

Supporting Information

Degree of Biomimicry of Artificial Spider Silk Spinning Assessed by NMR Spectroscopy

*Martins Otkovs, Marlene Andersson, Qiupin Jia, Kerstin Nordling, Qing Meng, Loren B. Andreas, Guido Pintacuda, Jan Johansson, Anna Rising, and Kristaps Jaudzems**

anie_201706649_sm_miscellaneous_information.pdf

Table of Contents

Experimental Procedures	2
Results and Discussion	3
Figure S1.....	3
Figure S2.....	4
Figure S3.....	5
Figure S4.....	6
Figure S5.....	7
Figure S6.....	7
Figure S7.....	7
Figure S8.....	8
Figure S9.....	8
Figure S10.....	8
References	9

Experimental Procedures

Structural analysis of NT2RepCT in solution. Solution NMR experiments were performed at 298 K on a 600 MHz Varian Unity Inova spectrometer equipped with an HCN triple-resonance pulsed-field-gradient cold probe or on an 800 MHz Bruker Avance III HD spectrometer equipped with an HCN triple-resonance pulsed-field-gradient room temperature probe (TXI). Variable pH spectra were recorded on a uniformly ^{13}C , ^{15}N labeled NT2RepCT sample dissolved in 20 mM sodium phosphate (NaPi), 20 mM NaCl, 0.02 % (w/v) NaN_3 , 5% D_2O (v/v). The pH was adjusted by addition of 1 M hydrochloric acid. Assignment of the 2Rep domain was performed at pH 6.7. The following 2D and 3D spectra were recorded: ^{15}N - ^1H -HSQC, HNCA, HNCO, HN(CA)CO, CBCA(CO)NH, HN(CA)NNH, HN(CO)N, ^{15}N -resolved NOESY-HSQC (60 ms mixing time), HNHA. $^3J_{\text{H}^{\alpha}\text{H}^{\beta}}$ couplings were determined from the relative intensities of diagonal and cross-peaks in the HNHA experiment.^[1] Correlation time τ_c was estimated from the ratio of T_1 and T_2 relaxation rates, using formula

$$\tau_c \approx 4\pi\omega_N \sqrt{6 \frac{T_1}{T_2} - 7},^{[2]}$$

where only $J(0)$ and $J(\omega_N)$ spectral density terms have been considered and higher frequency terms have been neglected. T_1 and T_2 rates were determined by 2D ^{15}N - ^1H TROSY-HSQC based T_1 and T_2 relaxation experiments^[3] with delays of 20, 100, 200, 400, 800, 1600, 3000 ms for T_1 and 10, 30, 50, 90, 150 and 250 ms for T_2 and using an interscan delay of 5 s. The $^{15}\text{N}\{^1\text{H}\}$ heteronuclear NOE experiments were recorded with a saturation period of 3.5 s and a total interscan delay of 5.0 s. Relaxation data were analyzed using the Relax software.^[4] The chemical shifts of the 2Rep domain in NT2RepCT at pH 6.7 have been deposited in the BioMagResBank with the accession number 27147.

Structural analysis of unlabeled NT2RepCT fibers. Solid-state ^1H - ^{13}C cross-polarization magic-angle spinning (CP-MAS) and 2D ^1H - ^{13}C heteronuclear correlation (HETCOR) NMR spectra for unlabeled fibers were recorded on an 800 MHz Bruker Avance III HD spectrometer equipped with a 3.2-mm $^{13}\text{C}/^{15}\text{N}\{^1\text{H}\}$ E-free MAS probe. The MAS frequency was 15 kHz and the temperature was regulated to 298 K. Cross-polarization (CP) from ^1H to ^{13}C was performed using a linear ramp from 32.2 to 64.4 kHz on ^1H , while the ^{13}C amplitude was held constant at 66.3 kHz. The CP was optimized at a contact time of 2 ms for ^1H - ^{13}C CP-MAS and 0.3 ms for 2D ^1H - ^{13}C HETCOR, respectively. Spinal64 decoupling^[5] was applied during acquisition with a ^1H radio-frequency (RF) field amplitude of 92.3 kHz and a pulse length of 5 μs . The acquisition time was 20 ms. The recycle delay was 5 s. As-spun fibers were stored in buffer (20 mM sodium acetate, pH 5.0) until they were packed into a 3.2 mm rotor.

