CIHR Canadian Institutes of Health Research
IRSC Instituts de recherche en santé du Canada
Submitted by CIHR
Déposé par les IRSC

Biochemistry. Author manuscript; available in PMC 2018 January 16.

Published in final edited form as:
Biochemistry. 2016 May 31; 55(21): 3048–3059. doi:10.1021/acs.biochem.6b00429.

Tracking Transitions in Spider Wrapping Silk Conformation and Dynamics by ^{19}F Nuclear Magnetic Resonance Spectroscopy

Muzaddid Sarker[†], **Kathleen E. Orrell**[†], **Lingling Xu**[†], **Marie-Laurence Tremblay**[†], **Jessi J. Bak**[†], **Xiang-Qin Liu**[†], and **Jan K. Rainey**[†,‡,*]

[†]Department of Biochemistry & Molecular Biology, Dalhousie University, Halifax, Nova Scotia, B3H 4R2, Canada

[‡]Department of Chemistry, Dalhousie University, Halifax, Nova Scotia, B3H 4R2, Canada

Abstract

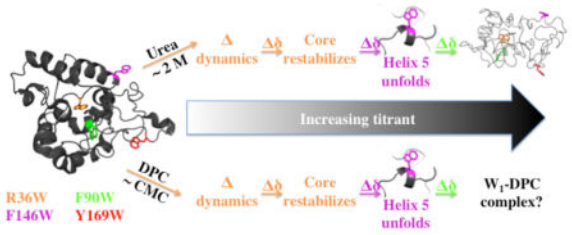
Aciniform silk protein (AcSp1) is the primary component of wrapping silk, the toughest of the spider silks due to combined high tensile strength and extensibility. *Argiope trifasciata* AcSp1 contains a core repetitive domain with at least 14 homogeneous 200-amino acid units ("W" units). Upon fibrillogenesis, AcSp1 converts from an α-helix-rich soluble state to a mixed α-helical/β-sheet conformation. Solution-state nuclear magnetic resonance (NMR) spectroscopy allowed demonstration of variable local stability within the W-unit, but comprehensive characterization was confounded by spectral overlap, exacerbated by decreased chemical shift dispersion upon denaturation. Here, ^{19}F-NMR spectroscopy, in the context of a single W-unit (W_1), is applied to track changes in structure and dynamics. Four strategic positions in the W-unit were mutated to tryptophan and biosynthetically labeled with 5-fluorotryptophan (5F-Trp). Simulated annealing-based structure calculations implied that these substitutions should be tolerated while circular dichroism (CD) spectroscopy and ^{1}H-^{15}N chemical shift displacements indicated minimal structural perturbation in W_1 mutants. Fiber formation by W_2 concatemers containing 5F-Trp substitutions in both W-units demonstrated retention of functionality, a somewhat surprising finding in light of sequence conservation between species. Each 5F-Trp-labeled W_1 exhibited a unique ^{19}F chemical shift, linewidth, longitudinal relaxation time constant (T_1), and solvent isotope shift. Perturbation to ^{19}F chemical shift and nuclear spin relaxation parameters reflected changes in conformation and dynamics at each 5F-Trp site upon addition of urea and dodecylphosphocholine (DPC). ^{19}F-NMR spectroscopy allowed unambiguous localized tracking throughout titration with each perturbant, demonstrating distinct behavior for each perturbant not previously revealed by heteronuclear NMR experiments.

Graphical abstract

[*]Corresponding Author. jan.rainey@dal.ca; phone: (902) 494-4632.
Author Contributions: M.S. and K.E.O. contributed equally to this work.

Supporting Information Available.
Table S1: Oligonucleotide primers for site-directed mutagenesis; Table S2: Details pertaining to simulated annealing structure calculations for W_1 mutants; Figure S1: W_1 protein expression and purification elaborated by SDS-PAGE; Figure S2: ^{1}H-NMR spectra of 5F-Trp mutants; Figures S3–S6: Chemical shift displacements for each W_1 mutant relative to WT; Figure S7: SDS-PAGE of purified W_2 mutants; Figures S8–S9: ^{19}F-NMR spectra (Figure S8) and relaxation properties (Figure S9) under titration with urea and DPC; Figure S10: R36W positioning; Figure S11: ^{1}H-^{15}N chemical shift displacements for WT W_1 in vicinity of each mutant under titration with urea and DPC.

CIHR Author Manuscript
CIHR Author Manuscript
CIHR Author Manuscript



Spider silks have long generated significant interest as biomaterials because of their exceptionally high tensile strength, extensibility, toughness, and biocompatibility.[1–3] Correspondingly, spider silks have many potential applications, including functioning as drug delivery vehicles,[4] in nerve regeneration,[5] and in tendon repair.[6] Spiders produce up to seven different protein-based silks, each with distinct mechanical properties.[7] Aciniform, or wrapping, silk is the toughest spider silk (i.e., it absorbs maximal mechanical energy prior to failing), and is used by spiders to wrap and immobilize prey and to build sperm webs, egg cases and stabilimenta.[8] The toughness of aciniform silk per unit weight surpasses even the toughest synthetic fibers available to date, such as nylon, Kevlar and high-tensile steel[9] but it is as light as cotton or nylon.[10]

Aciniform silk is composed primarily of the protein aciniform spidroin 1 (AcSp1). For AcSp1 proteins characterized to date, sizes depend upon the spider species and range from ~300–430 kDa, comprising a core repetitive domain flanked by short N- and C-terminal domains.[8] In *Argiope trifasciata*, the AcSp1 core repetitive domain is composed of a 200-amino acid unit (referred to here as the “W” unit, for wrapping) repeated at least 14 times.[8,11] Using solution-state nuclear magnetic resonance (NMR) spectroscopy, we recently demonstrated that a single W unit (W_1) consists of an ellipsoidally shaped, compactly folded five α-helix bundle flanked by intrinsically disordered N- and C-terminal tails.[12] Concatemers of two to four W units (W_2, W_3 and W_4) have been shown to be highly modular in terms of both structure and dynamics, with the linked C- and N-terminal tails of neighboring subunits retaining intrinsic disorder and giving rise to a “beads-on-a-string” architecture.[12]

Recombinant AcSp1 concatemers comprising at least two W units (i.e., W_2 and larger) can be induced to form silk-like fibers upon application of shear force.[13] The fact that fibers can be pulled directly from W_2 solution-state NMR samples implies that the observed beads-on-a-string concatemeric conformation acts as the soluble precursor for fibrillogenesis.[12] Fiber formation is accompanied by a concomitant secondary structural transition, with partial - but not complete - α-helix to β-sheet conversion.[12,13] This process can be inhibited by the addition of the zwitterionic detergent dodecylphosphocholine (DPC) above its critical micelle concentration (CMC ≈ 1.1 mM[14]) or by the chaotropic reagents urea or guanidinium chloride.[12]

In the course of titrations with these detergent and reagents, the W-unit exhibited both localized and progressive denaturation. Specifically, the fifth helix denatured prior to the remainder of the globular core alongside clear perturbation to the region of the core underneath this helix in the folded state. We therefore hypothesized that this localized

CIHR Author Manuscript CIHR Author Manuscript CIHR Author Manuscript

CIHR Author Manuscript CIHR Author Manuscript CIHR Author Manuscript

denaturation may provide an intermediate state *en route* to fiber formation, with the remainder of the globular helices retaining their secondary structure, given that the fibrous state exhibits loss of only a portion of the total helical content.[12] However, mapping of this structural transition employing chemical shift perturbation in heteronuclear ^{1}H, ^{13}C and/or ^{15}N correlation NMR experiments proved to be of limited utility due to spectral overlap exacerbated by loss of chemical shift dispersion during W_1 denaturation.

The ^{19}F nucleus is an outstanding NMR probe of local environment in proteins.[15–18] ^{19}F chemical shifts are extremely sensitive to van der Waals and electrostatic interactions, exhibiting ~100 times more sensitivity than the corresponding ^{1}H nuclei upon replacement.[19] The ^{19}F chemical shift,[18] longitudinal relaxation time constant (T_1),[18] and solvent isotope shift (SIS) difference between H_2O and D_2O environments[20] are good indicators of the extent of burial of the nucleus within a protein. In general, more buried residues are associated with more downfield chemical shifts, more rapid T_1 relaxation, and smaller values of SIS.[18,20] To date, a single study has employed ^{19}F-NMR to investigate silk, using 4-fluorophenylalanine-labeled *Bombyx mori* silkworm fibroin model peptides as a probe for side chain mobility.[21] This technique has not yet, to our knowledge, been applied to spider silk.

The exquisite modularity of AcSp1 makes W_1 an ideal model system to better understand the structural transition, orientation and self-assembly of W units that occur during fiber formation. Here, we have investigated the sensitivity and utility of probes of local conformational environment within W_1 using ^{19}F-NMR spectroscopy. Conveniently, the sequence of W_1 is devoid of tryptophan,[8,13] allowing for straightforward background-free site-directed mutagenesis. ^{19}F-labeling of tryptophan, in particular the incorporation of 5-fluorotryptophan (5F-Trp), is inexpensive and efficient.[22] This has allowed for straightforward position-specific incorporation of ^{19}F probes.

Four residues have been chosen for site-directed tryptophan mutagenesis of W_1 herein. Three were selected as probes of the globular domain: one in the most readily denatured helix 5 region, F146, potentially involved in conformational change to β-strand during fibrillogenesis; a mildly exposed phenylalanine in helix 3 on the surface of the globular domain, F90, possibly enhancing protein-protein interactions during fibrillogenesis; and, a buried, albeit charged residue, R36, within the center of the folded protein in the region most perturbed upon helix 5 unfolding. The fourth residue, Y169, is fully solvent exposed and located in the intrinsically disordered, C-terminal tail providing a probe that acts herein effectively as an internal control of an unperturbed 5F-Trp site. Using 5-fluoroindole,[22] we biosynthetically ^{19}F-labeled the tryptophan residues at each of these four positions (R36W, F90W, F146W and Y169W) in W_1 mutants. Through NMR and circular dichroism (CD) spectroscopy in combination with simulated-annealing structural calculations, we show that tryptophan mutation and ^{19}F-labeling do not affect the global fold of the protein. Furthermore, W_2 proteins containing 5F-Trp in both W-units at each of the four substitutions are capable of silk-like fiber formation in the same manner as wild type W_2. Each fluorinated mutant provides distinct ^{19}F chemical shift, T_1 relaxation time, and SIS values reflective of burial or solvent exposure and is sensitive to changes in local conformational environment and dynamics upon addition of urea or DPC. To our knowledge, this work

CIHR Author Manuscript

CIHR Author Manuscript

CIHR Author Manuscript

represents the first employment of ^{19}F labels in a spider silk alongside the first report of mutagenesis of an aciniform silk protein.

