This is a repository copy of *Intermolecular Packing in B. mori Silk Fibroin: Multinuclear NMR Study of the Model Peptide (Ala-Gly)15 Defines a Heterogeneous Antiparallel Antipolar Mode of Assembly in the Silk II Form.*

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/85871/

Version: Accepted Version

**Article:**

https://doi.org/10.1021/ma502191g

**Reuse**
Unless indicated otherwise, fulltext items are protected by copyright with all rights reserved. The copyright exception in section 29 of the Copyright, Designs and Patents Act 1988 allows the making of a single copy solely for the purpose of non-commercial research or private study within the limits of fair dealing. The publisher or other rights-holder may allow further reproduction and re-use of this version - refer to the White Rose Research Online record for this item. Where records identify the publisher as the copyright holder, users can verify any specific terms of use on the publisher’s website.

**Takedown**
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.
ABSTRACT: We have previously suggested that crystalline Bombyx mori silk in silk II form (the silk structure after spinning) is not a simple antiparallel β-sheet but is intrinsically heterogeneous. Using the peptide (AG)$_{15}$, we have obtained the first fully assigned high resolution solid state $^1$H NMR spectrum. Distinct heterogeneity was observed, in both $^1$H and $^{13}$C CP/MAS signals. Based on these results, a new model is proposed that contains two different packing arrangements of antiparallel β-sheets. The structures were energetically minimized by CASTEP calculation and used to calculate the solid state $^1$H, $^{13}$C, and $^{15}$N NMR chemical shifts using the GIPAW method. This new model was supported by good agreement between the calculated and observed $^1$H, $^{13}$C, and $^{15}$N chemical shifts and relative $^1$H--$^1$H proximities obtained from 2D $^1$H DQMAS experiments. We conclude that the intermolecular packing of B. mori silk fibroin has been finally resolved.

Because of the exceptional strength and toughness of the Bombyx mori (silkworm) silk fiber, and in view of increasing applications in the area of biomaterials, much attention has been paid to the structure of silk fibroin. Two crystalline forms, Silk I and Silk II, have been reported as dimorphs, essentially representing the regular domains of fibroin before and after spinning. By using several solid state NMR techniques, the Silk I form (as stored in the B. mori silkworm and dried under mild conditions) has been shown to possess a repeated type II β-turn structure. On the other hand, the precise intermolecular packing in the Silk II form (representing the core of the spun silk fiber) has not yet been determined. Using X-ray fiber diffraction of the crystalline region, the structure of Silk II was first characterized by Marsh, Corey, and Pauling as a regular array of antiparallel β-sheets: this structure remains the classic image of β-sheet silk. We call this model the “Marsh model”. Later, Fraser et al. Lotz and Keith, and Fossey et al. supported the general features of this antiparallel β-sheet model, but some of them also noted an irregular structure to be present in the silk fibers. Takahashi et al. proposed that a crystal site is statistically occupied by either of two antiparallel β-sheet chains with different relative orientations, in a 2:1 ratio, based on X-ray diffraction analysis of silk fibers. The latter analysis is more detailed and based on better data than the “Marsh model”. We call the model by Takahashi et al. the “Takahashi model”. There are no further reports about B. mori silk fiber in Silk II form at atomic level since Takahashi’s paper.

The Takahashi model is a better fit to the experimental data than the Marsh model but is not consistent with the distances of the intermolecular hydrogen bonds between the NH--OC groups of Ala and Gly, as explained below. It is therefore high time to come up with a new comprehensive model for the silk fiber that can satisfy all of the currently contradictory analytical data.

In the present work, a precise model for the crystalline structure of B. mori silk fibroin in the Silk II form is presented.