Preparation of selectively labeled NT2RepCT fibers. Stock solutions of amino acids were prepared in H_2O : 100 mg/ml for Ala, Arg, Gly, Lys Pro and Ser; 50 mg/ml for His, Thr, Val and Met; 25 mg/ml for Asn, Cys and Gln; and 20 mg/ml for Ile and Phe. Trp, Asp, Glu, Leu and Tyr could not be dissolved in H_2O and were added to the minimal medium as powders. For protein expression 0.5 g of each amino acid, except Tyr (for CT labeling) or Met/Phe/Trp (for NT labeling), respectively, was added to 500 ml of minimal medium containing ^{15}N labeled NH_4Cl and ^{13}C labeled glucose. Although unlabeled Met and Trp were not added in the medium for NT labeling, their aliphatic carbons are derived from serine and aspartate during biosynthesis in *E. coli*. As a result, the C_α and C_β atoms of Trp and all C atoms of Met were not labeled. NT2RepCT was purified and spun into fibers as described.^[6]

Structural analysis of selectively isotope enriched fibers. Solid-state NMR spectra for selectively isotope enriched fibers were recorded on 500 or 700 MHz Bruker Avance III spectrometers equipped, respectively, with a double resonance H/X or a triple-resonance HCN 1.3 mm MAS probe. The MAS frequency was 60 kHz. The temperature of the sample was regulated using a VT gas flow in the range of 230–243 K, to reach an estimated sample temperature of 300 K, using the chemical shift of KBr.^[7] The non-selective 90° pulses were set to 2.5 μs (^1H), 3.5 μs (^{15}N), and 3.5 μs (^{13}C). Forward and back CP from ^1H to ^{13}C were performed

SUPPORTING INFORMATION

using a linear ramp from 61.4 to 67.8 kHz on ^1H , while the ^{13}C amplitude was held constant at 16.6 kHz. The CP contact time for ^1H - ^{13}C CP and 2D ^{13}C - ^{13}C RFDR experiments was 2 ms. For 2D ^{13}C - ^1H CP-HSQC and (H)CCH-TOBSY^[8] experiments, the forward CP used a contact time of 0.6 ms, while the back CP used a contact time of 0.3 ms to select predominantly 1-bond transfer. Heteronuclear decoupling was implemented with WALTZ-16 at a 10 kHz RF-field amplitude, and applied on the ^1H , ^{13}C , and ^{15}N channels, specifically on ^1H during acquisition of the ^1H - ^{13}C CP spectrum, and during t1 and acquisition of the 2D ^{13}C - ^{13}C RFDR spectrum, as well as on the ^{13}C and ^{15}N channels during acquisition of proton. The MISSISSIPPI pulse sequence^[9] was used without homospoil gradients to suppress the water signal, with a 15 kHz irradiation for 100 ms. The RFDR mixing time was 2.67 ms. The TOBSY mixing time was 7.2 ms. The ^{13}C and ^1H chemical shifts were referenced externally relative to adamantane. The ^{13}C and ^1H chemical shifts of the adamantane CH_2 moiety relative to TMS were 38.48 ppm and 1.2 ppm, respectively. For comparison with the solid-state ^{13}C NMR chemical shifts, the values derived from solution NMR statistical data^[10] (listed in Table 1) were corrected for a calibration offset of -2.01 ppm due to the different reference standards (adamantane and DSS, respectively).^[11] All spectra were processed with Bruker Topspin 3.5 and analyzed either in Bruker Topspin 3.5 or in CARA.^[12]

