Science Advances
AAAS

Manuscript Template

FRONT MATTER

Title

Full Title:

Expanding the Tool Box for Native Structural Biology: ^{19}F Dynamic Nuclear Polarization with Fast Magic Angle Spinning

Short title:

^{19}F DNP MAS NMR for in-cell structural biology

Authors:

Kumar Tekwani Movellan[1,2], Wenkai Zhu[2,3], Daniel Banks[4], James Kempf[4], Brent Runge[1,2], Angela M. Gronenborn[1,2,3,5,6*], Tatyana Polenova[1,2,3*]

Affiliations

[1]University of Delaware, Department of Chemistry and Biochemistry, Newark DE 19716, United States.
[2]Pittsburgh Center for HIV Protein Interactions, University of Pittsburgh School of Medicine, 1051 Biomedical Science Tower 3, 3501 Fifth Avenue, Pittsburgh PA 15261, United States.
[3]Department of Structural Biology, University of Pittsburgh School of Medicine, 3501 Fifth Avenue, Pittsburgh PA 15261, United States.
[4]Bruker Biospin Corporation, 15 Fortune Drive, Billerica, Massachusetts MA 01821, United States
[5]Department of Bioengineering, Swanson School of Engineering, University of Pittsburgh, Pittsburgh PA 15261, United States.
[6]Department of Chemistry, Dietrich School of Arts and Sciences, University of Pittsburgh, Pittsburgh PA 15261, United States.

Abstract

Obtaining atomic-level information on components in the cell is a major focus in structural biology. Elucidating specific structural and dynamic features of proteins and their interactions in the cellular context is crucial for understanding cellular processes. We introduce ^{19}F dynamic nuclear polarization (DNP) combined with fast magic-angle-spinning (MAS) NMR spectroscopy as a powerful technique to study proteins in mammalian cells. We demonstrate our approach on the SARS-CoV-2 5F-Trp-N^{NTD} protein, electroporated into human cells. DNP signal enhancements of 30- to 40-fold were observed, translating into over 1000-fold time-savings in experiment time. High signal-to-noise ratio spectra were acquired on nanomole-quantities of a protein in cells in minutes. 2D ^{19}F-^{19}F dipolar correlation spectra with remarkable sensitivity and resolution were obtained, exhibiting ^{19}F line widths as narrow as ~2 ppm, and ^{19}F-^{19}F cross-peaks associated with fluorine atoms as far as ~10 Å apart. This work paves the way for ^{19}F DNP-enhanced MAS NMR applications in cells for probing protein structure, dynamics and ligand interactions.

Teaser

^{19}F DNP-enhanced MAS NMR is a powerful tool for protein structural characterization in cellular environments.

MAIN TEXT

Introduction

Advancing structural biology today requires the characterization of structure, dynamics and interactions of biomolecules in their native environment, the cell. Up to now, most of our structural knowledge has been garnered *in vitro*, yielding a tremendous number and variety of complex biomolecular structures that advanced our understanding of biological processes and biochemical pathways. The majority of structures up to date have been provided by X-ray crystallography, although cryo-electron microscopy (EM) and nuclear magnetic resonance (NMR) spectroscopy have also contributed (*1*, *2*). A major and unique feature of NMR is its ability to inform on protein conformational changes and dynamics at atomic resolution over a broad range of timescales (*3*, *4*).

Solution *in-cell* NMR has opened the way to transfer the unique capabilities of NMR into the cellular context (*5*), and over the last few years ^{19}F *in-cell* NMR has gained popularity (*6*–*8*), given the beneficial spectroscopic properties and high sensitivity of the ^{19}F spin, coupled to its virtual absence in biology. Moreover, fluorine can be readily introduced into biological macromolecules (*9*). Remarkably, ^{19}F resonances can be detected in the *in-cell* spectra of globular proteins when the commonly used ^{1}H-^{15}N HSQC spectra are invisible, with resonances broadened beyond detection due to protein interactions with cellular components (*6*).

Here, we introduce ^{19}F dynamic nuclear polarization (DNP)-enhanced magic angle spinning (MAS) NMR for investigating proteins in mammalian cells. We present results on the N-terminal domain of SARS-CoV-2 nucleocapsid protein, N^{NTD} (**Fig. 1**). We observed 30- to 40-fold signal enhancements of the ^{19}F resonance intensities by DNP in cells at fast spinning frequencies of 40 kHz, enabling the detection of low-nanomole quantities of 5F-Trp-N^{NTD} with a signal-to-noise ratio greater than five in spectra recorded in ~20 minutes. Similarly, ^{19}F-^{19}F 2D spectra, recorded in ~33 hours, would have required ~4.5 years without DNP, making such an experiment unrealistic. With the homogenous line widths as narrow as 2 ppm, well-resolved cross peaks were observed in 2D ^{19}F-^{19}F dipolar correlation spectra. A unique cross peak was unambiguously assigned to the correlation associated with residues W70 and W94 separated by ~10 Å.