Received: October 28, 2014
Revised: December 9, 2014

Supporting Information

Macromolecules
using a small (Ala-Gly)$_n$ peptide as the model. The alternating copolyptide (Ala-Gly)$_n$ has been generally accepted as a good model of the crystalline region, NMR spectra of (AG)$_n$ correspond closely to those obtained using the crystalline fraction of native silk II fibers, and the torsion angles of the straight backbone chains correspond to the typical angles of an antiparallel $\beta$-sheet. In previous $^{13}$C solid state NMR studies of (AG)$_n$ the $^{13}$C$\beta$ signal of the Ala residues has been reported to consist of three peaks. Recently, we have developed a 1 mm microcoil MAS NMR probe head for mass-limited solid samples. By combining the use of this microcoil probe head with ultrahigh field NMR at 920 MHz, we were able to obtain solid state $^1$H NMR spectra with excellent resolution for the (AG)$_15$ model peptide in the Silk I form as well as for several other related peptides.$^{9,21,22}$ Based on these advances, solid state $^1$H NMR can now be used to study the intermolecular arrangement of Silk II.

The key challenge lies in the ability to discern and resolve the two kinds of antiparallel $\beta$-sheet chains with different intermolecular packing arrangements, as detected here and in the earlier $^{13}$C CP/MAS NMR study.$^{15,16}$ We therefore carried out a search of packing arrangements, guided by crystallographic and NMR data; refined the resulting structures; and tested them against experimental data. The peptide (AG)$_n$ crystallizes in space group $P2_1$, a rectangular unit cell with the parameters $a = 9.38 \text{ Å}$, $b = 9.49 \text{ Å}$, and $c = 6.98 \text{ Å}$. The Marsh model places the molecular axis along $b$ but is otherwise very similar: $a = 9.40 \text{ Å}$, $b = 6.67 \text{ Å}$, and $c = 9.20 \text{ Å}$. In order to generate two kinds of $\beta$-sheet models with different intermolecular arrangements, we had the idea to calculate atomic coordinates for the chains, setting either $c$ or $b$ along the molecular axis. For each of these two models, energy optimization was performed. $^1$H, $^{13}$C, and $^{15}$N chemical shifts were then predicted for the two antiparallel $\beta$-sheet structures using the GIPAW method. Such GIPAW calculations have been widely applied to organic molecules, and their validity has been demonstrated by experimental solid state NMR analyses.$^{19,24-34}$ The $^{13}$C and $^{15}$N chemical shifts of Silk II are known from previous work and can thus be used to compare and validate the two different structural models based on their predicted chemical shift values.$^{35,36}$ The solid state $^1$H NMR chemical shift is particularly sensitive to the intermolecular packing arrangement of Silk II and could thus be used as a reliable tool to judge the validity of any previously proposed models and to propose a new intermolecular arrangement from this study.

**EXPERIMENTAL SECTION**

**Different Isotope-Labeled Peptides (AG)$_n$.** Isotope-labeled amino acids (2-3d-Ala, [3-13C]Ala, [U-13C]Gly, [U-13C]Ala) were purchased from Cambridge Isotope Laboratories Inc., Andover, MA. The synthesis of (AG)$_{15}$ peptides was performed with standard solid-phase Fmoc chemistry on an Agape Automated Peptide Synthesizer (AAPTec, Louisville, KY). An Fmoc-Gly-PEG-PS resin was used, and the Fmoc amino acids were coupled with HATU. Peptides were cleaved from the resin by treatment with 90% TFA for 2 h at room temperature. The crude peptide was precipitated and washed repeatedly with cold diethyl ether. The precipitate collected by centrifugation was dried under vacuum and then treated with formic acid to obtain the Silk II form. Confirmation of the Silk II form was obtained from the Ala CP peak pattern in the $^{13}$C CP/MAS spectrum as reported previously.$^{13}$ The peptides synthesized here are summarized in Table 1. Samples (a) and (b) were used for $^1$H DQMAS experiments to study the intermolecular arrangement. Samples (c–e) were used for spectral assignments by double CP $^{13}$H–$^{13}$C experiments.