EXPERIMENTAL PROCEDURES

Simulated W_1 Trp mutant structural ensembles and analysis

A single round of structure calculation was performed using Xplor-NIH 2.38[23,24] for each W_1 Trp mutant protein in identical manner to the final round of structural calculation for W_1 (PDB entry 2MU3[12]) prior to water refinement with the following two modifications: (1) the Xplor-NIH sequence input file was modified appropriately to convert R36, F90, F146, or Y169 to a Trp; and, (2) nuclear Overhauser effect (NOE) distance constraints to the native side chain were removed, leading to the removal of the following number of restraints: 49 (R36W), 36 (F90W), 14 (F146W), or 8 (Y169W). The per-residue and per-atom solvent accessible surface area (SASA) were calculated by the GETAREA web script (http://curie.utmb.edu/getarea.html) based upon the program FANTOM.[25] SASA values were averaged for the 20 lowest energy members of each 100 member mutant W_1 ensemble and weighted as a function of total Trp surface area[26] as for W_1.[12]

W_1 and W_2 mutagenesis, expression, 5F-Trp-labeling, and purification

W_1 Trp mutant genes were produced using previously described protocols for W_1[13] with slight modifications. In brief, individual tryptophan mutations were introduced on a previously descried pDW_1 plasmid (pDrive vector containing the W_1 gene) through inverse PCR[27] with mutagenic primers (Table S1). W_2 Trp mutant plasmids (pDW_2) were produced using the same cloning strategy as previously described[13], by ligation (using T4 DNA ligase) of two DNA fragments prepared from each of the above pDW_1 Trp mutant plasmids through restriction endonuclease digestions (*Bse*RI and *Bam*HI; *Bsg*I and *Bam*HI). pDW_1 (or pDW_2) Trp mutant plasmids were then digested by *Bsa*I and *Bfu*AI (New England Biolabs, Ipswich, MA), and the resulting W_1 (or W_2) Trp mutant genes were inserted downstream of the His_6-SUMO coding sequence in a modified pET-32 expression plasmid (EMD Millipore, Billerica, MA, USA) that was digested by the same enzymes.

Each expression plasmid encoding a His_6-SUMO-W_1 (or W_2) Trp mutant was introduced into *E. coli* BL21 (DE3) (Novagen, Darmstadt, Germany) using standard transformation protocols. The resulting *E. coli* cells containing the correct plasmid were grown in liquid LB medium with ampicillin (50 μg/mL) to OD_{600} ~ 0.6. For ^{19}F labeling, cells were then collected by centrifugation and re-suspended in M9 minimal growth medium[28] containing 1 g/L $(NH_4)_2SO_4$ (for ^{15}N labeling, $^{15}(NH_4)_2SO_4$ (Cambridge Isotope Laboratories, Tewksbury, MA)), 2 g/L D-glucose, 4 mL/L of 1M $MgSO_4$, and 1.8 ml/L of 1 mM $FeSO_4$. Following incubation in minimal medium at 37 °C for ~30 min, 60 mg/L of 5-fluoroindole (Sigma-Aldrich, Oakville, ON) in 1 mL dimethyl sulfoxide was added for biosynthetic ^{19}F-labeling of tryptophan.[22] Expression of a given ^{19}F-labeled W_1 (or W_2) Trp mutant protein was then induced through addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) to a concentration of 0.8 mM. After overnight induction at ~22 °C, protein purification and SUMO cleavage were performed as described previously for wild type W_1.[29]

CIHR Author Manuscript CIHR Author Manuscript CIHR Author Manuscript

Circular dichroism (CD) spectroscopy

Far-ultraviolet (UV) CD spectra were acquired on a J-810 spectropolarimeter (Jasco, Easton, MD) with a scan speed of 20 nm/min, response time of 1 s, acquisition interval of 1 nm, and a bandwidth of 2 nm. Spectra of 10–30 μM 5F-Trp-labeled W_1 Trp mutants in 20 mM sodium acetate buffer, pH 5.0 were collected at 22 ± 2 °C from 260 to 195 nm in a quartz cuvette (0.05 cm path length; Hellma Analytics, Plainview, NY). A blank solution containing the same buffer was measured under the same experimental conditions. For each sample, the blank solution was subtracted from the average of two scans, and the blank subtracted spectrum was converted to mean residue ellipticity on the basis of concentration determined by ultraviolet spectroscopy at 210 nm (estimated as ε_{210} = 270,858 M^{-1} cm^{-1} for W_1 and 541,716 M^{-1} cm^{-1} for W_2) in the same 0.05 cm path length quartz cuvette.

NMR characterization of W_1 mutants

Samples of individual W_1 Trp and 5F-Trp-labeled mutant proteins (0.06–0.2 mM), 2,2,2-trifluoroethanol-d_2 (TFE, 0.1 mM), and 5-fluoroindole (0.1 mM) were prepared in the buffer (20 mM CD_3COONa, 1 mM 2,2-dimethyl-2-silapentane-5-sulfonic acid (DSS), 1 mM NaN_3; pH 5.0; H_2O:D_2O at 9:1 (v:v), "NMR buffer") previously employed for W_1[12,29] and W_2[12] chemical shift assignment, structural and hydrodynamic characterization. For ^{19}F SIS effect measurement, the buffer for a given protein or 5-fluoroindole sample was exchanged to NMR buffer containing H_2O:D_2O at 1:9 (v:v). TFE was included as a reference for ^{19}F experimental setup standard in all samples, but not as a chemical shift standard (referencing detailed below).

NMR experiments were performed on an 11.7 T Avance spectrometer (Bruker Canada, Milton, ON) using a double-resonance 5-mm BBFO SmartProbe at 30 °C. 1D ^1H experiments (16–32 scans, sweep width 16 ppm, acquisition time 2 s, recycle delay 1–2 s) were performed with water suppression using presaturation or excitation sculpting.[30] 2D ^1H-^{15}N heteronuclear single quantum coherence (HSQC) experiments (16 scans, sweep width 26/14 ppm in F1/F2, size of fid 128/1024 in F1/F2, recycle delay 1 s) were performed using sensitivity improvement[31] and decoupling during acquisition. 1D ^{19}F experiments (128–5120 scans, sweep width 90 ppm centered at −105.0 ppm, acquisition time 387 ms, recycle delay 1–2 s) were performed with or without ^1H decoupling during acquisition. Inversion-recovery ^{19}F T_1 experiments (320–1136 scans, sweep width 50 ppm centered at −124.5 ppm, acquisition time 695 ms, recycle delay 5 s) were acquired with a series of 10 recovery delays (0.01, 0.03, 0.06, 0.10, 0.15, 0.25, 0.50, 1.00, 1.50 and 2.00 s) with or without ^1H decoupling during acquisition.

NMR data were processed and analyzed using Bruker TopSpin 3.1 and, for T_1 analysis, Bruker Dynamics Center. ^1H chemical shifts were referenced directly to DSS at 0.0000 ppm. ^{19}F and ^{15}N chemical shifts were indirectly referenced to the DSS frequency at 0.0000 ppm in the ^1H 1D spectrum, following IUPAC[32] and IUBMB[33] standards, respectively. ^{19}F chemical shifts and linewidths were determined using the Lorentz/Gauss Deconvolution in the Line Shape Fitting module of Topspin 3.1. ^{19}F SIS values were obtained as the difference between the chemical shift in 1:9 H_2O:D_2O and that in 9:1 H_2O:D_2O (i.e., $\delta(D_2O) - \delta(H_2O)$). T_1 relaxation time constants were determined from exponential fits of

CIHR Author Manuscript CIHR Author Manuscript CIHR Author Manuscript

the observed signal intensities as a function of delay for the inversion-recovery experiment, using the relationship:

$$I(t)=I(0)(1-2Ae^{-\frac{t}{T_1}}) \quad (1)$$

where $I(t)$ is the observed signal intensity at delay time t, $I(0)$ is the intensity of the full signal, and A is a normalization factor. Perturbation to structure by insertion of 5F-Trp were assessed through analysis of perturbation of ^{1}H-^{15}N HSQC peak positions through a gyromagnetic ratio-weighted combined chemical shift (CCS)[34] comparison to wild type W_1, as detailed previously.[12]

Fiber formation assays

Hand-drawn fibers were pulled from solutions of individual W_2 5F-Trp-labeled mutants (R36W, F90W, F146W or Y169W on both W units) either in 50 mM potassium phosphate buffer (pH 7.5) or in 20 mM sodium acetate buffer (pH 5.0), using a pipette tip as described previously[13] for wild type $W_{2–4}$. In brief, 10-μL aliquots of protein solution (at ~19 μM) were deposited on clean glass slides and ~1 cm long fibers were drawn and put directly back on the slide. Fibers were imaged using an inverted optical microscope (Axiovert 200M, Carl Zeiss Canada, Toronto, ON; 40× objective and 10× ocular).

W_1 mutant titrations

Titrations were carried out using mixtures of either all four or of two of the four W_1 5F-Trp-labeled mutants (individual protein concentrations of 0.06 mM–0.23 mM) prepared in the acetate buffer detailed above. Independent DPC (0, 0.5, 0.7, 1.0 1.3, 2.5, 3.7, 5.0, 10 and 20 mM; dilution negligible) and urea (0, 0.5, 1.0, 1.2, 1.4, 2.0, 2.3, 2.5, 2.7, 2.9, 3.1, 3.3, 3.5, 3.7 and 4.0 M; total W_1 concentration ~0.19–0.32 mM at 4.0 M urea endpoint) titrations were carried out. 1D ^{1}H and ^{19}F spectra were acquired at each titration point and ^{19}F inversion-recovery experiments were carried out at 0, 1.3 and 20.0 mM DPC and at 0 and 4.0 M urea, with T_1 values determined as for individual proteins using Bruker Dynamics Center.