In addition to improving sensitivity, the approach presented here overcomes other major challenges for *in-cell* ^{19}F NMR-based cellular structural biology, including safeguarding cell viability and the introduction of radicals for *in-cell* ^{19}F-DNP, as well as taking advantage of spectral simplification by using specifically introduced ^{19}F atoms. Taken together, our results demonstrate the promise of ^{19}F *in-cell* DNP-enhanced fast MAS NMR for structural investigations of proteins in mammalian cells.

Results

Sensitivity of *in-cell* ^{19}F DNP-enhanced MAS NMR experiments

N^{NTD} was labeled with 5F-Trp and delivered into human A2780 cells by electroporation. ^{19}F DNP-enhanced MAS NMR spectra for the samples containing 0.3 and 0.8 nanomoles of protein with 13 mM and 6.3 mM AMUPol, respectively, are shown in **Fig. 2A, B**. Remarkably, signals can be detected in a few scans (**Fig. 2A**), while no signal was present in the control 'microwave-off' experiment after 85 minutes of signal averaging on a sample containing 0.8 nanomoles of protein and 6.3 mM AMUPol (**Fig. 2B**). We

obtained ^{19}F DNP enhancements 30- to 40-fold at the MAS frequency of 40 kHz (Supplementary **fig. S1A)**.

Detection of ^{19}F signals in the *in-cell* MAS NMR experiments without DNP enhancement requires much longer measurement times and/or more sample. Signal-to-noise ratios (SNR) similar to those in DNP-enhanced MAS experiments at room temperature required 32.7 hours at 11.7 T and 5.3 days at 20.0 T (**Fig. 2C and D**). Note that the sample used for experiments at 11.7 T contained ~1.2-1.5 million cells (~1 nanomole of protein) in a 1.3 mm rotor while the sample used for experiments at 20.0 T contained ~4-5 million cells (~1.9 nanomoles of protein) in a 1.9 mm rotor. In principle, long acquisition times should not affect spectral quality adversely if the sample is maintained at -80 ºC or below with appropriate cryoprotectant (*10*, *11*). Unfortunately, in the 11.7 and 20.0 T systems without LT-MAS probes, the lowest attainable temperatures with the 1.3 mm HFX probe at 30 kHz MAS frequency and the 1.9 mm HX probe at 20 kHz MAS frequency are -11 ºC and -2 ºC, respectively, conditions under which cells are not viable for long times.

To systematically evaluate the *in-cell* ^{19}F DNP sensitivity, we examined the dependence of the normalized signal intensities and buildup times on the MAS frequency and AMUPol concentration. It is well known that nitroxide-based biradicals, commonly used for cross-effect DNP, are unstable in the reducing environment of the cell, as demonstrated by the groups of McDermott (*12*), Frederick (*13*), and Debelouchina (*14*). Therefore, we evaluated several procedures for introducing AMUPol into the sample: i) electroporating AMUPol solution together with the protein into the cells; ii) adding the AMUPol solution to the cells after protein electroporation and recovery; and iii) combining i) and ii) by electroporating AMUPol into the cell with the protein, followed by introducing additional AMUPol solution to the cells after recovery. The final concentration of AMUPol in the samples, as measured by EPR, depends critically on the incorporation procedure for the biradical, and the results are summarized in **Table 1** and detailed in Supplementary **Table S3 and Table S4**. The highest sensitivity in the DNP-enhanced experiments resulted from an in-cell concentration of 13 mM AMUPol, which was reached by electroporating 40 mM AMUPol solution with the protein into the cells, followed by the addition of 40 mM AMUPol solution to the sample after cell recovery.

The ^{19}F DNP signal buildup times for the *in-cell* samples containing <3, 6.3, and 13 mM AMUPol are similar at 14.6, 13.3, and 16.2 s, respectively (**Table 1** and **Fig. 2D**). Importantly, the ^{1}H signal buildup time for the sample containing 13 mM AMUPol is only ~1.7 s, i.e., approximately ten-fold slower (**Table 1**). These ^{19}F DNP signal buildup times are similar to those determined in our prior study on HIV-1 CA assemblies (*15*) and indicate that polarization transfer occurs directly from electrons to ^{19}F nuclei and is not driven by ^{1}H spin diffusion.

