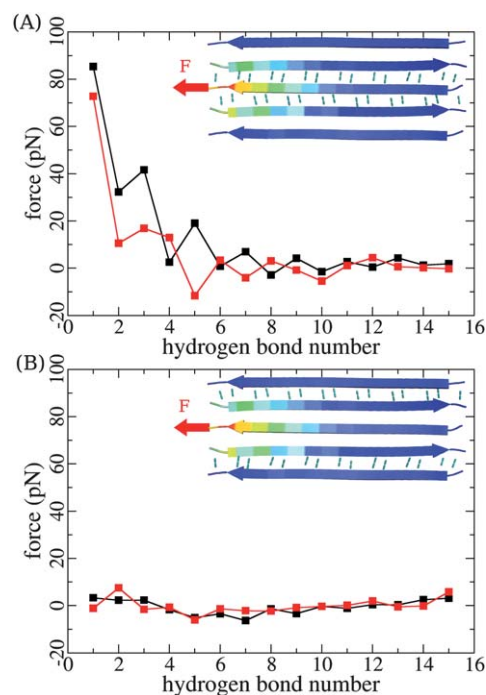
The main idea of the newly developed FDA is to analyze the inter-atomic forces as the structure is subjected to an external constant force in force-clamp molecular dynamics (FCMD) simulations. Each molecular structure was equilibrated in the absence of a force (relaxed state), and in presence of a constant force (stretched state). The atomic pair-wise forces monitored in the stretched state are averaged over time and subtracted from those forces observed in the relaxed state. Using a constant force has the advantage over a constant velocity that internal forces can be sampled at quasi-equilibrium without being loading rate dependent, as previously established for other systems.<sup>15</sup> The differences in internal forces due to the externally applied perturbation measure the strain distribution through the silk crystal, and were mapped onto the structure. FDA is reminiscent of finite element analysis performed for revealing the stress distribution in macroscopic objects. For more technical details, we refer to the original FDA publication.<sup>15</sup>

Here, we subjected the (Ala)<sub>16</sub> crystal to MD simulations in the absence of an external force, and in the presence of a constant force of 1660 pN. This external force is low enough to keep the structure intact within the nanosecond time scale of the simulations. We then calculated the internal force distribution in the (Ala)<sub>16</sub> crystal from the difference in the atomic pair-wise force between the resulting strained and relaxed state. The atomic forces have been averaged over 10 ns simulations for each state. The details of this method are published elsewhere.<sup>15</sup>

The forces carried by individual hydrogen bonds formed between the pulled central strand and its two adjacent strands in the central  $\beta$ -sheet layer are shown in Fig. 4A. The hydrogen bond forces, also shown as color codes on the  $\beta$ -sheet structure, are maximal at the point of force application, and decay along the strands. They reach almost zero around the eighth residue in the central strand. Residues more than 8 residues away from the point of force application are virtually not involved in the force distribution. Similarly, the next rows of hydrogen bonds in the structure do not carry any significant load (Fig. 4B). On the basis of this central result of the FDA, we predict the addition of alanine residues to strands of roughly 8 residues or longer not to add to the  $\beta$ -sheet crystal's mechanical stability. This is in agreement with and gives a molecular explanation for our results from force-probe MD simulations above. We conclude that  $\beta$ -strands of 8 residues in length are the most efficient building block of mechanically resistant structures.

We note that we obtained the consistent result of  $\sim 8$  residues as the optimal  $\beta$ -sheet length from two complementary, yet very distinct simulations techniques, namely FPMD and FDA. While in FPMD simulations, a constant loading rate leading to forces as high as 6 nN is employed, FDA makes use of a constant and significantly lower force (1.66 nN). This underlines that our finding is independent from the details of force application, but instead is largely determined by the nature of the molecular structure.