^{19}F T_1 and T_2 predictions

Two interactions are generally held to dominate ^{19}F spin relaxation: heteronuclear dipolar coupling of the dilute ^{19}F nucleus with surrounding protons (dip)[18,35] and chemical shift anisotropy (CSA).[36,37] Each contribution may be considered independently, with a given relaxation rate provided through the sum of the components:

$$R_1=\frac{1}{T_1}=(R_1)_{dip}+(R_1)_{CSA}=\left(\frac{1}{T_1}\right)_{dip}+\left(\frac{1}{T_1}\right)_{CSA} \quad (2)$$

CIHR Author Manuscript CIHR Author Manuscript CIHR Author Manuscript

$$R_2=\frac{1}{T_2}=(R_2)_{dip}+(R_2)_{CSA}=\left(\frac{1}{T_2}\right)_{dip}+\left(\frac{1}{T_2}\right)_{CSA} \quad (3)$$

where R_1 and R_2 are the longitudinal and transverse relaxation rates, respectively; T_1 and T_2 the longitudinal and transverse relaxation time constants, respectively; and, the bracketed subscripts indicate the contribution from a given relaxation mechanism. The formalism employed to estimate the dipolar and CSA contributions to eqs. 1–2 is detailed in the Appendix.

RESULTS

Putative structural ensembles of W_1 Trp mutants

Structural ensembles were calculated for each W_1 Trp mutant (positions in wild type W_1 in Figure 1A) in the same manner as for wild type W_1[12] employing all NMR restraints except distance constraints involving the mutated side chain. Despite the increased steric bulk of the Trp side chain, the total energies observed for the resulting 20-member structural ensembles were similar in magnitude to W_1. Both backbone and heavy atom root-mean-square deviations (RMSDs) of the ensembles of mutant structures, relative to wild type W_1, are comparable in magnitude to the RMSD for the wild type W_1 ensemble itself (Table S2). Furthermore, none of the mutant structural ensembles exhibited substantially increased NOE constraint violations relative to the converged W_1 ensemble (Table S2). As a whole, these structure calculations imply that W_1 should be able to accommodate mutation to Trp at each of the four selected locations without significant structural perturbation. Importantly, the ^{19}F probe in each mutant (blue spheres, Figure 1B–E) should experience a different local environment.

Relative ensemble-averaged SASA values were predicted for each Trp residue, with SASA ranked as R36W < F90W < F146W < Y169W (Table 1). Notably, the SASA values were only subtly perturbed relative to those of the corresponding residues in wild type W_1, with three mutants increasing in relative SASA (2% for R36W, 11% for F90W, and 12% for Y169W) and one decreasing (−4.5% for F146W). Following the relative degrees of surface exposure and packing, increased relative SASA corresponds directly to increased variability of Trp side chain positioning over the putative structural ensembles (Figures 1B–E). Localized solvent accessibility at the C-4 (Cε2), C-5 (Cζ3), and C-6 (Cη2) positions was used to infer potential for solvent exposure in proximity to the 5F position (Table 1); even in the most protected R36W position, ~25% of the ensemble members exhibited a SASA of ~6–9 $Å^2$ at both the C-5 and C-6 positions.

Properties of 5F-Trp-labeled W_1 tryptophan mutants

The expression levels of each His_6-SUMO-W_1 Trp mutant fusion protein (hereafter referred to by mutation as R36W, F90W, F146W or Y169W) in the presence of 5-fluoroindole were similar to that of His_6-SUMO-W_1 containing wild type W_1 in M9 minimal medium (~80 mg/L of cell culture). All fusion proteins were soluble after cell lysis and efficiently purified to >90% purity (Figure S1). After cleaving His_6-SUMO from the fusion proteins using

CIHR Author Manuscript CIHR Author Manuscript CIHR Author Manuscript

SUMO protease, the resulting 5F-Trp-labeled W_1 Trp mutants (~0.1 mM in solution) were stable at 4 °C for at least two weeks.

Far-UV CD spectra of the four W_1 5F-Trp-labeled mutants and wild type W_1 were practically identical (Figure 2). All mutants display the expected[12] characteristic negative α-helical bands at ~208 and ~222 nm convoluted by features from disordered and β-turn structures. Consistent with the CD spectroscopy results, 1D ^{1}H NMR spectra were highly similar for each mutant relative to wild type W_1 (Figure S2).

2D ^{1}H-^{15}N HSQC spectra provided greater insight into the extent and localization of structural perturbation. The same general peak pattern (or backbone “fingerprint”) was observed for each mutant relative to wild type W_1 (Figure 3), with varying degrees of perturbation apparent. Minimal spectral perturbation in the globular core was observed for R36W, F146W, and Y169W (Figure 3), with each mutant exhibiting a distinct subset of ^{1}H-^{15}N HSQC cross-peaks that was perturbed relative to W_1. CCS perturbations relative to the assigned W_1 chemical shifts[29] demonstrated that the cross-peaks that exhibit the highest degree of perturbation correspond to residues close in sequence and/or space (Figures 3, S3, S5 and S6). The F90W mutant, conversely, exhibited a higher degree of chemical shift perturbation, not necessarily localized to the site of mutation (Figures 3 and S4).

Fiber formation by 5F-Trp-labeled W_2 Trp mutants

The expression level of each double Trp mutant His_6-SUMO-W_2 fusion protein labeled with 5F-Trp was similar to that of the wild type His_6-SUMO-W_2 fusion protein in M9 minimal medium (~20–40 mg/L of cell culture), but different degrees of solubility were evident after cell lysis (wild type and F90W: ~80% soluble; R36W, F146W and Y169W: ~20% soluble). Only the soluble fraction of the cell lysate was subjected to immobilized metal affinity purification, allowing selective binding of the hexahistidine (His_6) tag of the fusion protein. Cleavage of His_6-SUMO from each fusion protein was carried out using SUMO protease, and the purity of each 5F-Trp-labeled W_2 mutant following reverse purification to remove the His_6-SUMO was >90% (Figure S7). All 5F-Trp-labeled W_2 mutants (~0.02 mM) were stable in both acetate and phosphate buffers at 4 °C for at least two weeks.

The ability to manually draw fibers from each W_2 double-mutant, with 5F-Trp at the same position in each W unit, was tested in both phosphate[13] and acetate[12] buffer. Each mutant exhibited similar fiber-forming capability to wild type W_2 (Figure 4). Therefore, despite varying degrees of backbone chemical shift perturbation (Figures 3 and S3–S6), implying varying degrees of structural perturbation, each of the tryptophan mutants retains full ability to form fibers.

^{19}F-NMR properties of W_1 mutants

Each 5F-Trp-labeled W_1 mutant exhibited a distinct ^{19}F chemical shift, ranging from −123.64 ppm to −126.27 ppm (Figure 5; Table 2). The residue predicted to be most buried, R36W, exhibited the most downfield chemical shift, while the mildly exposed F90W exhibited the most upfield chemical shift. In comparison, 5-fluoroindole under identical conditions exhibited a chemical shift between those of F90W and the next most upfield W_1 mutant, Y169W, the 5F-Trp residue predicted to be most solvent exposed (Figure 1 and

CIHR Author Manuscript CIHR Author Manuscript CIHR Author Manuscript

Table 1). Upon exchange from 10% to 90% D_2O, each W_1 mutant exhibited an SIS (Figure 5, Table 2; 5-fluoroindole and TFE shown for comparison), with the largest magnitude of SIS being observed for R36W. An upfield SIS was observed for every ^{19}F species except F90W, which exhibited a downfield shift.

Each W_1 mutant also exhibited a spin relaxation behavior (Table 2). The T_1 arising from CSA (eq. 11, Appendix) was calculated to be 20.6 s. Consistent with the study of protein J by Post et al.,[38] this value is ~25–50 times larger than our observed T_1 values, implying that T_1 is largely dominated here by dipole-dipole interactions (eq. 9, Appendix). Under isotropic tumbling conditions without a contribution from significant internal motion, the T_1 values (eq. 2) for buried vs. solvent exposed 5-fluoro moieties were calculated as 0.529 s and 0.507 s, respectively. For comparison, CSA is predicted to dominate for T_2 relaxation time constant (eq. 12, Appendix), adding ~11 Hz in linewidth in addition to ~4 Hz contribution from dipole-dipole interactions (eq. 10, Appendix). Quantitative determination of T_2 values was not carried out both due to the strong CSA dependence of T_2 (values at multiple field strengths would optimally be used to quantify dipolar vs. CSA contributions) and due to instrumental limitations in employing Carr-Purcell-Meiboom-Gill (CPMG) pulse trains for accurate T_2 measurement.

Following titration-induced W_1 perturbation

Distinct chemical shift and linewidth behavior of each 5F-Trp means that mixtures of W_1 mutants are amenable for detailed characterization of titration-induced perturbation. ^{19}F-NMR behavior of the mutant W_1 species in mixtures was indistinguishable from that of each individual species, indicating that mutant W_1 species did not interact with or affect each other (Figure S8).

Urea-induced W_1 denaturation gave rise to distinct behavior at each 5F-Trp position. R36W and F90W exhibited the largest chemical shift changes and, as with the SIS effect (Figure 5), the changes were opposite in sign (Figure 6B). Both of these ^{19}F probes behaved very similarly, with a urea concentration-dependent perturbation in the opposite direction prior to broadening beyond detection before a final, dramatic change in chemical shift (Figure 6C). F146W, conversely, transitioned in a unidirectional manner towards its final state. Each ^{19}F probe converged on the essentially unchanged F169W chemical shift, with a jump to this chemical shift at a distinct urea level (Figure 6C): first R36W, then F146W, and, finally, F90W. Following the jump, each probe exhibited a consistent, narrowed linewidth (Figure S9A for 4 M urea endpoint) while T_1 values became indistinguishable from Y169W at all sites at 4 M urea (Figure S9B).