| Table 1. Overview of the Isotope-Labeled (AG)$_{15}$ Samples Prepared Here |
|------------------|------------------|
| (a)              | (AG)$_{15}$      |
| (b)              | ([2-3d]AG)$_{15}$|
| (c)              | (AG)$_{15}$[3-13C]AG(AG)$_{15}$ |
| (d)              | (AG)$_{15}$[U-13C]G(AG)$_{15}$ |
| (e)              | (AG)$_{15}$[U-13C]C[U-13C]G(AG)$_{15}$ |

**Solid State DQMAS $^1$H NMR and Double CP $^1$H–$^{13}$C Correlation NMR.** DQMAS (double-quantum magic angle spinning) $^1$H NMR and double CP (cross-polarization) $^1$H–$^{13}$C correlation NMR experiments were performed at a $^1$H resonance frequency of 920 MHz, using a JEOL JNM-ECA920 spectrometer equipped with a $^1$H-X $^{13}$C double resonance and ultrahighspeed MAS probe at the Institute for Molecular Science (IMS) in Okazaki, Japan.$^6$ The sample spinning speed was actively stabilized by a pneumatic solenoid valve such that the spinning fluctuations were less than $\pm 10$ Hz at a spinning rate of 70 kHz. The temperature of the samples increases due to friction under fast MAS and was estimated to be around 333 K at 70 kHz MAS according to Pb(NO$_3$)$_2$ temperature calibration. The $^1$H rf field strength for the excitation $\pi/2$ pulse (1.29 $\mu$s) was 194 kHz. The $^1$H chemical shift was referenced to the peak of silicon rubber and set to 0.12 ppm from TMS. For the $^1$H DQMAS measurement, a dipolar homonuclear homogeneous Hamiltonian double-quantum/single-quantum correlation experiment (DH-DQSQ) was employed.$^{15,16}$ The $\tau_r$ delay was optimized, giving 0.3 ms for maximum S/N. The DQMAS spectra were obtained every 32 scans at each period in the DQ domain, and the recycle delay was set to 2 s. For $^1$H detection in the double CP $^1$H–$^{13}$C correlation measurements, the pulse sequence $^{90°}$C$\psi$-$^1$H$\phi$-$^90°$C$\psi$-$^1$H$\phi$-$^90°$ was used.$^{39}$ Here, 90 is a $\pi/2$ pulse, CP is a 4 ms cross-polarization pulse with a 10% (first) and -10% (second) ramp of $^{13}$C, $t_1$ is the evolution period, $t_2$ is a 5 ms period for dephasing of transverse $^{13}$C magnetization and $^1$H magnetization suppression, and $t_3$ is the detection period. Superscripts H and C indicate $^1$H and $^{13}$C, and subscripts $x$, $y$, and $\phi$ indicate rf phases, with $\phi = x$ and $y$ for quadrature detection in $t_1$. The $^1$H decoupling amplitude during $t_1$ was 27 kHz. The spectrum was obtained after 64 scans at each period in the $y$ domain with 512 points.

**DARR $^{13}$C NMR.** The $^{13}$C DARR spectrum (dipolar assisted rotational resonance) of (AG)$_{15}$[U-13C]$_n$[U-13C]$_n$G(AG)$_{15}$ was obtained after 32 scans at a $^{13}$C resonance frequency of 400 MHz, using a JEOL ECX400 spectrometer at a spinning speed of 8 kHz with a 4 mm rotor. The $\pi/2$ pulse was 3.8 $\mu$s for $^{13}$C and 3.4 $\mu$s for $^1$H. TPPM $^{13}$H decoupling was performed with a contact time of 2 ms. The mixing time was 500 ms, with a relaxation delay of 2 s. The indirect dimension consisted of 256 data points.