Results and Discussion

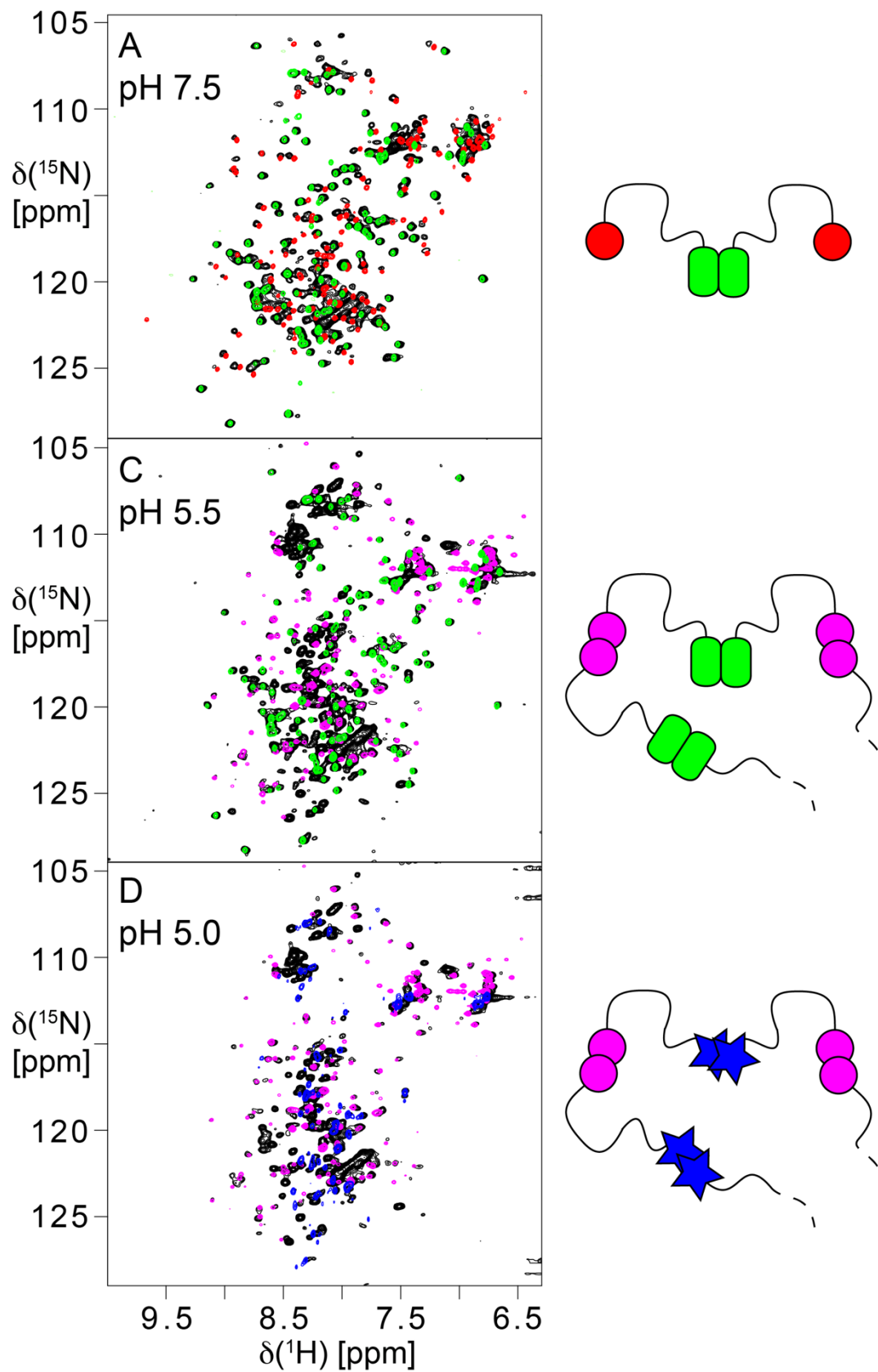
Figure S1. Domain structure and amino acid sequence of the NT2RepCT minispidroin. The poly-Ala segments are underlined. Residues derived from the expression vector and cloning sites are shown in grey.