DPC-induced ^{19}F transitions were generally unidirectional and, with the exception of R36W, gradual (Figure 6E–F). As with urea, each probe exhibited differences in DPC concentration-dependent perturbation, with R36W affected first, followed by F146W, and, finally, F90W. Although moderately perturbed, F146W became highly similar to the unchanged Y169W chemical shift, while R36W and F90W converged upon a distinct chemical shift that was upfield by ~0.6 ppm (Figure 6E). The R36W signal broadened beyond detection after the first DPC titration point, reappearing and remaining unchanged at its final, upfield shift over 2.5–20 mM DPC (Figure S8). At 1.3 mM DPC (just above the

CIHR Author Manuscript

CIHR Author Manuscript

CIHR Author Manuscript

CMC), F146W becomes significantly broadened but its T_1 is unchanged; Y169W, conversely, is only slightly broadened and T_1 is effectively unchanged (Figure S9). At 20 mM DPC, all three of these ^{19}F probes exhibited decreased linewidth and increased T_1 (Figure S9), thus becoming similar to Y169W.

DISCUSSION

Mutation tolerance of W unit

The tolerance to mutation within the W unit, exhibited through minimal structural perturbation in W_1 and ability to form fibers by the corresponding W_2 concatemers, is remarkable. The gene and, correspondingly, primary structuring of W units within AcSp1 are exceptionally homogeneous, with ~99.9% similarity at the DNA level for the *A. trifasciata* sequence.[8] Even more strikingly, Chaw et al. demonstrated that the ~150 amino acid structured globular domain, annotated on the basis of our chemical shift assignments[12,29], exhibits ~84% pairwise identity over three *Argiope* and one *Araneous* species.[11] Conversely, as might be anticipated from structural constraints, the intrinsically disordered linker region exhibits only ~54% pairwise identity alongside indels leading to variable length.[11] Notably, R36, F90 and Y146 all exhibit 100% conservation over these species. It is somewhat surprising, therefore, that the W unit is tolerant to mutation at all of these sites, particularly at R36 given the extreme modification of chemical character in an Arg to Trp substitution. It is certainly possible that one of more of these substitutions would affect the efficiency of fiber spinning or of protein storage at a high concentration in the spider aciniform gland, parameters not detectable in our hand-pulling assay. As a whole, however, the mutation tolerance of the W unit bodes well for rational engineering of modifications in this class of protein.

Y169W behavior

Residue Y169 is highly exposed, with Y169W correspondingly predicted to have 93% SASA. ^{19}F chemical shift and SIS values are fully consistent with this positioning (Table 2). Slow exchange on the NMR time scale is also apparent, with multiple distinct ^{19}F chemical shifts that vary in intensity as a function of conditions (Figure S8). The Y169W spin relaxation behavior is consistent with solvent exposure, exhibiting a narrow linewidth throughout both urea and DPC titrations (perhaps anomalously narrow at 4 M urea in comparison at 5.4 Hz) and a T_1 that remains unchanged at 4 M urea (Figure S9). Notably, the T_1 of Y169W is longer than anticipated (eq. 2). This would be consistent with relaxation arising from a spectral density with components originating both from molecular tumbling at the global correlation time of 7.9 ns and from internal motion at some other local correlation time (e.g., as described in the Lipari-Szabo model-free formalism[39]). Assuming a generalized order parameter (S^2) of ~0.5–0.6 consistent with a surface exposed Trp side chain and with the heteronuclear NOE (hetNOE) enhancement factor of −0.23 for the ^{1}H-^{15}N backbone amide at Y169,[12] a sub-ns effective correlation time would be required to recapitulate this in the model-free approach. However, slow ring flipping is likely giving rise to more complicated motional relationships than can be described by internal motion using the Lipari-Szabo model-free approach, such as the requirement for an exchange contribution.[40] In short, Y169W can be considered an internal ^{19}F-NMR control for the

CIHR Author Manuscript

CIHR Author Manuscript

CIHR Author Manuscript

other mutants, in that its behavior is generally unperturbed and consistent throughout with a solvent exposed 5F-Trp. This also provides a potentially highly valuable NMR probe for future studies tracking fiber formation and internal conformation by ^{19}F-NMR, since perturbation would be expected upon fibrillogenesis.

R36W behavior

Our previous finding that R36 is positioned in the W_1 protein core was unexpected, but very well supported by the NOE restraint network.[12] The structural and functional tolerance to Trp mutation implies that basic side chain functionality is not essential, at least in an in vitro fibrillogenesis setting. Consistent with a buried positioning (4% SASA), R36W exhibits the most downfield chemical shift of the four probes. Its linewidth is also ~double that of Y169W, consistent with a buried positioning. However, its T_1 is longer than predicted by eq. 2 and comparable to that of Y169W. Presuming an S^2 of ~0.8–0.9, consistent with a Trp side chain restricted through its location within in the protein core and consistent with the ^{1}H-^{15}N hetNOE enhancement of 0.62 observed at R36 in W_1,[12] local motional fluctuations in the sub-ns range are implied under model-free analysis to explain the lengthened T_1. Notably, fluctuations of this nature would also lead to a line narrowing relative to eq. 3 - this corresponds nicely to the fact that the less buried F90W probe has a broader linewidth than R36W.

The observed SIS at R36W is anomalous - it exhibits the largest SIS of all four probes, an upfield shift nearly as large as that of free 5-fluoroindole. It is notable that the putative mutant protein structure has two cavities leading from surface to the apparently buried R36W (Figure S10). Correspondingly, a subset of ensemble members exhibits solvent exposure in proximity to the 5F position (Table 1). This region of W_1 exhibits atypical ^{1}H, ^{15}N and ^{13}C chemical shifts[29] alongside evidence of intermediate exchange dynamics on the NMR time scale.[12] This is also consistent with the high frequency motion implied by the ^{19}F spin relaxation behavior of R36W. In the putative structure, the W36 side chain is positioned in close proximity (~4.5–7 Å in many ensemble members) to the side chain of R50 (Figure S10). Although not in an ideal “T-shaped” orientation, a cation-π interaction may still occur between the 6-membered ring of W36 and the guanidino moiety of R50.[41] None of the other three 5F-Trp probes are in close proximity to charged moieties. A combination of solvent accessibility through cavities, proximity to a charged group, and/or structural dynamics poorly reflected in the NOE- and chemical shift-derived NMR ensemble of W_1 may therefore all contribute to a large SIS value for this buried residue.

Consistent with urea and DPC titrations followed by heteronuclear NMR at the backbone (demonstrated for R36 in Figures 6A, D and for neighboring residues in Figure S11), R36W was the first position to undergo substantial titrant-induced perturbation. In ^{1}H-^{15}N HSQC-based titrations, individual backbone amide cross-peaks could not be tracked beyond loss of ~15% of α-helical character (2.5 M urea).[12] While this was sufficient to show greater than average perturbation at R36, it was not possible to follow it through to its final fate upon denaturation. Here, we observed a surprising sensitivity of R36W to urea relative to the backbone data in that the ^{19}F probe underwent severe broadening and was actually restored

CIHR Author Manuscript

CIHR Author Manuscript

CIHR Author Manuscript

as a narrow line, at its final chemical shift, before 2.5 M urea. This is illustrative of distinct differences in behavior at the side chain vs. main chain in this region.

F90W behavior

An unusual upfield chemical shift is observed for F90W, falling above even 5-fluoroindole. Notably, its SIS is also opposite in direction to all of the other 5F-Trp probes. Despite these issues, F90W exhibits the broadest line of the four probes, consistent with a relatively buried positioning (30% SASA) and a T_1 shorter than predicted by eq. 2. In wild type W_1, F90 exhibited well below average perturbation under titration by both urea and DPC (Figures 6A, D and S11). Correspondingly, F90W was the last residue to change in both ^{19}F-NMR titrations. Under urea titration, the jump in chemical shift for F90W occurs in the same concentration regime where the final α-helical character is lost according to CD spectroscopy,[12] indicating that this portion of the globular domain is relatively unperturbed until late in the denaturation process. Overall, the titration behavior of F90W is consistent with a buried side chain.

Charged groups in close proximity, such as neighboring carboxyls, have been noted in the literature to cause significant upfield chemical shifts.[18] Although electric field-induced ^{19}F chemical shifts for the 5F-Trp and 6F-Trp species may be relatively minimal on a protein surface,[42] buried residues experiencing a fixed dipole within the protein interior or residues interacting with solvent may exhibit significant and apparently anomalous chemical shift perturbation.[42,43] The relatively high SASA at the C-6 position on the F90W indole ring (Table 1) would support the possibility of solvent-induced changes in shielding. Short T_1 values may also be indicative of a strongly bound water molecule in close proximity to the ^{19}F, enhancing dipolar relaxation over and above the effect of the nearest (~2.6 Å) side chain ring proton (i.e., reducing the value of r_{min} in eq. 9, Appendix). If solvent exposure were the only consideration, however, one would expect an upfield shift when H_2O was replaced with D_2O rather than the observed downfield shift. Notably, F90W also exhibits the only downfield shift upon denaturation of all four 5F-Trp probes upon titration with either urea or DPC. The most probable explanation for a downfield shift upon solvent exchange or titration is that the 5F moiety is experiencing an electric field-induced chemical shift perturbation from the protein itself with only minor influence from solvent.

F146W behavior

The F146W side chain is predicted to be mostly solvent exposed (71% SASA), with a corresponding upfield chemical shift and narrow linewidth. Its SIS also follows the expected trend of an upfield shift, with a lesser magnitude than that of Y169W. Unlike Y169W, however, slow exchange is not apparent and the observed T_1 is shorter than predicted by eq. 2 for a solvent exposed 5F-Trp. This enhanced longitudinal relaxation is consistent with a relatively solvent exposed positioning in a setting which minimizes internal motion, as would be expected in its context in helix 5 of the W_1 globular domain.