**Construction of Two $\beta$-Sheet Models with Different Intermolecular Packing Arrangements.** The characteristic angles of $(\phi, \psi) = (-140°, 140°)$ for an antiparallel $\beta$-sheet structure were used for both Ala and Gly residues in straight (AG)$_n$ chains.$^{17}$ To make model 1, starting from the molecular arrangement of the Marsh model$^{18}$ viewed along its crystallographic b-axis (shown in Figure 1), we rotated strand b (see Figure 2) by 180° around its molecular axis and shifted it along the strand by one residue to change from polar to antiparabolic structure. Strands a and b were generated from a and b.
In order to avoid steric clash between strands a, b, and c, two molecules were generated from the lower two using the Marsh model, which all methyl groups in each sheet are pointing in the same direction. The Marsh model has this arrangement. Form (b) corresponds to model 1 and form (c) to model 2.

Figure 1. Marsh model of (AG)$_n$. The model is shown from three different orientations, with the relevant unit cell axes shown. Three β-sheet layers are shown. In the top layer, methyl groups are in magenta; in the middle layer they are in orange; and in the bottom layer they are in yellow. Interstrand hydrogen bonds are indicated for the central sheet. The directions of the strands are shown beneath panel (a), with the top strand in magenta and the central strand in orange. This structure corresponds to model (a) of Figure 2.

Figure 2. Possible arrangements for the four antiparallel β strands within the unit cell of a P2$_1$ space group, based on the Takahashi model. The strands are shown end-on. Strands a, b, and a’, b’ form antiparallel sheets linked by interstrand hydrogen bonds. An alternating (Ala-Gly)$_n$ structure has all methyl groups on each strand in the same direction, indicated by Me. (a) A polar arrangement, in which all methyl groups in each sheet are pointing in the same direction. The Marsh model has this topology. (b, c) Two alternative antipolar arrangements. The Takahashi models have this arrangement. Form (b) corresponds to model 1 and form (c) to model 2.

Figure 3. 1H DQMAS spectrum of (AG)$_5$ in the Silk II Form. By combining the use of a microcoil probe head with an ultrahigh-field NMR spectrometer at 920 MHz, we obtained a well-resolved solid state 1H NMR spectrum of (AG)$_5$. The 1H chemical shifts were assigned using a DQMAS 1H NMR experiment, as illustrated in Figure 3. From the high field to low field, the peaks are assigned as Ala Hβ, Gly Hα, Gly Hβ1, Gly Hβ2, Gly Hα, Ala Hα, Gly Hα, and Ala Hα.

NMR Chemical Shift Calculations. The chemical shifts of 1H, 13C, and 15N in the two antiparallel β-sheet structures with different intermolecular arrangements were calculated using the GIPAW method. The PBE approximation and “on the fly” pseudopotentials were used. The energy cutoff of the plane wave was set to 610 eV, and a $5 \times 2 \times 3$ Monkhorst-Pack k-point grid was used as described above. The chemical shift reference of the calculated chemical shifts was determined by minimizing the difference between the observed and calculated chemical shifts without changing the relative chemical shift differences between the peaks. The reference values were 30.51, 171.31, and 197.22 ppm for the 1H, 13C, and 15N nuclei, respectively. All calculations were carried out using the NMRCASTEP program.
separately and with a large chemical shift difference. This discrimination is generally feasible in the solid state, given the lack of motions around the backbone chains in silk fibroin. A more detailed assignment is performed below with the help of specifically isotope-labeled peptides, and the relative $\text{H}^1$−$\text{H}^1$ distances are measured and discussed in the last section.

**Determination of the $\text{H}^1$ and $\text{C}^{13}$ Chemical Shifts in the Heterogeneous Domains.** The $\text{C}^{13}−\text{C}^{13}$ DARR spectrum of $(\text{AG})_7[\text{U}−\text{C}^{13}]\text{A}[\text{U}−\text{C}^{13}]\text{G}(\text{AG})_7$ was obtained as shown in Figure 4. In agreement with our previous results, we see two well-resolved Ala Cβ peaks in an intensity ratio of approximately 2:1, which are named A and B, respectively, representing the two packing arrangements. From the correlations between these two Ala Cβ peaks and the Gly CO region, and based on the relative peak intensities, two peaks within the Gly CO signal could also be assigned as the A and B components. Within the Ala CO peak, on the other hand, there were no chemical shift differences resolved. Further assignment was obtained for the solid state NMR $\text{H}^1$ spectrum of $(\text{AG})_{15}$, a $\text{H}^1−\text{C}^{13}$ double CP spectrum of $(\text{AG})_7[\text{3−C}^{13}]$.