MGHHHHHMS HTTPWNPGL AENFMNSFMQ GLSSMPGFTA SQLDDMSTIA QSMVQSIQSL AAQGRTPNK LQALNMAFAS SMAEIAASEE GGGSLSTKTS
 SIASAMSNAP LQTTGVVNQP FINEITQLVS MFAQAGMNDV SAGNSGRGQG GYGQSGGNA AAAAAAAAAA AAAAGQGQG GYGRSQGAG SAAAAAAAAA
AAAAAGSGQG GYGGQGGY GQSGNSVTSG GYGYGTSAAA GAGVAAGSYA GAVNRLSSAE AASRVSSNIA ATASGGASAL PSVISNIYSG VVASGVSSNE
 ALIQALLELL SALVHVLSSA SIGNVSSVGV DSTLNVVQDS VGQYVG

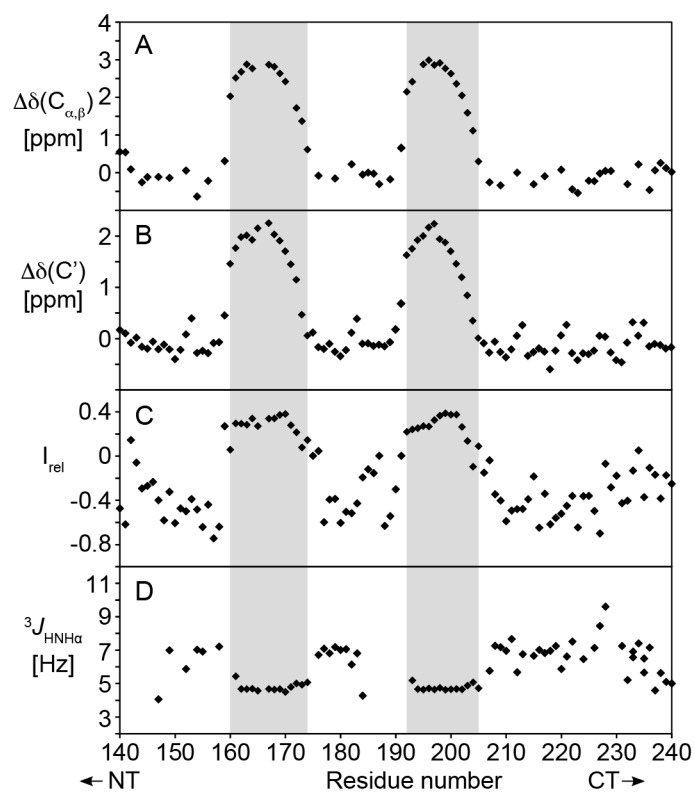
SUPPORTING INFORMATION

Figure S2. Enlarged spectral superpositions shown in Figure 1A, C and D. (A) Spectrum of NT2RepCT (black) at pH 7.5 overlaid with the spectra of monomeric NT (red) and folded dimeric CT (green). (C) Spectrum of NT2RepCT (black) at pH 5.5 overlaid with the spectra of dimeric NT (magenta) and folded dimeric CT (green). (D) Spectrum of NT2RepCT (black) at pH 5.0 overlaid with the spectra of dimeric NT (magenta) and unfolded CT at pH 5.0 (blue). A schematic presentation of the individual domains in NT2RepCT is shown on the right, where red/magenta circles represent the monomeric/dimeric NT, the black lines represent the 2Rep part, the green rounded rectangles represent the folded CT and the blue stars represent the unfolded CT.