Under urea titration, F146 in wild type W_1 is perturbed at a level slightly above average (Figure 6A), while other residues in helix 5 are significantly more perturbed (Figure S11). Following individual residues in helix 5 through the DPC is challenging due to its Ser-rich

CIHR Author Manuscript

CIHR Author Manuscript

CIHR Author Manuscript

sequence (Figure S11), with much of the helix being unassignable at 20 mM DPC even through triple-resonance experiments. Despite this behavior, F146 itself is relatively unperturbed, in keeping with a general trend of the perturbations along helix 5 being maximal for residues in close proximity to the W_1 helical bundle and minimal for those, like F146, that are solvent exposed. Under ^{19}F-NMR titration, F146W transitions to its final state at urea and DPC concentrations intermediate to R36W and F90W. Its dynamics are much more notably affected, with a significant line broadening at 1.3 mM DPC and restoration of linewidth at 20 mM DPC. Hence, although the chemical environment at F146W remains relatively consistent throughout titration, its dynamics are quite sensitive to the denaturation of helix 5 observed in heteronuclear experiments.

Urea denaturation vs. DPC perturbation

Comparing and contrasting the two titrations, urea denaturation appears to dramatically affect the folded region of the 5F-Trp-labeled W_1 mutant proteins by effecting global denaturation, with ^{19}F chemical shifts demonstrating differences in local solvent exposure/burial over the course of the structural transition. In contrast, DPC-induced perturbation of W_1 is gradual, with the exception of R36W. Given that urea is a chaotropic agent while DPC is an amphiphilic lysophospholipid analog, a difference in their denaturing behavior is to be expected.

We presume that the urea-denatured W_1 is completely unfolded, as reflected by highly similar ^{19}F chemical shifts and T_1 values alongside narrowed lines for all four probes beside a lack of ellipticity arising from α-helical (or β-strand) secondary structuring in CD spectroscopy. Unlike our previous study, use of ^{19}F-NMR allowed us to track denaturation of the protein to completion. Early denaturation of helix 5 (residues 135–149) is consistent with the observed perturbation to R36W. New insight into this process comes from the fact that the 5F-position in the R36W side chain is affected prior to the backbone and transitions to its final chemical shift and linewidth at a point when the backbone is only starting to be perturbed at an above average level. This implies that loosening of helix 5 away from the remainder of the globular bundle occurs even more readily than previously thought.

F90W is located in helix 3 with its side chain packed against helix 1. This residue remains practically unperturbed until near the titration endpoint, implying the late unfolding of this region of W_1, corresponding to the final ~20% of helicity according to CD. Although helix 5 denaturation and the resulting perturbation to the residues underneath it (residues 26–39) were apparent previously, the long-lived stability of the helix 1/helix 3 region as well as the perturbation and rearrangement implied in the vicinity of the R36 side chain could not be resolved.

The DPC-induced perturbation to W_1 appears much more subtle, in that CD spectra are hardly changed despite the loss of fiber formation ability.[12] This appears likely to be due to adsorption of DPC molecules to the W unit based upon the hydrophobicity of the residues and favorable phosphocholine head group-protein moiety interactions. This is consistent with the fact that the two ^{19}F probes from the globular core, R36W and F90W, converge to a chemical shift ~0.6 ppm upfield from that converged upon by the two solvent exposed probes proximal to and in the intrinsically disordered C-terminal tail, F146W and Y169W.

CIHR Author Manuscript

CIHR Author Manuscript

CIHR Author Manuscript

Such behavior is also consistent with dodecyl sulfate-ubiquitin complexes characterized by Shaw et al., where detergent molecules exhibited preferential binding to hydrophobic regions with a positive surface potential without the necessity for salt-bridge formation.[44] Similar preferential binding of DPC to the W_1 globular domain may take place, although the surface potential dependence may not be so clear cut given the zwitterionic nature of the phosphocholine moiety vs. the anionic sulfate moiety.

The role of helix 5 in fiber formation has already been postulated,[12] and is fully reflected by the perturbation of R36W and F146W ^{19}F chemical shifts and relaxation behavior over the course of both urea and DPC titrations. In all likelihood, DPC becomes associated with helix 5 residues as part of the mechanism by which fiber formation is prevented. F90W as a probe of helix 3 in the globular domain suggests that although this region is initially unperturbed, it becomes much more perturbed in the later stages of protein denaturation. Taken together, these data are consistent with a fibrillogenesis pathway requiring helix 5 unfolding, whereas loss of the core structuring of the globular domain cannot be accommodated.

Summary

Spider silk proteins are extremely difficult to study at a high resolution. To our knowledge, this work represents the first example of incorporation of ^{19}F probes in this class of protein. Our work suggests that strategic mutagenesis of single residues to tryptophan does not alter the global fold of the W_1 protein, nor does it inhibit fiber formation by W_2, despite high conservation at these sites in a variety of species. 5F-Trp-labeled W_1 mutants exhibited distinct ^{19}F chemical shifts and T_1 relaxation times that are reflective of the local environment of each fluorine probe, and are interesting in terms of solvent exposure, local protein-^{19}F interactions, and the effects of side chain dynamics. Correspondingly, perturbation of W_1 by addition of urea and DPC, in distinct manner, is reflected in corresponding changes of ^{19}F chemical shifts, T_1 relaxation times, and linewidths. ^{19}F probes allow for more precise tracking of transitions in both conformation and dynamics within the W unit as perturbants are added, which in turn provides additional clues as to the nature of the structural transition required for fiber formation.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding Source Statement: This work was supported by Discovery and Research Tools and Instruments Grants from the Natural Sciences and Engineering Research Council of Canada (NSERC; to JKR and XQL); a Leaders Opportunity Fund award from the Canadian Foundation for Innovation (to JKR); and, a Dalhousie Medical Research Foundation Capital Equipment Grant (to JKR and XQL). JKR is supported by a Canadian Institutes for Health Research New Investigator Award; and, KEO was supported by NSERC Undergraduate Student Research Awards and MLT by an NSERC Doctoral Postgraduate Scholarship.

We thank Dr. Mike Lumsden for 11.7 T NMR spectrometer support at the Nuclear Magnetic Resonance Research Resource Facility (NMR^3, Dalhousie University), Dr. David Waisman for CD spectropolarimeter access and Dr. Stephen Bearne for comments on the manuscript.

CIHR Author Manuscript

CIHR Author Manuscript

CIHR Author Manuscript

ABBREVIATIONS

5F-Trp	5-fluorotryptophan
AcSp1	aciniform spidroin 1
CCS	combined chemical shift
CD	circular dichroism
CMC	critical micelle concentration
CSA	chemical shift anisotropy
dip	dipolar
DPC	dodecylphosphocholine
DSS	4,4-dimethyl-4-silapentane-1-sulfonic acid
HSQC	heteronuclear single quantum coherence
IPTG	isopropyl β-D-1-thiogalactopyranoside
NMR	nuclear magnetic resonance
NOE	nuclear Overhauser effect
PCR	polymerase chain reaction
RMSD	root-mean-square deviation
SASA	solvent accessible surface area
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SIS	solvent isotope shift
SUMO	small ubiquitin-like modifier protein
T_1	longitudinal relaxation time constant
T_2	transverse relaxation time constant
TFE	2,2,2-trifluoroethanol
W	AcSp1 (“wrapping” silk) repeat unit
WT	wild type

References

1. Gosline JM, Guerette PA, Ortlepp CS, Savage KN. The mechanical design of spider silks: From fibroin sequence to mechanical function. J Exp Biol. 1999; 202:3295–3303. [PubMed: 10562512]
2. Vollrath F, Porter D. Spider silk as a model biomaterial. Appl Phys A Mater Sci. 2006; 82:205–212.