The dependence of the normalized ^{19}F signal intensity on the MAS frequency is provided in **Fig. 2E** and Supplementary **fig. S1C** and shows that increasing the MAS frequency from 20 to 30 kHz results in ~10% signal intensity-gain in all the samples. Interestingly, increasing the spinning frequency from 30 to 40 kHz produced no enhancement in signal intensity for the <3 and 6.3 mM AMUPol-containing samples, whereas the signal intensity increased by about 25% in the sample containing 13 mM AMUPol. Conversely, a ~60% drop in ^{13}C DNP signal enhancements was observed upon increasing the MAS frequency from 20 to 40 kHz, as illustrated for the carbonyl signal intensity in the sample containing 6.3 mM AMUPol (Supplementary **fig. S6**). This finding is in agreement with results from non-DNP-based experiments (*16–18*) and underscores the benefits of fast spinning frequencies (40 kHz) also for ^{19}F DNP-enhanced MAS NMR .

To elucidate the best conditions for the *in-cell* ^{19}F DNP experiments, we calculated SNR_{norm}, the SNR of the most intense peak, per nanomole of protein per square root of

experimental time. A maximum of SNR_{norm} of ~0.5 was seen in samples in which AMUPol was added extracellularly, while SNR_{norm}>3 was observed for samples where AMUPol was introduced into the cell by electroporation as well as those in which AMUPol was electroporated into the cell, followed by the addition of AMUPol solution to the cells after recovery (**Table 2**).

Overall, our results suggest that electroporation of the biradical into the cells together with the protein, followed by additional biradical addition to the extracellular buffer, is best for attaining highest AMUPol concentrations. It should be pointed out, however, that it is unclear at present whether the variation in the DNP signal enhancements seen here is representative, since the spectra were recorded at different times, i.e., weeks apart, and may solely reflect the varied performance of the instrument. Further systematic experiments beyond the scope of this manuscript are necessary to fully evaluate and select an optimal procedure. Meanwhile, it is both noteworthy and encouraging that the results presented here demonstrate that AMUPol-mediated DNP-gains are achieved either by introducing this widely used polarizing agent into the cell or the surrounding medium, or a combination of both.

Signal assignments, spectral resolution, and homogeneous line widths

The apparent resolution of the ^{19}F DNP-enhanced *in-cell* MAS NMR spectra of the WT N^{NTD} is high, with the overall spectral envelope spanning over 12 ppm and with multiple partially resolved resonances of ~2-3 ppm line width (**Fig. 3A** and **3C**). Resonance assignments of the ^{19}F signals were obtained using three protein variants in which single tryptophan residues were substituted by phenylalanine, referred to as 5F-Trp,U-^{15}N-N^{NTD} W14F, 5F-Trp,U-^{15}N-N^{NTD} W70F, and 5F-Trp,U-^{15}N-N^{NTD} W90F (**Fig. 3B**). In all variants, like for the WT, ^{19}F incorporation was over 95%, as assessed by mass spectrometry (Supplementary **fig. S2**). The structural integrity of the variants was assessed via solution ^{1}H-^{15}N HSQC spectroscopy and RNA binding by electrophoretic mobility shift assays. No notable differences were noted in the ^{1}H-^{15}N HSQC spectra, compared to the spectrum of the WT N^{NTD} (Supplementary **fig. S3**). Similarly, there are no noteworthy effects on the RNA binding (Supplementary **fig. S4**).

Note that only small chemical shift differences are observed between the *in-vitro* and *in-cell* spectra and the overall peak pattern in the MAS spectra is similar. ^{19}F chemical shifts for all samples are summarized in Supplementary **table S5**.

To determine the homogeneous ^{19}F line widths in the MAS spectra, a series of ^{19}F DNP-enhanced spectra with selective magnetization inversion pulses were recorded, using "delays alternating with nutation for tailored excitation" (DANTE) sequence (*19*). The inverted individual peaks are 2-3 ppm broad (**Fig. 3C**), which corresponds to the upper limit of the homogeneous line widths.

The above results are very encouraging for future *in-cell* ^{19}F applications. They demonstrate that inhomogeneous line broadening is the main factor determining the line widths and, therefore, further improvements in resolution can be expected with dedicated ^{19}F DNP MAS NMR probes that permit ^{1}H decoupling, faster spinning frequencies and use at higher magnetic fields.