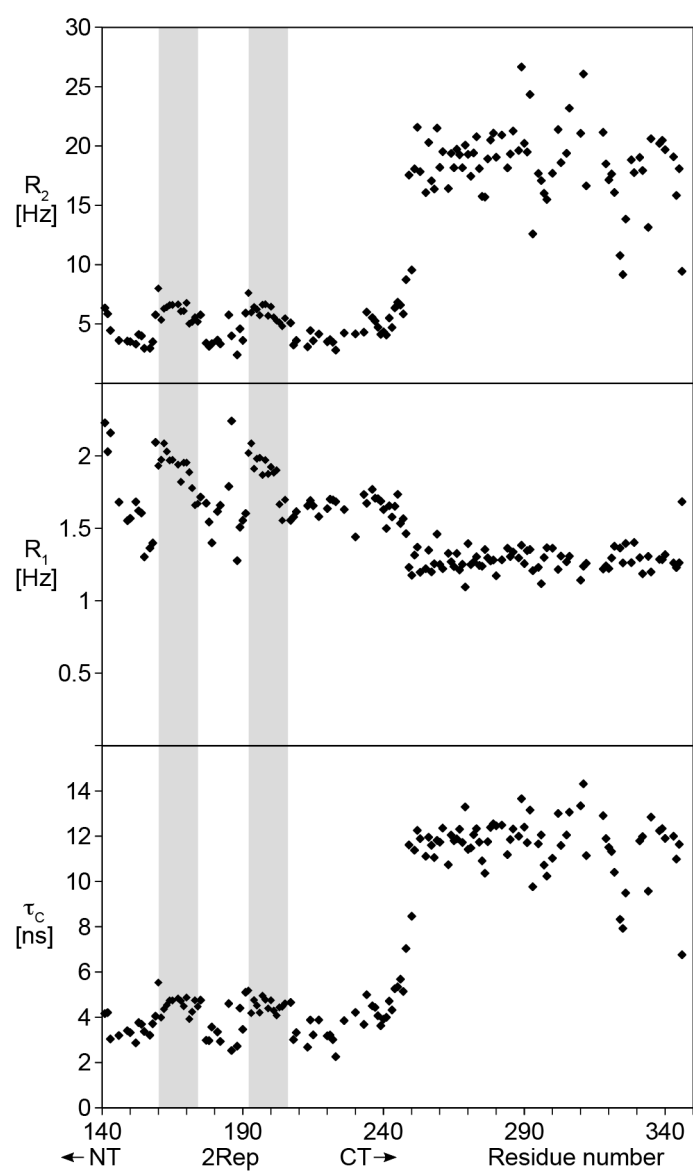
SUPPORTING INFORMATION

Figure S3. Secondary structure analysis of the 2Rep domain in NT2RepCT at pH 6.7 based on $C_{\alpha,\beta}$ and C' chemical shift deviations from random coil values (A, B), backbone amide $^{15}\text{N}\{^1\text{H}\}$ -NOE values (C) and $^3J_{\text{H}^{\text{N}}\text{H}^{\alpha}}$ coupling constants (D). The poly-Ala segments are shaded in grey and the C- and N-termini of the terminal domains are indicated below the plot.

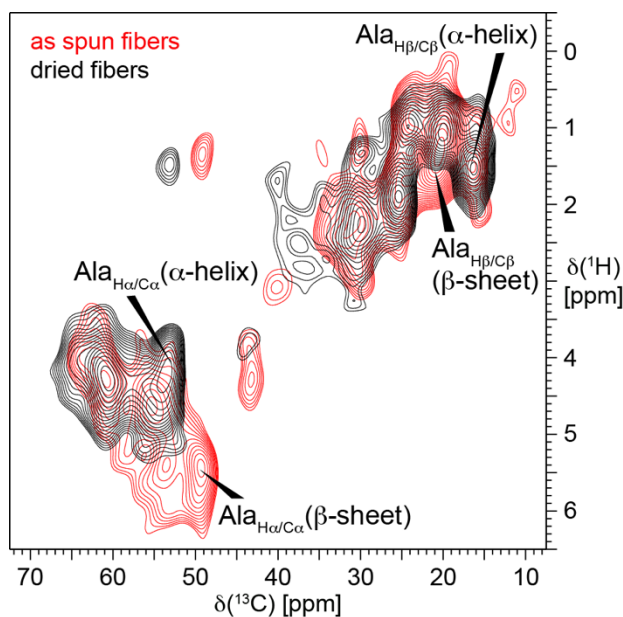
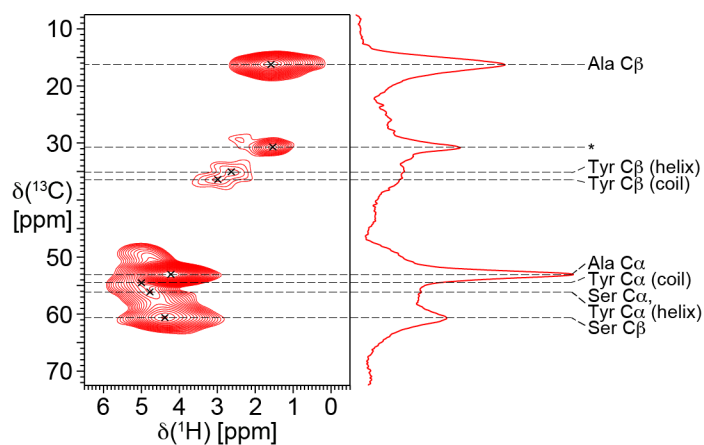
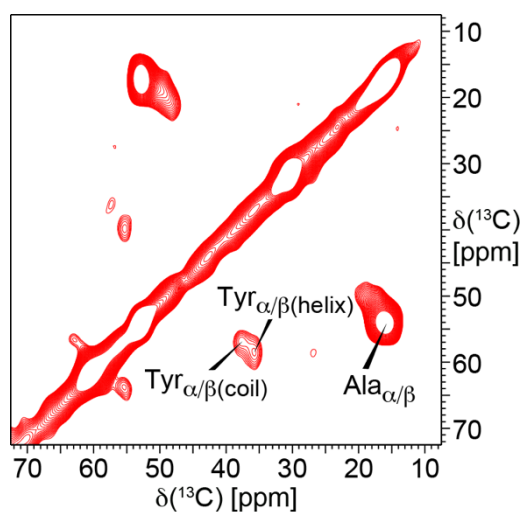


