Electronic Supplementary Material (ESI) for ChemComm.

Supporting Information

Measurement of ^{14}N Quadrupole Couplings in Biomolecular Solids Using Indirect-Detection ^{14}N Solid-State NMR with DNP.

James A. Jarvis[a,d], Ibraheem M. Haies[b,e], Moreno Lelli[c,f], Aaron J. Rossini[c,g], Ilya Kuprov[b], Marina Carravetta[b], Philip T.F. Williamson[a*]

[a]Centre for Biological Sciences, University of Southampton, Southampton, SO17 1BJ, United Kingdom.

[b]School of Chemistry, University of Southampton, Southampton, SO17 1BJ, United Kingdom.

[c]Centre de RMN à Tres Hauts Champs, Institut de Sciences Analytiques, Université de Lyon (CNRS/ENS Lyon/UCB Lyon 1), 69100 Villeurbanne, France.

[d]Current Affiliation: Department of Life Sciences, Imperial College London, London, UK

[e]Current Affiliation: Department of Chemistry, Mosul, Iraq.

[f]Current Affiliation: Center for Magnetic Resonance, University of Florence, Via L. Sacconi 6, 50019 Sesto Fiorentino, Italy.

[g]Current Affiliation: Department of Chemistry, Iowa State University, 0205 Hach Hall, 2438 Pammel Drive, Ames, US.

Contents

1. Sample Preparation

Samples of GB3 were prepared as follows. U-^{13}C labelled GB3 samples were prepared by growing recombinant GB3 transformed BL21(DE3) cells in minimal media enriched with ^{13}C glucose (CIL), and U-^{13}C,^{15}N labelled samples were prepared in minimal media enriched with ^{13}C glucose and ^{15}N ammonium chloride (CIL). In either case; BL21(DE3) cells were transformed with a pET expression vector containing a synthetic gene for GB3 under the control of a T7 promoter and an ampicillin resistance gene. A single colony of cells grown on agar containing 100 µg/mL ampicillin was added to 10 mL LB containing 100 µg/mL ampicillin and grown at 37 °C with 200 rpm shaking for 14 hours. This overnight colony was centrifuged for 5 mins at 4,000 g, and harvested cells were transferred to 500 mL 1xM9 media, containing 100 µg/mL ampicillin, 2 g L^{-1} $^{13}C_6$-glucose, 1 g L^{-1} NH_4Cl, and grown at 37 °C with 200 rpm shaking until the OD_{600}=0.6, at which point expression was induced by addition IPTG to a final concentration of 1 mM. After 4 hours, cells were harvested by centrifugation at 6900 g for 20 mins at 4 °C.

The cell pellet was re-suspended in 30 mL phosphate buffered saline (PBS) and lysed by sonication using a Heat Systems sonicator with a stud probe at power level 7 for 10 minutes, with a 20 seconds on 30 seconds off cycle. The insoluble fraction was pelleted by centrifugation at 16,000 g, 4 °C for 20 minutes.

The soluble protein containing supernatant was heated to 80 °C for 5 mins in a waterbath and then immediately chilled on ice for 10 minutes. Precipitated matter from the heating was pelleted by centrifugation at 20,000 g, 4 °C for 30 minutes. GB3 was isolated from the soluble fraction by precipitation with ammonium sulphate ($(NH_4)_2SO_4$) at 25 °C. $(NH_4)_2SO_4$ was added to the soluble fraction to 60% and stirred for 2 hours. Precipitated proteins were removed by centrifugation at 5500 g for 20 minutes. $(NH_4)_2SO_4$ was then added to the soluble fraction up to 90% with stirring for 2 hours. At this $(NH_4)_2SO_4$ concentration, GB3 was precipitated, and was recuperated by centrifugation at 5500 g for 20 minutes, before being dissolved in 10 mL 25 mM Tris, pH = 8.0.