DNP-enhanced 2D ^{19}F-^{19}F correlation spectroscopy

Given the high sensitivity of the ^{19}F *in-cell* DNP-enhanced MAS NMR experiments and the narrow homogeneous line widths, we recorded *in-cell* 2D ^{19}F-^{19}F spin diffusion (SD) spectra on the 5F-Trp,U-^{15}N-N^{NTD} sample, containing 13 mM AMUPol. The spectrum acquired in 32 hours with a SD mixing time of 2 s is shown in **Fig. 3D**, left panel. Strong cross peaks are present on both sides of the diagonal, corresponding to a correlation between 5F-Trp-70 and 5F-Trp-94. In contrast, the control spectrum recorded in 28 hours with no mixing is devoid of cross peaks. The corresponding interfluorine distance between 5F atoms

of Trp-70 and Trp-94 is ~9.6±0.9 Å, much shorter than those between the 5F atoms of Trp-14 and Trp-70 (14.2±1.2 Å) or those between Trp-14 and Trp-94 (18.4±3.0 Å). Importantly, we employed a SD-based magnetization transfer since the probe does not permit simultaneous use of ^{1}H and ^{19}F channels, thus limiting the accessible mixing sequences. Moreover, given that ^{19}F isotropic chemical shifts of 5F-Trp-70 and 5F-Trp-94 are less than 3 ppm apart, other dipolar-mixing schemes, such as RFDR, are inefficient (*20*).

Discussion

NMR spectroscopy is currently one of the few non-destructive techniques that can provide atomic details of protein structure, dynamics, and interactions in living cells without the need for potentially perturbing labels. Yet, the low sensitivity and the high cellular background from naturally occurring ^{1}H, ^{13}C and ^{15}N atoms presents numerous challenges in the application of *in-cell* NMR. These challenges can be overcome by the use of ^{19}F *in-cell* DNP-enhanced fast MAS NMR spectroscopy. The here detailed 30- to 40-fold sensitivity enhancements by DNP, combined with ~2-3 ppm homogeneous line widths, observed without ^{1}H decoupling in 40 kHz MAS spectra, open doors for performing 2D and 3D spectroscopy with only ~0.3 nanomoles of protein in the MAS NMR rotor, as demonstrated here.

Overall, our study established a proof of concept for and will inspire further *in-cell* ^{19}F DNP applications. Indeed, we anticipate rapid improvements, especially with advent of DNP dedicated fluorine probes, capable of ^{1}H decoupling and ^{1}H-^{19}F cross polarization transfers. These added capabilities will further improve sensitivity and resolution, reduce polarization buildup times, and will permit the acquisition of heteronuclear-based correlation experiments. Additional sensitivity and resolution gains are expected to arise from higher magnetic fields and faster MAS frequencies. Coupled to the development of superior radicals for *in-cell* applications as well as expanding ^{19}F labeling strategies will make ^{19}F *in-cell* DNP-enhanced MAS NMR more broadly accessible to non-specialized researchers.

Materials and Methods

Sample preparation

Expression and purification of SARS-CoV-2 N^{NTD}

All 5F-Trp-N^{NTD} proteins, 5F-Trp,U-^{13}C,^{15}N-N^{NTD} wild type (WT) as well as 5F-Trp,U-^{15}N-N^{NTD} WT and W14F, W70F, and W94F variants were expressed and purified using a similar protocol as previously described for N^{NTD} (*21*, *22*). Briefly, *E. coli* BL21 Rosetta (DE3) cells harboring the recombinant plasmid from GenScript for expressing SARS-CoV-2 N^{NTD} (residues 40-174, current construct residue numbering 2-136) sub-cloned into a pET28a(+) vector fused with an N-terminal hexahistidine tag, followed by a TEV cleavage site, His_6-TEV-N^{NTD}, were used. Cells were grown to an OD_{600} of ~1 in ~5 mL, of Luria Bertani (LB) medium, supplemented with 50 µg/mL Kanamycin and 30 µg/mL chloramphenicol. 1 mL of LB starting culture was added to ~50 mL of M9 media, supplemented with 1 g/L $^{15}NH_4Cl$ (U-^{15}N-N^{NTD}) or 1 g/L $^{15}NH_4Cl$ and 2 g/L U-$^{13}C_6$-glucose (U-^{13}C,^{15}N-N^{NTD}) and grown overnight at 37 ºC with shaking at 170 rpm. 1 L of M9 medium, supplemented with 1 g/L $^{15}NH_4Cl$ (U-^{15}N-N^{NTD}), or 1 g/L $^{15}NH_4Cl$ and 2 g/L U-$^{13}C_6$-glucose (U-^{13}C,^{15}N-N^{NTD}), was seeded with the M9 starting culture to an OD_{600} of 0.1 and grown at 37 ºC to OD_{600} of 0.7-0.8. At that time 20-25 mg of 5-fluoroindole (in 70% ethanol) were added, and the temperature was lowered from 37 to 25 ºC. After 45 min, protein expression was induced by the addition of 0.5 mM IPTG and the culture was grown for an additional ~16-18 hours at 25 ºC. The cells were harvested by centrifugation at 5,000 x g for 10 min at 4 ºC, and the cell pellet was resuspended in the lysis buffer (Buffer A: 20