![Figure 4](image1.png)

**Figure 4.** $\text{C}^{13}−\text{C}^{13}$ DARR spectrum of $(\text{AG})_7[\text{U}−\text{C}^{13}]\text{A}[\text{U}−\text{C}^{13}]\text{G}(\text{AG})_7$ in the Silk II form. The inset shows the correlation between the CO and Ala Cβ region.

![Figure 5](image2.png)

**Figure 5.** Double CP $\text{H}^1−\text{C}^{13}$ spectrum of $(\text{AG})_7[\text{3−C}^{13}]\text{AG}(\text{AG})_7$ in the Silk II form, showing the correlations of Ala Cβ with Ala H/β and Hα.
The chemical shifts of the A and B components within the Ala Hβ and Hα peaks were determined from their correlation with the two well-resolved Ala Cβ signals. A small chemical shift difference of 0.3 ppm was clearly discernible in the Ala Hβ peak. Within the Ala Hα region, on the other hand, chemical shift differences were not resolved. Similarly, we used the 1H−13C double CP spectrum of (AG)₇A[U-13C]G(AG)₇ in Figure 6 to assign the two components A (3.9 ppm) and B (3.4 ppm) within the Gly Hα signal, while any chemical shift differences in the Gly Hα₂ region could not be resolved. The observed chemical shift data are summarized in Table 2.

AG(AG)₇ in the Silk II form was acquired as shown in Figure 5. The chemical shifts of the A and B components within the Ala Hβ and Hα peaks were determined from their correlation with the two well-resolved Ala Cβ signals. A small chemical shift difference of 0.3 ppm was clearly discernible in the Ala Hβ peak. Within the Ala Hα region, on the other hand, chemical shift differences were not resolved. Similarly, we used the 1H−13C double CP spectrum of (AG)₇A[U-13C]G(AG)₇ in Figure 6 to assign the two components A (3.9 ppm) and B (3.4 ppm) within the Gly Hα₁ signal, while any chemical shift differences in the Gly Hα₂ region could not be resolved. The observed chemical shift data are summarized in Table 2.

**Construction of Two Antiparallel β-Sheet Structures with Different Intermolecular Packing Arrangements.** It has been previously reported that the 13C CP/MAS NMR spectra of both the model peptide (AG)₁₅ as well as the natural Cp-fraction of B. mori silk fibroin in the Silk II form show a multicomponent Ala methyl peak. This Ala Cβ peak was resolved and assigned to three components, namely two different kinds of β-sheet structure (19.2 and 22.3 ppm), plus a distorted β-sheet and/or random coil conformation (16.1 ppm), the latter presumably originating from loops and turns at...
each end of the antiparallel crystalline regions. However, there has been no further discussion so far about the meaning of the first two peaks assigned to the two types of antiparallel β-sheet structures. The large chemical shift difference of about 3 ppm within the Ala Cβ peak cannot be interpreted in terms of different torsion angles for the Ala residue in the β-sheet region. Therefore, the chemical shift difference must be attributed to differences in the intermolecular packing of the β-

Figure 7. Model 1 (top row) and model 2 (bottom row), shown from three orthogonal orientations. The same color scheme is used as for Figure 1. Model 1 corresponds to structure (b) in Figure 2 and model 2 to structure (c). In model 1, the molecular axis is along the crystallographic axis c, and in model 2, the molecular axis is along the crystallographic axis b.