Following desalting with a GE Healthcare PD-10 desalting column, anion exchange chromatography was used to remove contamination from nucleic acids. A GE Healthcare HiTrap Q HP 5 mL column packed with Q-sepharose on an ÄKTA Prime liquid chromatography system was used for anion exchange. The column was equilibrated with 2 column volumes of 25 mM Tris pH = 8.0, before loading the GB3 solution. The protein was eluted with gradient of NaCl concentration from 0-1 M, in Tris buffer, over 25 mL.

Microcrystals of purified GB3 were prepared by precipitation at 4 °C. Purified GB3 was buffer exchanged into 25 mM bis-Tris buffer, pH = 6.5 and concentrated to 60 mg/mL by centrifugal

filtration in a 5,000 Da MWCO Vivaspin spinning at 3700 g. To this, hexylene glycol (Sigma Aldrich, UK) was added to a final concentration of 20% (v/v). Microcrystals formed immediately, and were pelleted by centrifugation at 4000 g for 5 mins. For $^{13}C/^{15}N$ studies 30 mg pelleted microcrystals were typically packed into a Varian 3.2 mm rotor for ssNMR.

For samples containing DNP radical, the protein was buffer exchanged into 90% D_2O, 10% 25 mM bis-Tris buffer, pH 6.5, before being concentrated to 60 mg/ml prior to the addition of 20% (v/v) hexylene glycol. Microcrystals formed immediately, and were pelleted by centrifugation at 4000 g for 5 mins. 30 mg pelleted crystals were resuspended in the 90% D_2O, 10% 25mM bis-Tris supernatant, and a solution of 20 mM AMUPol (15-{[(7-oxyl-3,11-dioxa-7-azadispiro[5.1.5.3]hexadec-15-yl)carbamoyl][2-(2,5,8,11-tetraoxatridecan) in 70% glycerol-d_6:30% D_2O was added to the sample to a final concentration of 2 mM and 12.5 mM AMUPOL. The samples were then packed into Bruker 3.2 mm sapphire rotors.

2. Experimental Conditions and Pulse Sequences

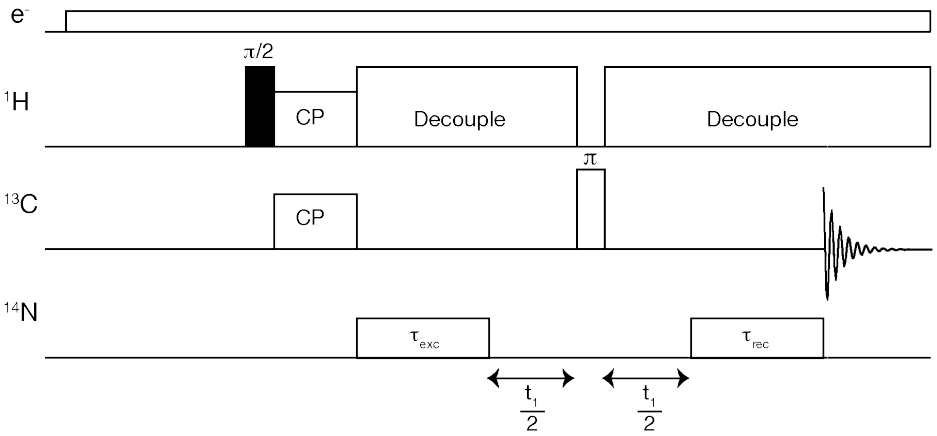

Figure S1 – Pulse sequence used for acquisition of $^{13}C/^{14}N$ correlation spectra