On the basis of the above results, we can explain the length of  $\beta$ -strands found in spider dragline silk. The segment of the sequence forming  $\beta$ -sheet crystals is poly(Ala)<sub>8</sub>. Apparently, during evolution, spiders optimized the sequence such that the resource of the natural



**Fig. 4** Force distribution in the hydrogen bonds of a  $\beta$ -sheet layer. (A) Forces taken up by the hydrogen bonds between the central pulled strand and the two adjacent strands, as indicated as dashed lines in the structure. The upper and lower strand hydrogen bonds are shown in red and black, respectively. The same data is shown as a color code mapped onto the  $\beta$ -sheet structure, where red indicates high and blue indicates low strain. (B) Force distribution of hydrogen bonds between the outer strands of the  $\beta$ -sheet layer. Otherwise same as (A).

material silk is most efficiently used, resulting in the maximal resilience of the crystalline subunits at a given sequence length. Similarly, intra and extra-cellular proteins with mechanical function often feature  $\beta$ -sheet sandwich domains with  $\beta$ -strands in the range of 8 residues. Among others, the immunoglobulin domains of the muscle protein titin are among the most mechanically resistant domains known and consist of strands typically showing a length of  $6 \pm 1$  residues as shown in Fig. 2. We note that the optimum of 8 residues is an approximate result. Differences in the rupture forces conferred per residue for 6 and 12 residues, though significant, are not large and might depend on the particular sequence and packing geometry. However, in contrast to earlier results,<sup>10</sup> we clearly show that  $\beta$ -sheets formed from four residues or even less perform poorly in comparison and as a consequence have not been the length of choice during evolution.

In conclusion, we compared the mechanical efficiency of  $\beta$ -sheet crystals of different strand lengths in terms of rupture force per residue and found  $\beta$ -strands of around 8 residues in length to be the optimum. Our results explain the prevalence of  $\beta$ -strands  $\sim 8$  residues in length in silk. For synthetic silk-mimetic polymers, we predict 8 residues, *i.e.* hydrogen bonds, in the crystalline phase an optimal choice for the cross-linking and force-bearing properties of the crystals.

## References

- 1 M. Pfuhl and A. Pastore, *Structure*, 1995, **3**, 391–401.
- 2 D. Bisig, P. Weber, L. Vaughan, K. H. Winterhalter and K. Piontek, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 1999, **55**, 1069–1073.

- 
- 3 J. T. Nielsen, M. Bjerring, M. D. Jeppesen, R. O. Pedersen, J. M. Pedersen, K. L. Hein, T. Vosegaard, T. Skrydstrup, D. E. Otzen and N. C. Nielsen, *Angew. Chem., Int. Ed.*, 2009, **48**, 2118–2121.
  - 4 P. A. Guerette, D. G. Ginzinger, B. H. Weber and J. M. Gosline, *Science*, 1996, **272**, 112–115.
  - 5 Y. Takahashi, M. Gehoh and K. Yuzuriha, *Int. J. Biol. Macromol.*, 1999, **24**, 127–138.
  - 6 A. G. Murzin, S. E. Brenner, T. Hubbard and C. Chothia, *J Mol Biol*, 1995, **247**, 536–540.
  - 7 D. Frishman and P. Argos, *Proteins: Struct., Funct., Genet.*, 1995, **23**, 566–579.
  - 8 W. Kabsch and C. Sander, *Biopolymers*, 1983, **22**, 2577–2637.
  - 9 J. M. Gosline, P. A. Guerette, C. S. Ortlepp and K. N. Savage, *J. Exp. Biol.*, 1999, **202**, 3295–3303.
  - 10 S. Keten and M. J. Buehler, *Nano Lett.*, 2008, **8**, 743–748.
  - 11 S. Xiao, W. Stacklies, M. Cetinkaya, B. Markert and F. Gräter, *Biophys. J.*, 2009, **96**, 3997–4005.
  - 12 I. Krasnov, I. Diddens, N. Hauptmann, G. Helms, M. Ogurreck, T. Seydel, S. S. Funari and M. Müller, *Phys. Rev. Lett.*, 2008, **100**, 048104–048104.
  - 13 H. Grubmüller, B. Heymann and P. Tavan, *Science*, 1996, **271**, 997–999.
  - 14 B. Hess, C. Kutzner, D. van der Spoel and E. Lindahl, *J. Chem. Theory Comput.*, 2008, **4**, 435–447.
  - 15 W. Stacklies, M. C. Vega, M. Wilmanns and F. Gräter, *PLoS Comput. Biol.*, 2009, **5**(3), e1000306.