CIHR Author Manuscript CIHR Author Manuscript CIHR Author Manuscript

3. Tokareva O, Jacobsen M, Buehler M, Wong J, Kaplan DL. Structure-function-property-design interplay in biopolymers: Spider silk. Acta Biomater. 2014; 10:1612–1626. [PubMed: 23962644]
4. Pritchard EM, Valentin T, Panilaitis B, Omenetto F, Kaplan DL. Antibiotic-releasing silk biomaterials for infection prevention and treatment. Adv Funct Mater. 2013; 23:854–861. [PubMed: 23483738]
5. Wendt H, Hillmer A, Reimers K, Kuhbier JW, Schafer-Nolte F, Allmeling C, Kasper C, Vogt PM. Artificial skin--culturing of different skin cell lines for generating an artificial skin substitute on cross-weaved spider silk fibres. PLoS One. 2011; 6:e21833. [PubMed: 21814557]
6. Hennecke K, Redeker J, Kuhbier JW, Strauss S, Allmeling C, Kasper C, Reimers K, Vogt PM. Bundles of spider silk, braided into sutures, resist basic cyclic tests: Potential use for flexor tendon repair. PLoS One. 2013; 8:e61100. [PubMed: 23613793]
7. Scheibel T. Spider silks: Recombinant synthesis, assembly, spinning, and engineering of synthetic proteins. Microb Cell Fact. 2004; 3:14. [PubMed: 15546497]
8. Hayashi CY, Blackledge TA, Lewis RV. Molecular and mechanical characterization of aciniform silk: Uniformity of iterated sequence modules in a novel member of the spider silk fibroin gene family. Mol Biol Evol. 2004; 21:1950–1959. [PubMed: 15240839]
9. Fu C, Shao Z, Fritz V. Animal silks: Their structures, properties and artificial production. Chem Commun. 2009:6515–6529.
10. Agnarsson I, Dhinojwala A, Sahni V, Blackledge TA. Spider silk as a novel high performance biomimetic muscle driven by humidity. J Exp Biol. 2009; 212:1990–1994. [PubMed: 19525423]
11. Chaw RC, Zhao Y, Wei J, Ayoub NA, Allen R, Atrushi K, Hayashi CY. Intragenic homogenization and multiple copies of prey-wrapping silk genes in *Argiope* garden spiders. BMC Evol Biol. 2014; 14:31. [PubMed: 24552485]
12. Tremblay ML, Xu L, Lefevre T, Sarker M, Orrell KE, Leclerc J, Meng Q, Pezolet M, Auger M, Liu XQ, Rainey JK. Spider wrapping silk fibre architecture arising from its modular soluble protein precursor. Sci Rep. 2015; 5:11502. [PubMed: 26112753]
13. Xu L, Rainey JK, Meng Q, Liu XQ. Recombinant minimalist spider wrapping silk proteins capable of native-like fiber formation. PLoS One. 2012; 7:e50227. [PubMed: 23209681]
14. Palladino P, Rossi F, Ragone R. Effective critical micellar concentration of a zwitterionic detergent: A fluorimetric study on *n*-dodecyl phosphocholine. J Fluoresc. 2010; 20:191–196. [PubMed: 19756982]
15. Ge X, MacRaild CA, Devine SM, Debono CO, Wang G, Scammells PJ, Scanlon MJ, Anders RF, Foley M, Norton RS. Ligand-induced conformational change of *Plasmodium falciparum* AMA1 detected using ^{19}F NMR. J Med Chem. 2014; 57:6419–6427. [PubMed: 25068708]
16. Kitevski-LeBlanc JL, Evanics F, Prosser RS. Approaches for the measurement of solvent exposure in proteins by ^{19}F NMR. J Biomol NMR. 2009; 45:255–264. [PubMed: 19655092]
17. Evanics F, Bezsonova I, Marsh J, Kitevski JL, Forman-Kay JD, Prosser RS. Tryptophan solvent exposure in folded and unfolded states of an SH3 domain by ^{19}F and ^{1}H NMR. Biochemistry. 2006; 45:14120–14128. [PubMed: 17115707]
18. Hull WE, Sykes BD. Fluorotyrosine alkaline phosphatase^{19}F nuclear magnetic resonance relaxation times and molecular motion of the individual fluorotyrosines. Biochemistry. 1974; 13:3431–3437. [PubMed: 4602295]
19. Kitevski-LeBlanc JL, Prosser RS. Current applications of F-19 NMR to studies of protein structure and dynamics. Prog Nucl Magn Reson Spectrosc. 2012; 62:1–33. [PubMed: 22364614]
20. Hull WE, Sykes BD. Fluorine-19 nuclear magnetic resonance study of fluorotyrosine alkaline phosphatase: the influence of zinc on protein structure and a conformational change induced by phosphate binding. Biochemistry. 1976; 15:1535–1546. [PubMed: 4091]
21. Asakura T, Suita K, Kameda T, Afonin S, Ulrich AS. Structural role of tyrosine in *Bombyx mori* silk fibroin, studied by solid-state NMR and molecular mechanics on a model peptide prepared as silk I and II. Magn Reson Chem. 2004; 42:258–266. [PubMed: 14745806]
22. Crowley PB, Kyne C, Monteith WB. Simple and inexpensive incorporation of ^{19}F-tryptophan for protein NMR spectroscopy. Chem Commun. 2012; 48:10681–10683.
23. Schwieters CD, Kuszewski JJ, Tjandra N, Clore GM. The Xplor-NIH NMR molecular structure determination package. J Magn Reson. 2003; 160:65–73. [PubMed: 12565051]

CIHR Author Manuscript CIHR Author Manuscript CIHR Author Manuscript

24. Schwieters CD, Kuszewski JJ, Clore GM. Using Xplor-NIH for NMR molecular structure determination. Prog Nucl Magn Reson Spectrosc. 2006; 48:47–62.
25. Fraczkiewicz R, Braun W. Exact and efficient analytical calculation of the accessible surface areas and their gradients for macromolecules. J Comput Chem. 1998; 19:319–333.
26. Samanta U, Bahadur RP, Chakrabarti P. Quantifying the accessible surface area of protein residues in their local environment. Protein Eng. 2002; 15:659–667. [PubMed: 12364580]
27. Dominy CN, Andrews DW. Site-directed mutagenesis by inverse PCR. Methods Mol Biol. 2003; 235:209–223. [PubMed: 12904664]
28. Marley J, Lu M, Bracken C. A method for efficient isotopic labeling of recombinant proteins. J Biomol NMR. 2001; 20:71–75. [PubMed: 11430757]
29. Xu L, Tremblay ML, Meng Q, Liu XQ, Rainey JK. ^{1}H, ^{13}C and ^{15}N NMR assignments of the aciniform spidroin (AcSp1) repetitive domain of *Argiope trifasciata* wrapping silk. Biomol NMR Assign. 2012; 6:147–151. [PubMed: 21989955]
30. Hwang TLS, AJ. Water suppression that works. Excitation sculpting using arbitrary waveforms and pulsed field gradients. J Magn Reson A. 1995; 112:275–279.
31. Palmer AG, Cavanagh J, Wright PE, Rance M. Sensitivity improvement in proton-detected 2-dimensional heteronuclear correlation NMR spectroscopy. J Magn Reson. 1991; 93:151–170.
32. Harris RK, Becker ED, De Menezes SMC, Granger P, Hoffman RE, Zilm KW. Further conventions for NMR shielding and chemical shifts (IUPAC recommendations 2008). Pure Appl Chem. 2008; 80:59–84.
33. Wishart DS, Bigam CG, Yao J, Abildgaard F, Dyson HJ, Oldfield E, Markley JL, Sykes BD. ^{1}H, ^{13}C and ^{15}N chemical shift referencing in biomolecular NMR. J Biomol NMR. 1995; 6:135–140. [PubMed: 8589602]
34. Schumann FH, Riepl H, Maurer T, Gronwald W, Neidig KP, Kalbitzer HR. Combined chemical shift changes and amino acid specific chemical shift mapping of protein-protein interactions. J Biomol NMR. 2007; 39:275–289. [PubMed: 17955183]
35. Solomon I. Relaxation processes in a system of two spins. Phys Rev. 1955; 99:559–565.
36. Abragam, A. The Principles of Nuclear Magnetism. Clarendon Press; Oxford, UK: 1961.
37. Hull WE, Sykes BD. Fluorotyrosine alkaline phosphatase: Internal mobility of individual tyrosines and the role of chemical shift anisotropy as a ^{19}F nuclear spin relaxation mechanism in proteins. J Mol Biol. 1975; 98:121–153. [PubMed: 1195374]
38. Post JF, Cottam PF, Simplaceanu V, Ho C. Fluorine-19 nuclear magnetic resonance study of 5-fluorotryptophan-labeled histidine-binding protein J of Salmonella typhimurium. J Mol Biol. 1984; 179:729–743. [PubMed: 6389886]
39. Lipari G, Szabo A. Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. J Am Chem Soc. 1982; 104:4546–4559.
40. Clore GM, Szabo A, Bax A, Kay LE, Driscoll PC, Gronenborn AM. Deviations from the simple 2-parameter model-free approach to the interpretation of N-15 nuclear magnetic relaxation of proteins. J Am Chem Soc. 1990; 112:4989–4991.
41. Gallivan JP, Dougherty DA. Cation-pi interactions in structural biology. Proc Natl Acad Sci U S A. 1999; 96:9459–9464. [PubMed: 10449714]
42. Lian C, Le H, Montez B, Patterson J, Harrell S, Laws D, Matsumura I, Pearson J, Oldfield E. Fluorine-19 nuclear magnetic resonance spectroscopic study of fluorophenylalanine- and fluorotryptophan-labeled avian egg white lysozymes. Biochemistry. 1994; 33:5238–5245. [PubMed: 8172898]
43. Pearson JG, Oldfield E, Lee FS, Warshel A. Chemical shifts in proteins: A shielding trajectory analysis of the fluorine nuclear magnetic resonance spectrum of the *Escherichia coli* galactose binding protein using a multipole shielding polarizabilit-local reaction field-molecular dynamics approach. J Am Chem Soc. 1993; 115:6851–6862.
44. Shaw BF, Schneider GF, Arthanari H, Narovlyansky M, Moustakas D, Durazo A, Wagner G, Whitesides GM. Complexes of native ubiquitin and dodecyl sulfate illustrate the nature of hydrophobic and electrostatic interactions in the binding of proteins and surfactants. J Am Chem Soc. 2011; 133:17681–17695. [PubMed: 21939262]

CIHR Author Manuscript CIHR Author Manuscript CIHR Author Manuscript

45. Hull WE, Sykes BD. Dipolar nuclear spin relaxation of F-19 in multispin systems - Application to F-19 labeled proteins. J Chem Phys. 1975; 63:867–880.
46. Tremblay, ML. PhD Thesis. Dalhousie University; Halifax, NS: 2015.
47. Bloembergen N, Purcell EM, Pound RV. Relaxation effects in nuclear magnetic resonance absorption. Phys Rev. 1948; 73:679–712.
48. Lever JE. Purification and properties of a component of histidine transport in *Salmonella typhimurium* The histidine-binding protein J. J Biol Chem. 1972; 247:4317–4326. [PubMed: 4556308]
49. Lide, DR., editor. CRC Handbook of Chemistry and Physics. 89. CRC Press; Boca Raton, FL: 2008.
50. Zhao X, DeVries JS, McDonald R, Sykes BD. Determination of the ^{19}F NMR chemical shielding tensor and crystal structure of 5-fluoro-DL-tryptophan. J Magn Reson. 2007; 187:88–96. [PubMed: 17475524]

APPENDIX

The dipolar and CSA contributions to T_1 and T_2 (eqs. 3 and 4) were estimated as follows. To calculate the contribution of heteronuclear dipolar coupling to relaxation of the ^{19}F nucleus by the surrounding protons, we followed the formalism of Hull and Sykes,[18] as shown to be appropriate for a dilute ^{19}F spin interacting with several ^{1}H nuclei,[45] assuming isotropic motion with a rotational correlation time (τ_c) of 7.9 ns[46] the dipolar contributions to T_1 and T_2 for a spin i (^{19}F) by unlike spins j (^{1}H) are given by:

$$\left(\frac{1}{T_{1i}}\right)_{dip}=\frac{\gamma_i^2\gamma_j^2\hbar^2}{10}\left(\sum r_{ij}^{-6}\right)F_1 \quad (4)$$

$$\left(\frac{1}{T_{2i}}\right)_{dip}=\frac{\gamma_i^2\gamma_j^2\hbar^2}{20}\left(\sum r_{ij}^{-6}\right)F_2 \quad (5)$$

where γ_x is a gyromagnetic ratio ($\gamma_i = 2.5176\times10^4$ rad^{-1}·s^{-1}·G^{-1}; $\gamma_j = 2.6753\times10^4$ rad^{-1}·s^{-1}·G^{-1}); $\hbar$ is the reduced Planck constant; and, Σr_{ij}^{-6} is the sum over all internuclear distances between spin i and all interacting spins j. F_1 and F_2 are given by:

$$F_1=J(\omega_j-\omega_i)+3\,J(\omega_i)+6J(\omega_j+\omega_i) \quad (6)$$

$$F_2=F_1+4J(0)+6J(\omega_j) \quad (7)$$

where ω_i and ω_j are the angular frequencies of the given nuclei ($\omega_i = 2\pi \times 470.59$ *MHz*; $\omega_j = 2\pi \times 500.13$ *MHz*) and the spectral density function at angular frequency ω, $J(\omega)$, is given by:

CIHR Author Manuscript CIHR Author Manuscript CIHR Author Manuscript

$$J(\omega)=\frac{\tau_c}{1+\omega^2\tau_c^2} \quad (8)$$

To account for the surrounding proton bath,[38,47] the sum in eqs. 4 and 5 can be approximated as an integral over a sphere containing protons at a defined density (ρ). T_1 and T_2 for spin *i* can then be written as:

$$\left(\frac{1}{T_{1i}}\right)_{dip}=\frac{\gamma_i^2\gamma_j^2\hbar^2}{10}\rho F_1\int_{r_{min}}^{\infty}\frac{4\pi r^2}{r_{ij}{}^6}dr_{ij}=\frac{\gamma_i^2\gamma_j^2\hbar^2}{10}\rho\left(\tfrac{4}{3}\pi r_{min}^3\right)F_1 \quad (9)$$

$$\left(\frac{1}{T_{2i}}\right)_{dip}=\frac{\gamma_i^2\gamma_j^2\hbar^2}{20}\rho F_2\int_{r_{min}}^{\infty}\frac{4\pi r_{ij}{}^2}{r_{ij}{}^6}dr_{ij}=\frac{\gamma_i^2\gamma_j^2\hbar^2}{20}\rho\left(\tfrac{4}{3}\pi r_{min}^3\right)F_2 \quad (10)$$

where r_{min} is defined as the distance from the nearest proton neighbor (assumed, unless otherwise noted, to be 2.6 Å, corresponding to the nearest protons at positions 4 and 6 on the 5-fluoroindole ring). The value of ρ for a buried site was taken as 5.73×10^{22} mL^{-1} (as for the 29 kDa globular J protein[38,48]) and that for a solvent-exposed site approximated as 5.99×10^{22} mL^{-1} (based on the density of pure H_2O at 30 °C[49] with a population of 90% ^1H).

The contributions to T_1 and T_2 relaxation for spin *i* from CSA[36,37] are given by (with prefactors appropriately modified to use $J(\omega)$ as defined in eq. 8):

$$\left(\frac{1}{T_{1i}}\right)_{CSA}=\frac{6}{20}\omega_i^2\delta_{z'}^2\left(1+\frac{\eta^2}{2}\right)J(\omega_i) \quad (11)$$

$$\left(\frac{1}{T_{2i}}\right)_{CSA}=\frac{1}{20}\omega_i^2\delta_{z'}^2\left(1+\frac{\eta^2}{2}\right)[3J(\omega_i)+4J(0)] \quad (12)$$

where the components of the traceless chemical shift tensor for nucleus *i* are $\delta_{x'}$, $\delta_{y'}$, and $\delta_{z'}$ with the asymmetry parameter, η, defined as

$$\eta=\frac{\delta_{x'}-\delta_{y'}}{\delta_{z'}}. \quad (13)$$

Values of $\delta_{z'}$ = 49.3 ppm and η = −0.397, based on single crystal ^{19}F-NMR analysis of 5-fluoro-DL-tryptophan,[50] were employed.

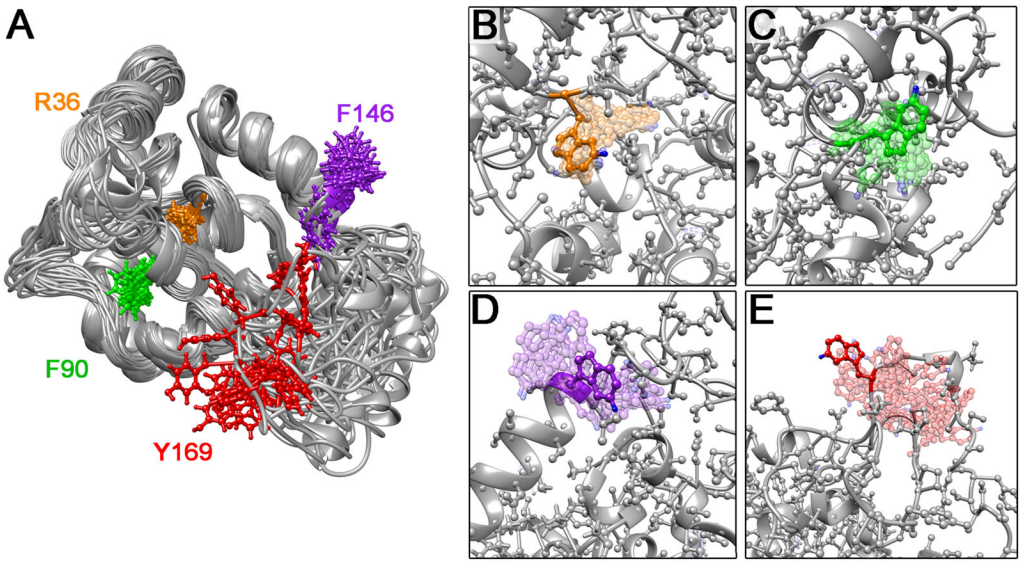


Figure 1.

(A) Superposed solution NMR structural ensemble of W_1 (PDB entry 2MU3[12]) showing side chains only for residues mutated to tryptophan. Putative structural ensembles for (B) R36W, (C) F90W, (D) F146W and (E) Y169W tryptophan mutants (20 lowest energy members of 100 member ensembles superposed onto the lowest energy structure; calculated using wild type W_1 NMR restraints with removal of distance restraints involving the mutated side chain). 5-fluoro moieties are represented by blue spheres.

CIHR Author Manuscript CIHR Author Manuscript CIHR Author Manuscript

CIHR Author Manuscript
CIHR Author Manuscript
CIHR Author Manuscript

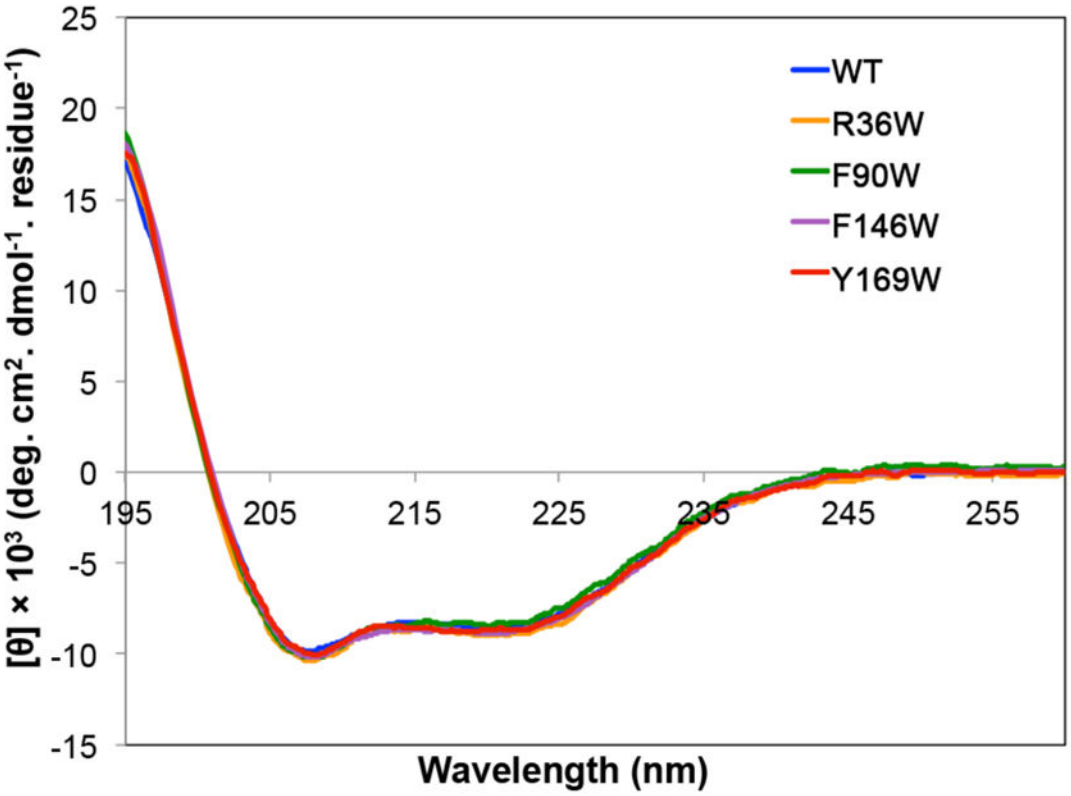


Figure 2.
Far-UV CD spectra of fluorinated W_1 tryptophan mutants (5F-Trp-labeled at R36W, F90W, F146W and Y169W) and wild type (WT) W_1.