All DNP experiments were performed on a Bruker Avance III spectrometer operating at 18.8 T (800 MHz of ^{1}H Larmor frequency) equipped with a gyrotron oscillator at 527 GHz and a low temperature, triple-resonance, 3.2 mm MAS probe tuned to $^{1}H/^{13}C/^{14}N$ Larmor frequencies. Spectra were recorded at a temperature of 100 K and at a MAS frequency of 13.5 kHz. Data processing was performed in NMRPipe[1] and MatNMR[2]. In all experiments, TPPM[3] proton decoupling was employed at 100 kHz. Cross polarization (CP) from protons to ^{13}C was performed used ramped-CP with a ^{13}C spin-lock field of 50 kHz. The proton amplitude was ramped from 70 to 100% and adjusted to give maximal intensity at the n=+1 Hartmann-Hahn condition. The CP contact time was 1.75 ms

The pulse sequence shown in Figure S1 [4] was used for acquisition of $^{13}C/^{14}N$ filtered and 2D correlation spectra. The excitation and reconversion pulses applied to ^{14}N were rotor synchronized and applied for 1.7 ms (23 rotor periods) at an RF amplitude of 30 kHz. The refocusing pulse to ^{13}C was applied with a 80 kHz rf amplitude. ^{14}N-filtered 1D spectra were acquired without t_1 evolution. Two dimensional data were acquired in a phase sensitive manner using States-TPPI[5], with 76 rotor synchronized increments acquired with 640 scans per increment. The recycle delay was 6 s. Spectra were zerofilled to 2048 points in either dimension, and exponential linebroadening of 40 Hz was applied in the F1 dimension before 2D Fourier transform.

MAS-DNP ^{13}C-^{13}C proton driven spin-diffusion spectra were acquired under identical conditions with a 50 ms mixing period to facilitate transfer of magnetization between ^{13}C sites within the protein. Data were acquired with States-TPPI, with 128 points in the indirect dimension which was subsequently linear predicted by a further 128 points. Spectra were zerofilled to 2048 points in either dimension, and exponential linebroadening of 70 Hz was applied in the F1 dimension before 2D Fourier transform.

Spectra are referenced to liquid ammonia at 0 ppm in the ^{14}N dimension, using a secondary reference of solid ammonium chloride at 39.3 ppm[6]. The ^{13}C dimension is referenced to DSS, using a secondary reference of the downfield peak of adamantane at 40.48 ppm [7].

$^{13}C/^{15}N$ experiments were performed on a Varian DD2 spectrometer operating at 14.1 T equipped with a triple resonance 3.2 mm MAS probe. Both NCO and NCA spectra were acquired at 0 °C with 12.5 kHz spinning. Spectra were acquired using the BioSolidsPack, with spin lock fields of $(5/2)\omega_r$ and $(3/2)\omega_r$ for the ^{15}N and ^{13}C during cross polarization. During both direct and indirect acquisition 120 kHz SPINAL proton decoupling was applied. The spectra were acquired with States-TPPI[5] with 128 increments in the indirect dimension and 32 scans per increment. Spectra were processed in NMRPipe[1] using 150 Hz linebroadening in the both dimensions to mimic the effects of inhomogeneous broadening in the DNP measurement and analyzed using CCPN Analysis[8].

3. Effects of DNP conditions on spectral resolution

The resolution observed in the CP spectra of GB3 at 100 K shows reduced resolution compared to spectra obtained at 298 K, where the linewidths are in the order of 30 Hz. Despite this, there appears to be little difference in the resolution between spectra obtained from samples doped with 2 mM and 12.5 mM AMUPol. To better assess the nature of the broadening observed, a 2D proton driven spin-diffusion (PDSD) spectra of microcrystalline GB3 was acquired under DNP conditions with a concentration of AMUPol biradical of 12.5 mM (Supplementary Figure S2).