mM HEPES, 500 mM NaCl, pH 8) and stored at -80 ºC until further use. Cells were opened by sonication (Branson, Digital Sonifier 450) at 30% power for ~20 minutes total time (20 s pulse on and 40 s pulse off), cells were kept on ice during sonification. The cellular lysate was clarified by centrifugation at 10,000 x g for 30 min at 4 ºC. The supernatant was passed over a HisTrap column (Cytiva 5 mL column) pre-equilibrated with buffer A containing 20 mM imidazole (buffer B). Protein elution was achieved using a linear gradient from 10% to 100% of buffer A containing 500 mM imidazole (buffer C). Protein fractions containing 6xHis-5F-Trp,U-^{13}C,^{15}N-N^{NTD} or 6xHis-5F-Trp,U-^{15}N-N^{NTD} were pooled and TEV protease was added with the protein at a ~1 to 30 molar ration (TEV to fusion protein) to cleave the 6xHis N-terminal. (TEV plasmid was kindly provided by Sharon Rozovsky (University of Delaware); TEV was expressed and purified in house according to (*23*)). The cleavage was performed while dialyzing the sample against buffer A overnight at 4 ºC. The cleaved protein was passed through a HisTrap column (Cytiva 5 mL column) pre-equilibrated with buffer B. 5F-Trp-N^{NTD}-containing flowthrough was collected, diluted 3 times (v/v) with 20 mM HEPES buffer, pH 8, and passed over a heparin column (Cytiva 5 mL column) pre-equilibrated in 20 mM HEPES, 0 mM NaCl, pH 8 (buffer D). The protein was eluted from the heparin column using a linear gradient (0 to 100 %) of 20 mM HEPES, 1 M NaCl, pH 8 (buffer E).

Preparation of in-cell samples for solution NMR, MAS NMR, and DNP-enhanced MAS NMR

Purified proteins, 5F-Trp,U-^{15}N-N^{NTD} WT, W14F, W70F, and W94F, were delivered into A2780 mammalian cells following the electroporation protocol developed by Selenko and coworkers (*5*). Briefly, A2780 cells were seeded in RPMI-1640 medium, supplemented with 10% FBS, in a 175 cm^3 plate. After reaching 80% confluency, cells were washed with phosphate buffer saline (PBS) followed by trypsin treatment (Gibco, 0.05%), for 5 min at 37 ºC at 5% CO_2. After trypsinization, cells were pelleted by centrifugation at 150 x g for 5 min at room temperature. The supernatant was discarded, and cells were washed twice with 10 mL PBS buffer. The number of live and dead cells was determined using Trypan blue (*24*). Live cells were harvested and washed once with 1 mL of electroporation (EP) buffer without ATP (100 mM sodium phosphate, 5 mM KCl, 15 mM $MgCl_2$, 15 mM HEPES, 2 mM glutathione reductase at pH 7). After the EP wash, the cells were pelleted and mixed with the recombinant protein solution (~25-60 mg/mL, in EP buffer containing 2 mM ATP). The volume for resuspension was calculated to reach ~15 to 20 million cells per cuvette, and the cell/protein suspension was placed into the cuvette and electroporated twice using the B-028 program in a Lonza Amaxa Nucleofactor IIb instrument. After EP, the cells were immediately resuspended in 1 mL warm rich RPMI-1640 medium and transferred to a plate containing warm medium. Cells were allowed to recover for 4 to 6 hours. After recovery, cells were washed three times with PBS and then collected using trypsinization. The number of cells and their viability were accessed using Trypan blue exclusion with a Neubauer hemocytometer.

For delivering protein and AMUPol into the cells, the above protocol was used with minor changes. The electroporation buffer containing 20 or 40 mM AMUPol was prepared without glutathione reductase, to avoid radical reduction (final buffer: 100 mM sodium phosphate, 5 mM KCl, 15 mM $MgCl_2$, 15 mM HEPES, 2 mM ATP and 20 or 40 mM AMUPol, pH 7). After electroporation, 1 mL of warm rich RPMI-1640 medium was added to the cuvette and the mixture was transferred to a plate containing warm medium. For this sample a short recovery period was used, with recovery monitored every 10 min and stopped when cells were starting to attach to the plate (~30-60 minutes). After recovery, the cells were packed

into the rotor (see below). These resulting samples had final AMUPol concentration of <3 mM for $[AMUPol]_{in/out}$ and 13 mM AMUPol for $[AMUPol]_{in/out}$ in the rotor. An analogous sample was prepared as the one with 13 mM AMUPol for $[AMUPol]_{in/out}$, using 30 mM AMUPol electroporated into the cell with the protein, followed by introducing additional 30 mM AMUPol solution to the cells after recovery.