Figure 8. Stick spectra of the calculated and observed ¹H, ¹³C, and ¹⁵N chemical shifts (in ppm) for (AG)₁₅ in the intrinsically heterogeneous Silk II form. The observed shifts are colored red and blue to correspond to the set of peaks A and B, respectively, which are in an intensity ratio of approximately 2:1. The calculated shifts are colored green and orange for models 1 and 2, respectively.
Table 3. Closest $^1$H−$^1$H Distances of Protons Evaluated for the Two Different Models of the Silk II Structure$^a$

<table>
<thead>
<tr>
<th>Model</th>
<th>Gly Hα2</th>
<th>Gly Hα1</th>
<th>Ala Hα</th>
<th>Ala Hβ</th>
<th>Ala Hα</th>
<th>Ala Hβ</th>
<th>Ala Hα</th>
<th>Ala Hβ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marsh model</td>
<td>I</td>
<td>II</td>
<td>I</td>
<td>II</td>
<td>I</td>
<td>II</td>
<td>I</td>
<td>II</td>
</tr>
<tr>
<td>Gly Hα2</td>
<td>1.99</td>
<td>3.43</td>
<td>4.30</td>
<td>3.75</td>
<td>5.90</td>
<td>5.65</td>
<td>5.13</td>
<td>6.59</td>
</tr>
<tr>
<td>Gly Hα1</td>
<td>3.43</td>
<td>4.38</td>
<td>3.75</td>
<td>4.05</td>
<td>5.31</td>
<td>6.59</td>
<td>4.16</td>
<td>4.50</td>
</tr>
</tbody>
</table>

$^a$: two $^1$H nuclei in different strands that are located within the same β-sheet plane. II: two $^1$H nuclei in different strands that are located in neighboring β-sheet planes. III: two $^1$H nuclei within the same strand. The italic numbers indicate distances of less than 4 Å.

Figure 9. $^1$H DQMAS spectrum of deuterium-labeled ([2-d$_1$]AG)$_{15}$ in the Silk II form to remove potential overlap in the Hα region (cf. Figure 3). It is clear that there are no Gly Hα−Gly Hα peaks on the diagonal.

strands. In the Marsh model, all molecules are in identical environments. It is therefore not readily reconciled with the experimental data. We have previously tried to interpret these two peaks using the Takahashi model, which has two different packing arrangements within the crystal, occupied statistically in the ratio 2:1, and is therefore in much better agreement with the NMR data.$^{15,16}$ However, the Takahashi model is clearly not correct in detail. In particular, the interstrand NH−OC hydrogen bond lengths are 2.1 Å for Ala and 2.6 Å for Gly, whereas the experimentally observed Ala and Gly H$_N$ chemical shifts are both 8.7 ppm, implying hydrogen bond lengths of around 1.8 Å for both Ala and Gly.$^{11}$ We therefore explored alternative models based on the Takahashi model, but with better geometry.

Takahashi et al.$^{14}$ note that an antiparallel β-sheet composed of alternating Gly and Ala can be constructed in two ways: a $\text{polar}$ arrangement in which the backbone hydrogen bonds are Ala−Ala and Gly−Gly, and the methyl groups in one sheet are all pointing in the same direction; or an $\text{anti-polar}$ arrangement in which the backbone hydrogen bonds are Ala−Gly and, the methyl groups in a sheet alternate, pointing up in one strand and down in the next (Figure 2). They concluded that the crystallographic data fit an antipolar model better. By contrast, the Marsh model (Figure 1) is polar, which forces sheets to be alternately close and distant. We therefore constructed models with antipolar sheets.