The diagonal in the 2D-PDSD provides insights into the nature of the linebroadening, with inhomogeneous broadening resulting in a distribution of peaks along the diagonal direction and any homogeneous contribution resulting in a thickening of the diagonal measured in the direction of the antidiagonal [9]. Inspection of the anti-diagonal reveals peaks with a homogeneous linewidth of ~1 ppm distributed along the diagonal. The absence of any variation in the homogeneous linewidth along the diagonal, suggests that the homogeneous broadening experienced throughout the protein is relatiavely uniform; although the absence of site specific relaxation measurement prohibits a more detailed analysis. With homogeneous broadening contributing only 1 ppm to the linewidth, this indicates that the major influence on the linewidths observed in the spectra arises from an increase in inhomogeneous broadening resulting from the trapping of a range of conformational states upon freezing the sample at 100 K. In contrast, paramagnetic relaxation enhancement arising from the presence of the biradical in the sample, which would manifest in an increase in the homogeneous linewidth, appears to contribute less than 1 ppm to the linewidth. These observations are consistent with previous DNP enhanced solid-state NMR studies on biomolecules at cryogenic temperatures which have clearly shown that the broadening of solid-state NMR signals is primarily inhomogeneous in nature.

Interestingly though, the contributions to inhomogeneous broadening does not appear to be uniform throughout the protein. For example, in the aromatic region the resonances attributed to a unique tryptophan sidechain can be well resolved, with linewidths of ~0.9 ppm. Analysis of the crystal structure (PDB code: 2LUM) reveals this tryptophan to be buried in the core of the protein with little conformational freedom. Such an observation would be consistent with the analysis of other rigid proteins structures, such as the Type-III secretion needle, that exhibit limited conformational heterogeneity at low temperature and well resolved resonances up to ~1 ppm of linewidth can be observed under DNP conditions[10].

We note that the resolution apparent in these spectra appears somewhat more favorable than that observed from the homologous protein GB1 at 100 K[11]. Given the correlation

observed between protein dynamics room temperature and the degree of inhomogeneous broadening observed at 100 K [10] this would suggest that a careful comparison of the dynamics between GB1 and GB3 may provide additional insights into how protein dynamics influence the range of conformational states upon freezing to cryogenic temperature and the potential degree of inhomogeneous broadening observed.