SUPPORTING INFORMATION

Figure S4. Relaxation analysis of 2Rep and CT in NT2RepCT at pH 6.7. The three panels show (from top to bottom) backbone amide ^{15}N R_2 and R_1 relaxation profiles and calculated correlation time τ_c . The positions of the domains are indicated below the plot.



SUPPORTING INFORMATION

Figure S5. Overlay of 2D ^1H - ^{13}C HETCOR spectra of as-spun and dried NT2RepCT fibers.**Figure S6.** 2D ^{13}C - ^1H CP-HSQC spectrum of Y-reverse labeled NT2RepCT fibers. The ^1H - ^{13}C CP-MAS spectrum with peak assignments is shown on the right, * denotes an impurity.**Figure S7.** 2D ^{13}C - ^{13}C RFDR spectrum of Y-reverse labeled NT2RepCT fibers.

SUPPORTING INFORMATION

Figure S8. 2D ^{13}C - ^1H CP-HSQC spectrum of F-reverse labeled NT2RepCT fibers. The ^1H - ^{13}C CP-MAS spectrum with peak assignments is shown on the right, * denotes an impurity.

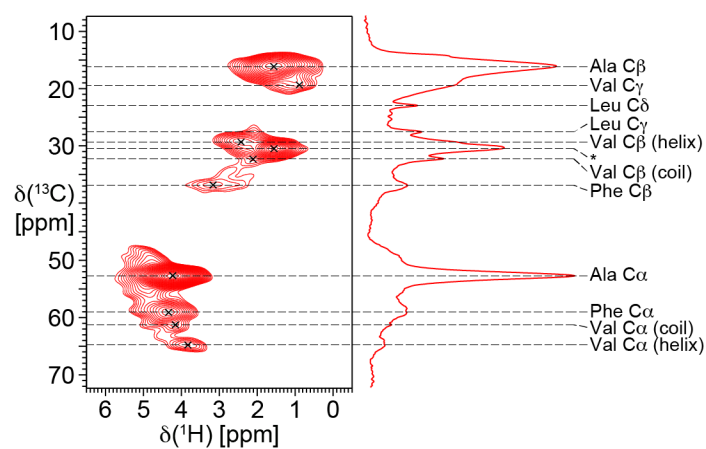


Figure S9. 2D ^{13}C - ^{13}C projection of a 3D (H)CCH TOBSY spectrum of F-reverse labeled NT2RepCT fibers.