Previously,$^{17}$ we determined the torsion angles in the B. mori silk fibroin fiber from solid state NMR orientational constraints to be $(-140^\circ, 142^\circ)$ for Ala and $(-139^\circ, 135^\circ)$ for Gly, within an experimental error of $\pm 5^\circ$. We thus used the typical β-sheet torsion angles of $(-140^\circ, 140^\circ)$ for both the Ala and Gly residues to generate model structures of (AG)$_{15}$. We also used the unit cell dimensions of the B. mori silk fibroin fiber as reported by Takahashi et al.$^{14}$ Given these constraints, the problem is limited to how one sheet packs on top of its neighbor. Any stereochemically viable model must have the strands in one sheet displaced by roughly half an interstrand spacing compared to its neighbor (Figure 2b,c).

On this basis, we constructed and refined two structural models with different intermolecular packing of the β-strands in the unit cell: model 1 and model 2 (Figure 7). Model 1 was consistently of slightly lower energy than model 2. Figure 7 shows both models with the central sheet in the same orientation to emphasize the difference in packing of the top sheet against the middle one. A key difference is that the Ala methyls are positioned differently. Model 1 has the packing shown in Figure 2b. The methyls of the top sheet that point down to the central sheet point roughly toward the Gly Hα, in the spaces between the pairs of interstrand Gly−Ala hydrogen bonds. By contrast, in model 2 (corresponding to Figure 2c), the methyls point to the center of the pair of interstrand Gly−Ala hydrogen bonds and are thus shifted along the strand by 3 Å relative to one residue. We note that both models were energy minimized against the crystal dimensions of Takahashi et al.$^{14}$ they are both consistent with the crystallographic data.

$^1$H, $^{13}$C, and $^{15}$N Chemical Shift Calculation of Model (AG)$_{15}$ Structures. $^1$H, $^{13}$C, and $^{15}$N chemical shifts were calculated for models 1 and 2 using GIPAW and are summarized in Table 2. The output files after CASTEP calculations are listed in the Supporting Information: Tables 1S and 2S (model 1 = A) and 2S (model = B). Figure 8 shows the corresponding stick spectra for the calculated and observed chemical shifts, from which it can be seen that the calculated shifts for model 1 fit the positions of experimental peaks A well, and calculated model 2 shifts fit experimental peaks B well, while the alternative assignment (model 1 = B and model 2 = A)
A) fits poorly. In particular, the covariance\textsuperscript{42} for 1 = A and 2 = B is 0.23 ppm\textsuperscript{2}, while for 1 = B and 2 = A it is 0.57 ppm\textsuperscript{2}, clearly indicating that the correct assignment is 1 = A and 2 = B. This pairing is in agreement with the calculated lower energy for model 1. We therefore equate model 1 with peaks A and model 2 with peaks B. The agreement between calculated and observed $^1$H shifts is reasonable. The $^1$H chemical shift calculation of model 1 makes it possible now to assign the two Gly Hα peaks. Namely, the Hα of the Gly residue observed at lower field at 4.6 ppm can be assigned to the Hα located in the β-sheet plane. This feature is important when we come to discuss the β-sheet assembly in the light of the DQMAS $^1$H NMR data, where the Ala methyl signal corresponding to model 1 was obtained at higher field than for model 2.

For the $^{13}$C chemical shifts, the agreement between the observed and calculated chemical shifts is excellent, given that the entire chemical shift range from the highest field Ala Cβ to the lowest field Ala CO peak could be well reproduced, and the chemical shift differences between the different carbons also agree very well. In addition, the peak of Ala Cβ was correctly found to appear at a higher field in model 1 than in model 2. The experimental finding that components A and B were not resolved in the $^{13}$C signals of Gly Cα, Ala Cα, and Ala CO is also in agreement with the small calculated chemical shift differences between models 1 and 2.