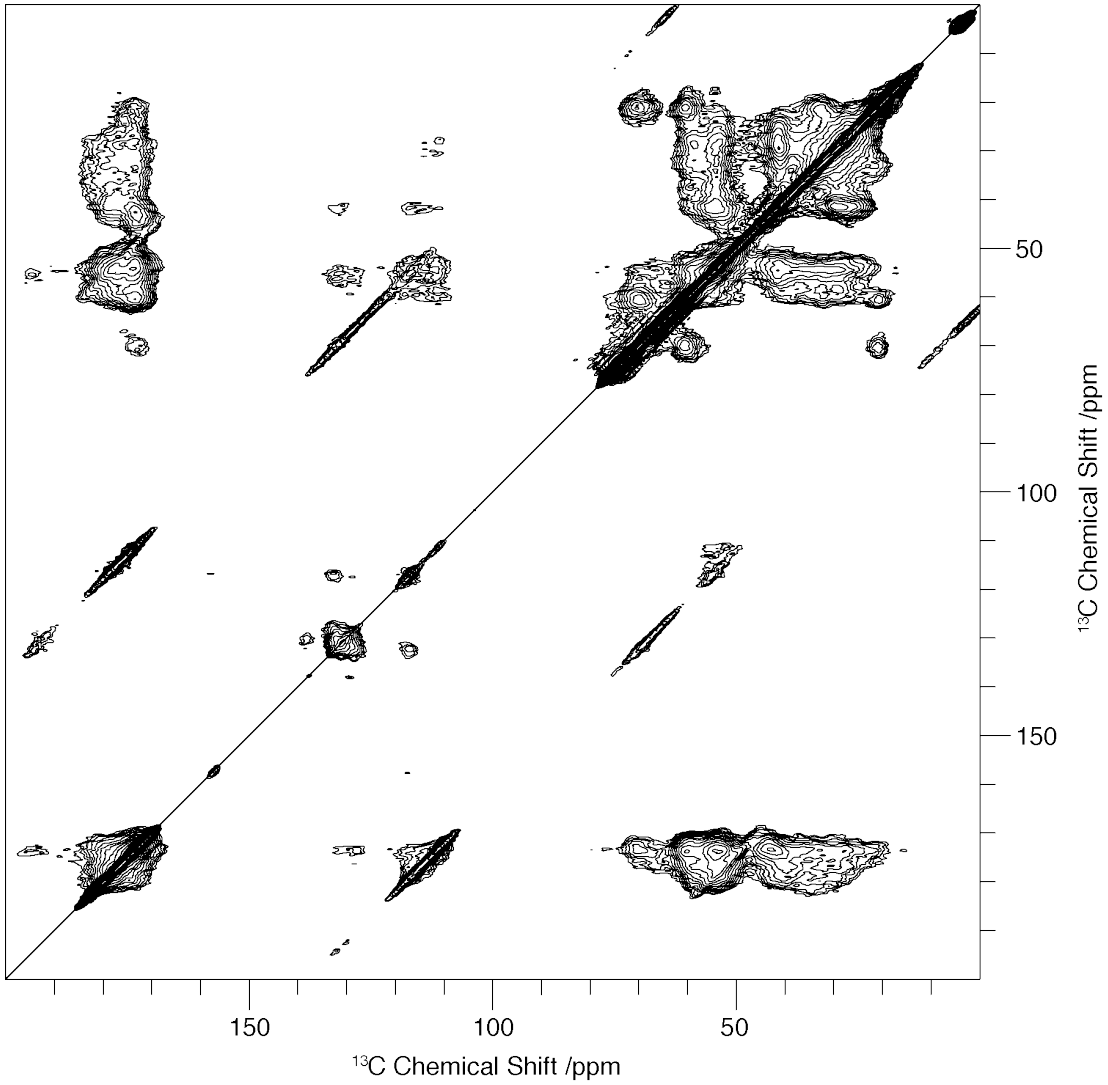

Figure S2. DNP enhanced 2D proton driven spin-diffusion (PDSD) ^{13}C-^{13}C homonuclear correlations spectrum of GB3 in 12.5 mM AMUPol at 100 K, acquired at 13.5 kHz.

4. Calculated $^{13}C/^{14}N$ Spectra of GB3

To aid in interpretation of 2D $^{13}C/^{14}N$ spectra recorded on GB3 with DNP, calculated $^{13}C/^{14}N$ spectra of GB3 were constructed using the $^{13}C/^{15}N$ NCa and NCO spectra shown in Figure 1. To each of the assigned peaks in the NCA and NCO spectra a SOIQS was added based on the secondary structure of each residue as ascertained from the CSI of the assigned spectra. The expected shifts were estimated from the known ^{14}N EFG tensor parameters of NAV, C_Q=3.21MHz, η=0.32[12], whose ^{14}N site is structurally similar to those found in the peptide backbone of proteins. This was assumed to be the magnitude of the ^{14}N C_Q if the site was expected to be a β-sheet. According to the study of Fukazawa *et al*.[13], 200 kHz was added to the β-sheet C_Q of 3.21 MHz if the site was expected to be in an α-helix. Those sites not predicted to be in an α-helical or β-sheet conformation we assigned a C_Q of 130 kHz less than that of the β-sheet, due to the observation of a number of resonances on the high-field edge of the experimental ^{14}N spectra which appear to have a C_Q too low for the expected values for α-helix or β-sheet. At all sites, η was kept constant and assumed to be η=0.3, which is within η=±0.01 of amide sites found in a number of peptides and acetylated amino acids, such as NAV[12], triglycine[14], AGG and AAG[15]. From the equation for the ^{14}N δ_Q^{iso} given in Equation (1) in the main text, the expected ^{14}N δ_Q^{iso} at B_0 = 18.8 T for each type of secondary structure was calculated to be ^{14}N δ_Q^{iso} = 261 ppm for an α-helix (C_Q = 3.41 MHz, η = 0.3), ^{14}N δ_Q^{iso} = 237 ppm for a β-sheet (C_Q = 3.21 MHz, η=0.3), and ^{14}N δ_Q^{iso}= 219ppm for an random coil (C_Q = 3.08 MHz, η = 0.3). Spectra were calculated in the time domain using ^{13}C shifts and ^{15}N shifts with secondary structure dependent 'expected' ^{14}N shifts added. A sinebell squared window function was applied in the time domain to the generated signals and a further 150 Hz linebroadening added in either dimension to simulate the effect of cryogenic temperatures on ^{13}C linewidths, and cryogenic temperatures and second order broadening effects on ^{14}N linewidths. Data were zerofilled to 1024 points and Fourier transformed in both dimensions.