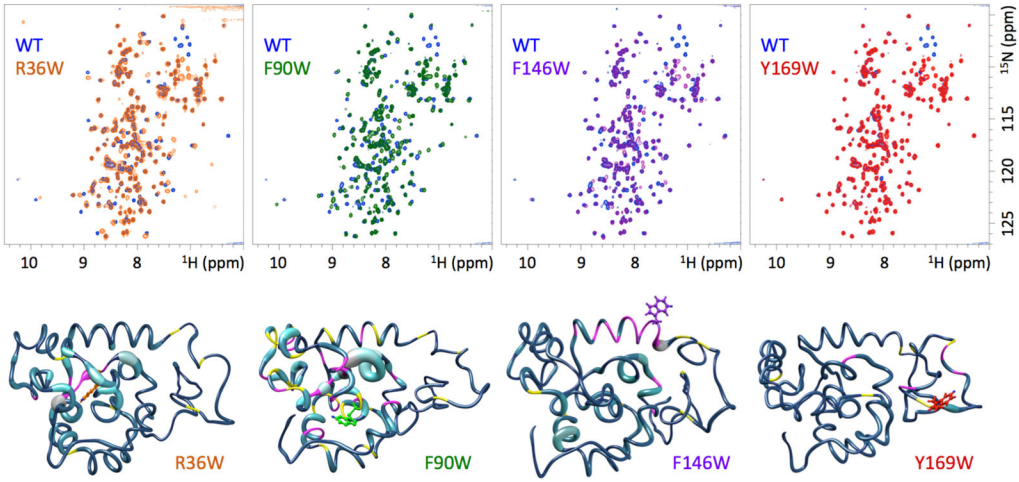


Figure 3.

2D ^{1}H-^{15}N HSQC spectra of indicated W_1 5F-Trp-labeled mutants overlaid on wild type (WT) W_1. Note that some side chain ^{1}H-^{15}N cross-peaks are aliased differently in W_1 WT vs. W_1 mutants (WT W_1 data were acquired previously using a different field strength and sweep witdth[12]). Combined chemical shift[34] perturbations as a function of position are mapped onto the lowest energy member of each calculated mutant protein structural ensemble, with the thickness of the cylinder representing the degree of ^{1}H-^{15}N HSQC peak displacement. Positions corresponding to cross-peaks that broadened beyond detection under experimental conditions are colored in magenta and ambiguous cross-peaks are colored in yellow.

CIHR Author Manuscript
CIHR Author Manuscript
CIHR Author Manuscript

CIHR Author Manuscript
CIHR Author Manuscript
CIHR Author Manuscript

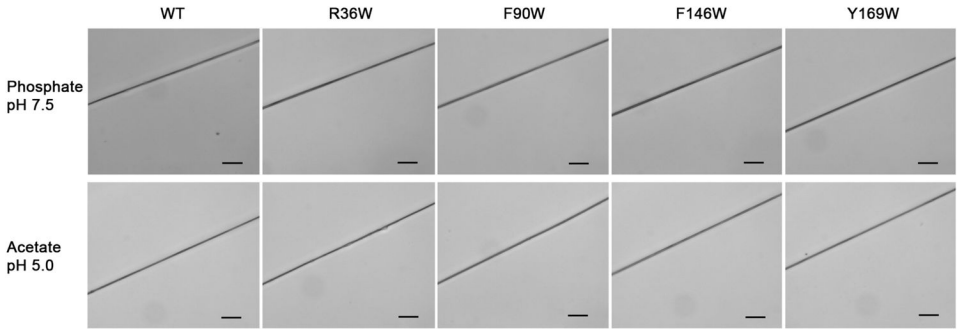


Figure 4.

Optical micrographs showing manually drawn fibers formed by wild type (WT) and mutant W_2 proteins in indicated buffer (scale bars: 20 μm).

CIHR Author Manuscript

CIHR Author Manuscript

CIHR Author Manuscript

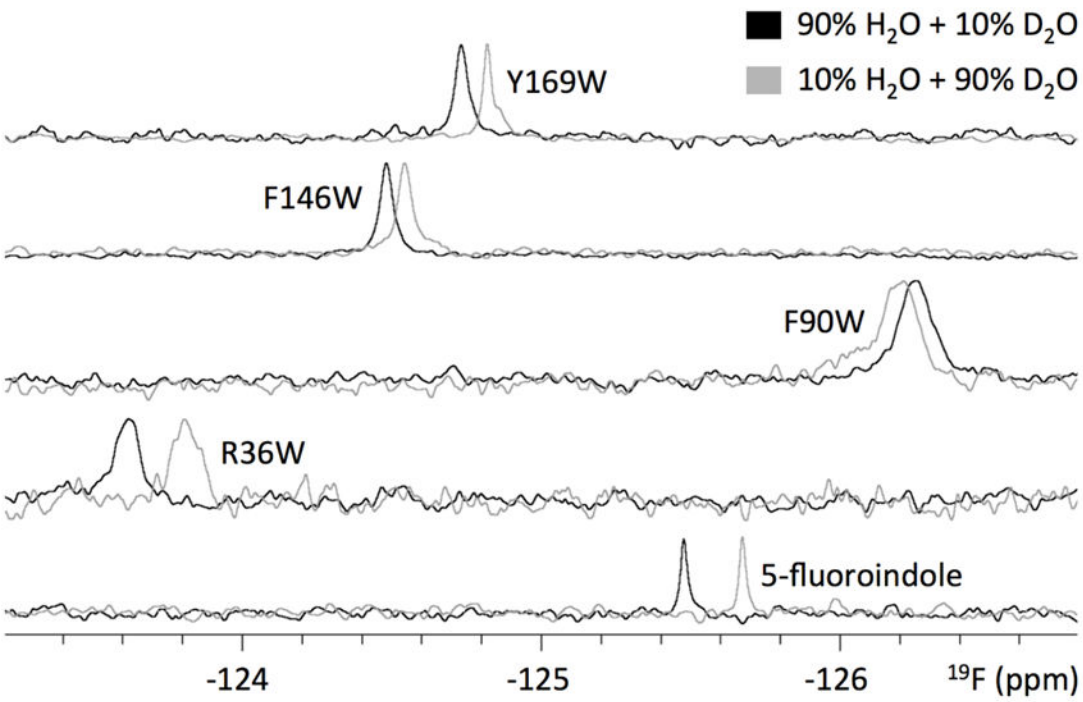


Figure 5.

1D ^{19}F-NMR spectra of W_1 5F-Trp-labeled mutants and free 5-fluoroindole as observed in indicated H_2O/D_2O mixture (W_1 mutants: 512 scans; 5-fluoroindole: 128 scans).

CIHR Author Manuscript

CIHR Author Manuscript

CIHR Author Manuscript

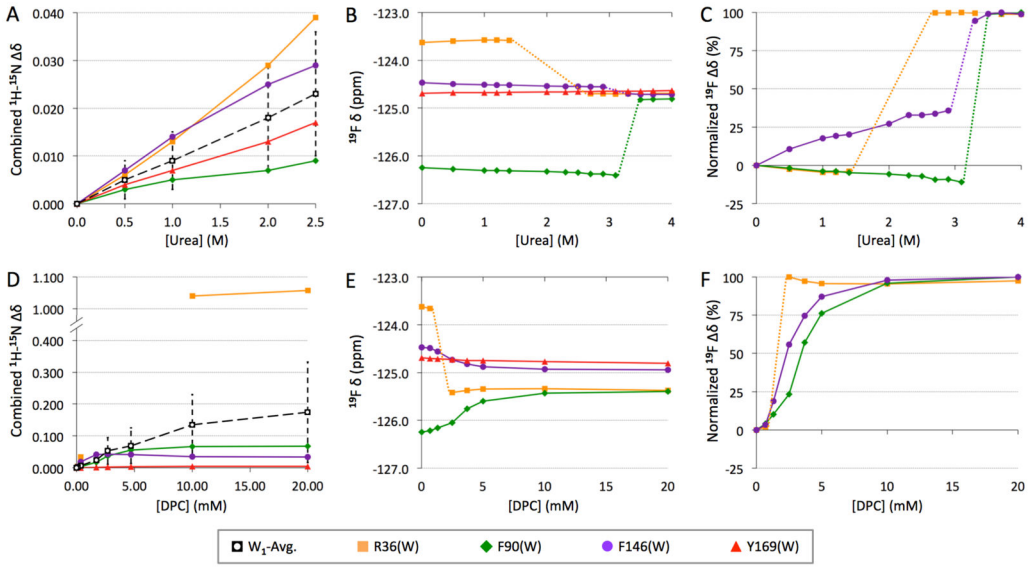


Figure 6.

Titration of wild type W_1 followed at indicated backbone amide site by heteronuclear NMR (A, D; combined chemical shift[34] perturbation based on our previously reported data[12]) and of indicated 5F-Trp-labeled mutants followed by ^{19}F NMR, with absolute (B, E) and normalized (C, F) ^{19}F chemical shift changes illustrated (spectra shown in Figure S8). W_1-Avg. represents average from all assigned residues (± average deviation). Dashed lines are used to join titration points when signals could not be tracked due to severe line broadening.

CIHR Author Manuscript

CIHR Author Manuscript

CIHR Author Manuscript

Table 1

Solvent accessible surface area (SASA) relative to total Trp side chain area and for indicated carbon in each W_1 mutant (average ± average deviation over 20-member structural ensembles).

	R36W	F90W	F146W	Y169W
Predicted side chain SASA (%)	4±4	30±3	71±14	93±10
C-4 SASA (Å^2)	0.0±0.0	1.2±1.1	14.4±5.7	17.4±6.9
C-5 SASA (Å^2)	1.7±2.5	8.0±5.6	30.4±8.2	33.2±6.0
C-6 SASA (Å^2)	2.8±3.2	24.1±7.5	34.7±4.3	36.3±3.0

CIHR Author Manuscript

CIHR Author Manuscript

CIHR Author Manuscript

Table 2

^{9}F chemical shifts (±0.01 ppm; δH_2O in 90%H_2O+10%D_2O and δD_2O in 10%H_2O+90%D_2O), SIS ($\Delta(\delta D_2O - \delta H_2O)$), T_1 relaxation time constant, and linewidth for 5F-Trp-labeled W_1 mutants alongside TFE and 5-fluoroindole controls.

	R36W	**F90W**	**F146W**	**Y169W**	**TFE**	**5-fluoroindole**
δH_2O (ppm)	−123.64	−126.27	−124.48	−124.73	−76.53	−125.48
δD_2O (ppm)	−123.81	−126.23	−124.54	−124.82	−76.64	−125.67
SIS (ppm)	−0.17	+0.04	−0.06	−0.09	−0.11	−0.19
T1 (s)	0.64±0.18	0.44±0.06	0.45±0.05	0.61±0.07		
Linewidth (Hz)	38.2±0.9	44.8±0.9	25.3±1.1	19.5±0.7		