For *in-cell* MAS NMR, the cells where pelleted by centrifugation at 150 x g for 5 min at room temperature, and the pellet was resuspended in RPMI-1640 medium supplemented with 20% FBS and 10% DMSO. After resuspension, about 3 million cells were transferred into the rotor using a 200 µL pipette tip that was adjusted in size to fit into a 1.5 mL Eppendorf tube. Cells were loaded into the tip and pelleted inside the rotor by centrifugation at 500 x g for 5 min at 4 ºC. The latter step was repeated as many times as necessary to fill the rotor.

For *in-cell* DNP-enhanced MAS NMR, the cells were pelleted by centrifugation at 150 x g for 5 min at room temperature, and the pellet was resuspended in RPMI-1640 medium, supplemented with 20% FBS and 10% DMSO. After resuspension, for fast packing, about 3 million cells were transferred to a 1.5 mL Eppendorf and centrifuged at 150 x g for 5 min at room temperature. The cell pellet was resuspended in 30 µL of cold medium (RPMI-1640 medium, supplemented with 20% FBS, 10% DMSO and AMUPol), and this step was repeated twice to ensure complete buffer exchange, see **table S3 and S4** for specific buffer compositions. The AMUPol concentration varied from 20 to 40 mM. To optimize the sample preparation protocol, we tested pre-loading the rotor with 10 µL of cold buffer containing AMUPol, prior to loading 10 µL of cells. Other tests included limiting the amounts of cells to increase the final AMUPol concentration (performed only for W14F, W70F and W94F variant samples, see below). During all steps, cells and buffers were kept on ice and centrifuges were pre-cooled at 4 ºC.

The rotors were transferred to -80 ºC in a Styrofoam box, designed to freeze the cells slowly at ~-1 ºC/min. The approach was tested with cells and yielded about 90% cell viability.

The concentration of AMUPol in each 1.3 rotor was determined by EPR. The measurements were performed on an ESR5000 instrument, sweeping the magnetic field from 300 to 370 mT and recording 2000 points at room temperature (~25 ºC). Each measurement took approximately 5 minutes, ensuring minimal AMUPol reduction.

Electrophoretic mobility shift assay (ESMA)

All N^{NTD} samples were subjected to EMSA analysis to assess RNA binding of the different protein variants. We used a 32-nt RNA oligo of the 5'UTR of SARS-CoV-2 RNA (Genscript). EMSA were performed mixing a constant RNA amount (0.5 to 1 µM), and protein concentrations up to 62 µM in 20 mM phosphate buffer, 150 mM NaCl, pH 7.4. Mixtures were incubated at 37 ºC for 30 min and loaded onto a 1.6 % agarose gel containing GelGreenR Nucleic Acid Stain. The agarose gel was run at 60-70 V for approximately 1.5 hours in 0.5X TBE buffer (50 mM Tris base, 50 mM boric acid, 0.4 mM EDTA). The gel was imaged using a FluroChem Q device (Cell Biosciences). RNA binding was qualitatively assessed, based on band intensities, measured by ImageJ (*25*).

NMR spectroscopy

Solution NMR spectroscopy

Protein integrity was checked by solution NMR, for WT, W14F, W70F, and W94F 5F-Trp,U-^{15}N-N^{NTD}. ^{1}H-^{15}N HSQC spectra were recorded at 14.1 T (^{1}H Larmor frequency of 600.13 MHz) on a Bruker NEO spectrometer equipped with 5 mm QCI Bruker CryoProbe. ^{19}F 1D NMR spectra were acquired at 298 K at 14.1 T (^{1}H Larmor frequency of 600.32 MHz) on a Bruker AVIII spectrometer, outfitted with a 5 mm Bruker Prodigy CryoProbe. The sample volume was 150 µL, containing ~5 mg/mL N^{NTD} in EP buffer with 10% D_2O. The ^{19}F chemical shifts are referenced to trifluoracetic acid. Other data acquisition and processing parameters are specified in the **fig. S1, 3** legends and **table S6**.

MAS NMR spectroscopy

MAS NMR experiments were performed at 20.0 T (^{1}H Larmor frequency of 850.17 MHz) on a Bruker AVIII spectrometer equipped with a 1.9 mm HX MAS probe with the ^{1}H channel tuned to ^{19}F. MAS NMR spectra were also acquired at 11.7 T (^{1}H Larmor frequency of 500.13 MHz) on a Bruker AVIII spectrometer equipped with a 1.3 mm HFX MAS probe. For the measurements at 20.0 and 11.7 T, prior to sample insertion, the probes were pre-cooled to the lowest temperature possible using a gas flow of 1500 to 1700 L/h and temperature set to 225 K at the VT control unit, reaching temperatures ≤-20 ºC. After the desired sample temperature was reached, the gas flow was quickly reduced to 400 L/min, and the sample was inserted using a pre-chilled Bruker sample extraction/insertion tool, to ensure that the sample remained frozen during the insertion. Once the rotor was inserted, the gas flow was set to the maximum and spinning was started slowly, increasing in 5 kHz steps, until the desired MAS frequency at the desired sample temperature were reached. ^{19}F chemical shifts were referenced to mefloquine used as a secondary reference (the most shielded peak of mefloquine at 8.8 ppm). The typical 90 pulse lengths were 2.4 µs (^{19}F) and 2.0 µs (^{1}H). All spectra were processed using TopSpin 3.6 or MNova. Other data acquisition and processing parameters are specified in the legend of **Fig. 2C, 2D** and summarized in **table S7**.