As demonstrated in Figure S3, the range and intensity distribution in ^{14}N shifts observed experimentally is reflected in the spectra calculated spectra of GB3 based on C_Q's obtained from model compounds and predicted variability in EFG's determined from Overtone studies of model peptides with known secondary structure[13]. Furthermore, the absence of significant differences between the C_Q's measured in GB3 and model crystalline systems[14, 15], suggests limited dynamic averaging of the ^{14}N quadrupolar interactions within the protein backbone at these temperatures, allowing the SOIQS to be interpreted largely in terms of backbone conformation. Despite similarities in the distribution of intensity of the ^{14}N dimension there are some notable differences including differences in the intensity profile of the signal envelope and the presence and absence of peaks in the simulations that are

present in the experimental data. This in part may reflect the rather simple approximations of the C_Q used in these calculations, which fails to adequately represent the subtle variations in C_Q within the protein structure, leading to changes in the intensity profile of the ^{14}N signal envelope and the presence of a small number of additional resonances with smaller ^{14}N shifts (e.g. resonances with ^{14}N shifts between 320 to 340 ppm and a ^{13}C shift of ~54 ppm). However, some resonances that are expected to be resolved in the ^{13}C spectrum such as those of glycine residues (~40-45 ppm) and the well resolved threonine residue at ~68 ppm are absent in the $^{14}N/^{13}C$ correlation spectrum. Studies on model peptides indicate that the quadrupolar interaction for such residues should be similar to other amino acids, and given the broadbanded nature of these experiments it is unlikely that these residues are absent due to inefficient excitation. Accordingly we attribute their absence to the signal to noise obtained in these experiments. This is particularly acute for glycine resonances as the due to the shorter T_2 they exhibit due to the strong $^{1}H/^{13}C$ dipolar couplings within the backbone CH_2 groups which are poorly suppressed with the ^{1}H decoupling fields available with the current experiment apparatus.