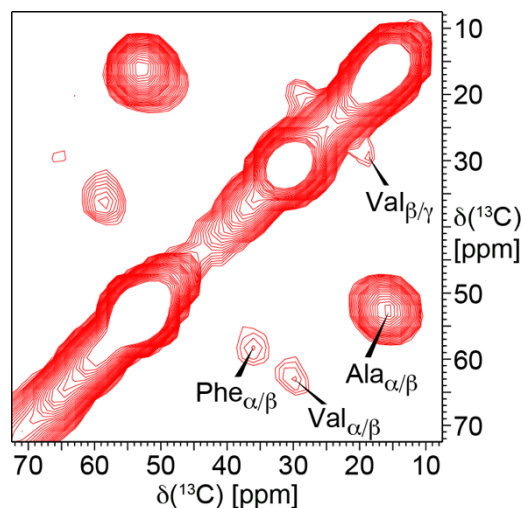
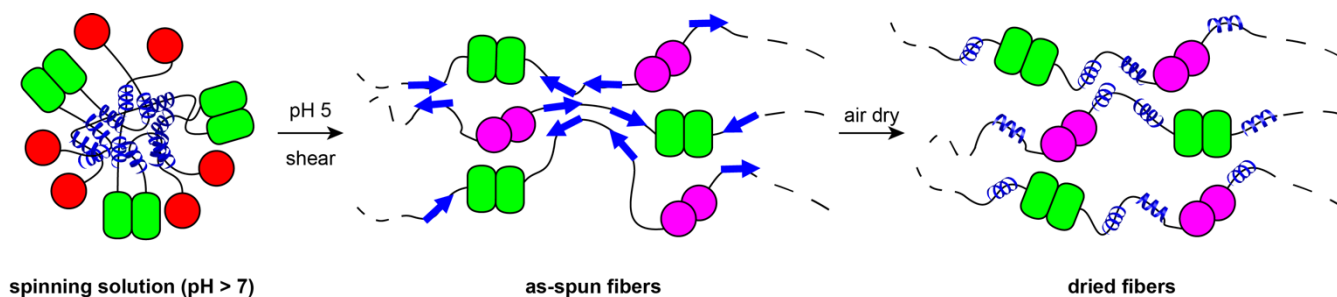


Figure S10. Schematic presentation of the structural changes experienced by the different domains of NT2RepCT upon fiber spinning and drying. Red/magenta circles represent the monomeric/dimeric NT, the black lines represent the 2Rep part, the green rounded rectangles represent the folded CT and the blue helices/arrows represent the poly-Ala segments in α -helical/ β -sheet conformation.



SUPPORTING INFORMATION

References

- [1] G. W. Vuister, A. Bax, *J. Am. Chem. Soc.* **1993**, *115*, 7772-7777.
- [2] L. E. Kay, D. A. Torchia, A. Bax, *Biochemistry* **1989**, *28*, 8972-8979.
- [3] a) G. Zhu, Y. L. Xia, L. K. Nicholson, K. H. Sze, *J. Magn. Reson.* **2000**, *143*, 423-426; b) C. Renner, M. Schleicher, L. Moroder, T. A. Holak, *J. Biomol. NMR.* **2002**, *23*, 23-33.
- [4] a) E. J. d'Auvergne, P. R. Gooley, *J. Biomol. NMR* **2008**, *40*, 107-119; b) E. J. d'Auvergne, P. R. Gooley, *J. Biomol. NMR* **2008**, *40*, 121-133.
- [5] B. M. Fung, A. K. Khitrin, K. Ermolaev, *J. Magn. Reson.* **2000**, *142*, 97-101.
- [6] M. Andersson, Q. Jia, A. Abella, X. Y. Lee, M. Landreh, P. Purhonen, H. Hebert, M. Tenje, C. V. Robinson, Q. Meng, G. R. Plaza, J. Johansson, A. Rising, *Nat. Chem. Biol.* **2017**, *13*, 262-264.
- [7] K. R. Thurber, R. Tycko, *J. Magn. Reson.* **2009**, *196*, 84-87.
- [8] M. Baldus, B. H. Meier, *J. Magn. Reson.* **1996**, *A121*, 65-69.
- [9] D. H. Zhou, C. M. Rienstra, *J. Magn. Reson.* **2008**, *192*, 167-172.
- [10] Y. Wang, O. Jardetzky, *Protein Sci.* **2002**, *11*, 852-861.
- [11] C. R. Morcombe, K. W. Zilm, *J. Magn. Reson.* **2003**, *162*, 479-486.
- [12] R. Keller, Cantina Verlag, Goldau, **2004**.