DNP-enhanced MAS NMR spectroscopy

Sample insertion was performed as described above, except that both the rotor insert holder and the rotor were pre-chilled on dry ice prior to transferring the rotor to the holder. Following the pre-chilling step, the holder was immediately placed into the magnet, and the sample rotor was inserted. MAS NMR experiments with and without DNP enhancement were performed at 9.4 T (^{1}H Larmor frequency of 400.56 MHz) on a Bruker Avance NEO spectrometer, equipped with a klystron microwave source operating at 263 GHz electron frequency and <5 W power input to the probe, and a 1.3 mm HCN DNP MAS probe where the ^{1}H channel was tuned to ^{19}F. The instrument performance was checked for each run when setting up the magnet for the in-cell experiments. DNP enhancements were ascertained on a standard proline sample, which consistently yields ^{13}C DNP signal enhancements of over 200. Adamantane and mefloquine samples were used for ^{13}C and ^{19}F chemical shift reference, respectively. The typical 90 pulse lengths were 1.1 µs (^{19}F) and 1.1 µs (^{1}H). The ^{19}F and ^{1}H signal buildup curves were recorded with a pseudo 2D pulse sequence, which was modified to incorporate a three-pulse scheme before signal acquisition to remove ^{19}F background (*26*). ^{19}F chemical shifts were referenced to mefloquine as described above. The buildup curves were recorded for samples containing <3, 6.3, and 13 mM AMUPol (**table S3**) with the following delays: 0.05, 0.5, 1.0, 2.0, 3.0, 5.0, 7.5, 10.0, 12.5, 15.0, 20.0, 25.0, 30.0 and 60.0 seconds for both, the <3 and the 6.3 mM AMUPol-containing samples and 1.0, 2.0, 3.0, 5.0, 7.5, 10.0, 12.5, 15.0, 20.0, 25.0, 30.0, 40.0, 50.0, 60.0, 70.0, 80.0, 90.0, 105.0, 120.0, 150.0 and 180.0 seconds for the 13 mM AMUPol-containing sample. The ^{19}F DANTE pulse length was 0.1 µs. The DANTE interpulse delay was set to one rotor cycle, 25 µs. A total of 22 DANTE pulses were used for the selective inversion. The ^{1}H buildup curve was recorded for the sample containing 13 mM AMUPol

(**table S9**) using the following delays: 0.01, 0.05, 0.10, 0.25, 0.50, 0.75, 1.00, 2.00, 3.00, 5.00, 10.00 and 15.00 seconds. Other data acquisition and processing parameters are summarized in **table S8 and S9** and **fig. S8** legend.

References

1. H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne, The Protein Data Bank. *Nucleic Acids Res.* **28**, 235–242 (2000).

2. RCSB Protein Data Bank, https://www.rcsb.org.

3. A. G. Palmer, NMR Characterization of the dynamics of biomacromolecules. *Chem. Rev.* **104**, 3623–3640 (2004).

4. T. R. Alderson, L. E. Kay, NMR spectroscopy captures the essential role of dynamics in regulating biomolecular function. *Cell* **184**, 577–595 (2021).

5. F.-X. Theillet, A. Binolfi, B. Bekei, A. Martorana, H. M. Rose, M. Stuiver, S. Verzini, D. Lorenz, M. Van Rossum, D. Goldfarb, P. Selenko, Structural disorder of monomeric α-synuclein persists in mammalian cells. *Nature* **530**, 45–50 (2016).

6. W. Zhu, A. J. Guseman, F. Bhinderwala, M. Lu, X. Su, A. M. Gronenborn, Visualizing proteins in mammalian cells by ^{19}F NMR spectroscopy. *Angew. Chem. Int. Ed.* **61**, e202201097 (2022).[*Angewandate Chemie International Edition*]

7. C. Li, G.-F. Wang, Y. Wang, R. Creager-Allen, E. A. Lutz, H. Scronce, K. M. Slade, R. A. S. Ruf, R. A. Mehl, G. J. Pielak, Protein ^{19}F NMR in *Escherichia coli*. *J. Am. Chem. Soc.* **132**, 321–327 (2010).