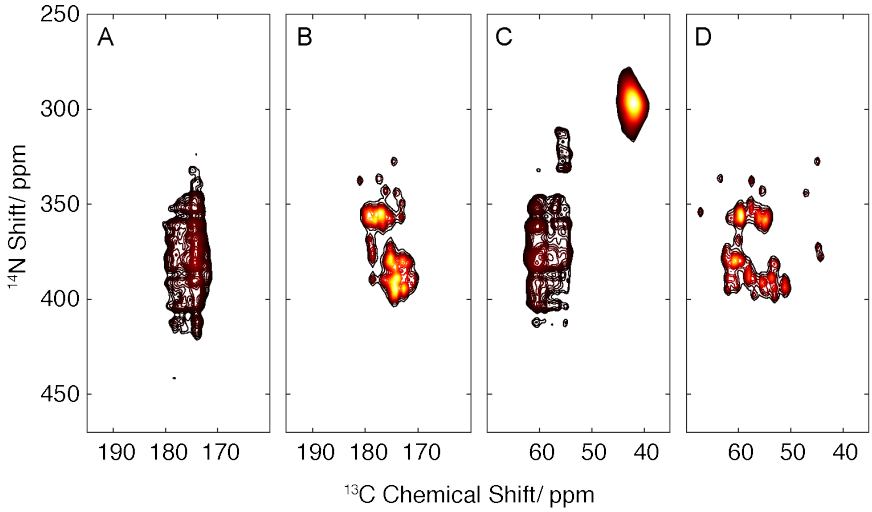

Figure S3 – Experimental 2D $^{13}C/^{14}N$ spectrum of microcrystalline GB3 containing 2 mM AMUPol is shown in panels (A) and (C), which are carbonyl and aliphatic regions, respectively. 'Calculated' spectra for the same regions are shown in panels (B), carbonyl, and (D), alipahtic. Experimental spectrum acquired at 18.8 T, 100 K, 13.5 kHz MAS with DNP enhancement. Experimental data plotted with a lower contour threshold than in the main paper to emphasize the full distribution of ^{14}N shifts observed.

References

1. Delaglio, F., et al., *NMRPipe: a multidimensional spectral processing system based on UNIX pipes.* J Biomol NMR, 1995. **6**(3): p. 277-93.
2. van Beek, J.D., *matNMR: a flexible toolbox for processing, analyzing and visualizing magnetic resonance data in Matlab.* J Magn Reson, 2007. **187**(1): p. 19-26.
3. Bennett, A.E., et al., *Heteronuclear decoupling in rotating solids.* The Journal of Chemical Physics, 1995. **103**(16): p. 6951.
4. Jarvis, J.A., et al., *An efficient NMR method for the characterisation of 14N sites through indirect 13C detection.* Phys Chem Chem Phys, 2013. **15**(20): p. 7613-20.
5. Marion, D., et al., *Rapid Recording of 2d Nmr-Spectra without Phase Cycling - Application to the Study of Hydrogen-Exchange in Proteins.* Journal of Magnetic Resonance, 1989. **85**(2): p. 393-399.
6. Bertani, P., J. Raya, and B. Bechinger, *15N chemical shift referencing in solid state NMR.* Solid State Nucl Magn Reson, 2014. **61-62**: p. 15-8.
7. Morcombe, C.R. and K.W. Zilm, *Chemical shift referencing in MAS solid state NMR.* Journal of Magnetic Resonance, 2003. **162**(2): p. 479-486.
8. Vranken, W.F., et al., *The CCPN data model for NMR spectroscopy: development of a software pipeline.* Proteins, 2005. **59**(4): p. 687-96.
9. Taylor, G.F., et al., *Morphological differences between beta(2) -microglobulin in fibrils and inclusion bodies.* Chembiochem, 2011. **12**(4): p. 556-8.
10. Fricke, P., et al., *High resolution observed in 800 MHz DNP spectra of extremely rigid type III secretion needles.* J Biomol NMR, 2016. **65**(3-4): p. 121-6.
11. Lewandowski, J.R., et al., *Protein dynamics. Direct observation of hierarchical protein dynamics.* Science, 2015. **348**(6234): p. 578-81.
12. Stark, R.E., R.A. Haberkorn, and R.G. Griffin, *14N NMR determination of NH bond lengths in solids.* The Journal of Chemical Physics, 1978. **68**(4): p. 1996.
13. Fukazawa, J., et al., *(14)N Quadrupolar Coupling of Amide Nitrogen and Peptide Secondary Structure As Studied by Solid-State NMR Spectroscopy.* Journal of the American Chemical Society, 2010. **132**(12): p. 4290-4294.
14. Strohmeier, M., D.W. Alderman, and D.M. Grant, *Obtaining molecular and structural information from C-13-N-14 systems with C-13 FIREMAT experiments.* Journal of Magnetic Resonance, 2002. **155**(2): p. 263-277.
15. Dib, E., T. Mineva, and B. Alonso, *Recent Advances in 14N Solid-State NMR.* 2015.