Finally, the calculated and previously observed $^{15}$N chemical shifts\textsuperscript{43} are compared for the two models. In this case the two peaks of Ala and Gly were well resolved, so their relative peak positions as well as the chemical shift difference could be compared. The agreement is also excellent, and the two peaks corresponding to models 1 and 2 could be assigned for both $^{15}$N nuclei. The goodness of fit for $^{13}$C and $^1$H can be compared to literature values. Using the assignment of model 1 = A and model 2 = B, the root-mean-square difference between calculated and observed shifts is 2.2 ppm for $^{13}$C and 0.68 ppm for $^1$H (or 0.4 ppm omitting amide protons, for which chemical shift calculations are particularly difficult because of their great sensitivity to hydrogen bonding). This can be compared to other comparisons between GIPAW calculations and experimental solid-state shifts for small organic compounds: 2.5 ppm for $^{13}$C and 0.3 ppm for $^1$H (penicillin G)\textsuperscript{44}, 3.4 ppm for $^{13}$C (testosterone)\textsuperscript{45}, an average of 3.1 ppm for $^{13}$C and 0.3 ppm for $^1$H (thymol).\textsuperscript{46} Thus, the chemical shift correlation overall reproduces the observed chemical shifts very well for all three nuclei, giving us confidence in the accuracy of the models. We therefore propose that B. mori (Ala-Gly)\textsubscript{n} silk II consists of antipolar antiparallel sheets arranged statistically in the arrangements shown in models 1 and 2, with a preference of about 2:1 for model 1 vs model 2.

Validation of the New Heterogeneous Model from DQMAS $^1$H NMR. A further test for the validity of the models derived here comes from $^1$H–$^1$H distances observed in DQMAS $^1$H NMR spectra, which typically must be within about 4 Å to give rise to observable cross-peaks.\textsuperscript{19} A set of nine $^1$H–$^1$H correlation signals is indicated in Figure 3. We examined the $^1$H–$^1$H distances underlying these observed $^1$H–$^1$H correlations by inspecting the list of $^1$H–$^1$H distances calculated from our models, and comparing them to the Marsh model. Particularly diagnostic are the $^1$H–$^1$H distances in which either Gly Hα1 or Gly Hα2 protons are involved, which are listed in Table 3. All distances calculated to be less than 4 Å in one or both models (underlined in Table 3) are present in the spectrum, as expected. By contrast, several distances that are very short in the Marsh model do not give rise to observable peaks in the spectrum, providing strong evidence that the Marsh model does not correspond with the experimental data: A given contact for Gly Hα1–Gly Hα1 or Gly Hα2–Gly Hα2 indicates a distance between two $^1$H nuclei that are located in different strands.

1. Cross-peak v is between Gly Hα2 and Ala Hα. This distance is very short in both models, but is longer than 4 Å in the Marsh model.

2. Cross-peaks viii and ix are from Ala Hβ to Gly Hα1 and Hα2. Both these distances are short in models 1 and 2. However, in the Marsh model these distances are both well over 4 Å.

3. In the Marsh model, the Gly Hα2 protons in adjacent β-strands are very close to one another, so a diagonal peak for Gly Hα2 should be detected. This feature, however, is difficult to judge from Figure 1 because the Gly Hα2 and Ala Hα+2 peaks overlap in the relevant spectral region. We therefore synthesized deuterium-labeled [(2-d)AG]\textsubscript{15} and acquired another DQMAS $^1$H NMR spectrum. As seen in Figure 9, there is clearly no Gly Hα2 peak on the diagonal, now that the Ala Hα signal at around 5.0 ppm has been removed. This observation provides very strong evidence that polar models, such as in the Marsh model, cannot be correct.

In summary, we have shown that B. mori (Ala-Gly)\textsubscript{n} silk II exists in two packing arrangements A and B in a ratio of approximately 2:1. We have presented two models (1 and 2, corresponding respectively to A and B), which fit all experimental data, in particular crystallographic, chemical shifts and $^1$H–$^1$H dipolar contacts. We have demonstrated that silk II must be an antipolar, not a polar, packing. We propose that crystalline Silk II is a statistical mixture of these two packing arrangements, in a ratio 2:1. (The coordinates of the new Silk II model are listed in the Supporting Information: Table 1S (model 1 = A) and 2S (model = B).)