8. S. S. Stadmiller, A. H. Gorensek-Benitez, A. J. Guseman, G. J. Pielak, Osmotic shock induced protein destabilization in living cells and its reversal by glycine betaine. *J. Mol. Biol.* **429**, 1155–1161 (2017).

9. A. M. Gronenborn, Small, but powerful and attractive: ^{19}F in biomolecular NMR. *Structure* **30**, 6–14 (2022).

10. R. Ghosh, Y. Xiao, J. Kragelj, K. K. Frederick, In-cell sensitivity-enhanced NMR of intact viable mammalian cells. *J. Am. Chem. Soc.* **143**, 18454–18466 (2021).

11. S. A. Overall, A. B. Barnes, Biomolecular perturbations in in-cell dynamic nuclear polarization experiments. *Front. Mol. Biosci.* **8,** 743829 (2021).

12. K. M. McCoy, R. Rogawski, O. Stovicek, A. E. McDermott, Stability of nitroxide biradical TOTAPOL in biological samples. *J. Magn. Reson.* **303**, 115–120 (2019).

13. R. Ghosh, R. Dumarieh, Y. Xiao, K. K. Frederick, Stability of the nitroxide biradical AMUPol in intact and lysed mammalian cells. *J. Magn. Reson.* **336**, 107150 (2022).

14. B. E. Ackermann, B. J. Lim, N. Elathram, S. Narayanan, G. T. Debelouchina, A comparative study of nitroxide-based biradicals for dynamic nuclear polarization in cellular environments. *ChemBioChem* **23**, e202200577 (2022). [*ChemBioChem]*

15. M. Lu, M. Wang, I. V. Sergeyev, C. M. Quinn, J. Struppe, M. Rosay, W. Maas, A. M. Gronenborn, T. Polenova, ^{19}F Dynamic nuclear polarization at fast magic angle spinning for NMR of HIV-1 capsid protein assemblies. *J. Am. Chem. Soc.* **141**, 5681–5691 (2019).

16. M. Fritz, J. Kraus, C. M. Quinn, G. P. A. Yap, J. Struppe, I. V. Sergeyev, A. M. Gronenborn, T. Polenova, Measurement of accurate interfluorine distances in crystalline organic solids: A high-frequency magic angle spinning NMR approach. *J. Phys. Chem. B* **123**, 10680–10690 (2019).

17. G. Porat-Dahlerbruch, J. Struppe, C. M. Quinn, A. M. Gronenborn, T. Polenova, ^{19}F fast MAS (60–111 kHz) dipolar and scalar based correlation spectroscopy of organic molecules and pharmaceutical formulations. *Solid State Nucl. Magn. Reson.* **122**, 101831 (2022).

18. M. Lu, S. Sarkar, M. Wang, J. Kraus, M. Fritz, C. M. Quinn, S. Bai, S. T. Holmes, C. Dybowski, G. P. A. Yap, J. Struppe, I. V. Sergeyev, W. Maas, A. M. Gronenborn, T. Polenova, ^{19}F magic angle spinning NMR spectroscopy and density functional theory calculations of fluorosubstituted tryptophans: Integrating experiment and theory for accurate determination of chemical shift tensors. *J. Phys. Chem. B* **122**, 6148–6155 (2018).

19. G. Bodenhausen, R. Freeman, G. A. Morris, A simple pulse sequence for selective excitation in Fourier transform NMR. *J. Magn. Reson.* **23**, 171–175 (1976).

20. A. E. Bennett, R. G. Griffin, J. H. Ok, S. Vega, Chemical shift correlation spectroscopy in rotating solids: Radio frequency-driven dipolar recoupling and longitudinal exchange. *J. Chem. Phys.* **96**, 8624–8627 (1992).

21. P. B. Crowley, C. Kyne, W. B. Monteith, Simple and inexpensive incorporation of ^{19}F-Tryptophan for protein NMR spectroscopy. *Chem. Commun.* **48**, 10681 (2012).

22. S. Sarkar, B. Runge, R. W. Russell, K. T. Movellan, D. Calero, S. Zeinalilathori, C. M. Quinn, M. Lu, G. Calero, A. M. Gronenborn, T. Polenova, Atomic-resolution structure of SARS-CoV-2 nucleocapsid protein N-terminal domain. *J. Am. Chem. Soc.* **144**, 10543–10555 (2022).

23. P. G. Blommel, B. G. Fox, A combined approach to improving large-scale production of tobacco etch virus protease. *Protein Expr. Purif.* **55**, 53–68 (2007).

24. W. Strober, Trypan Blue Exclusion Test of Cell Viability. *CP in Immunology* **21,** A3.B.1–A3.B.3. (1997).

25. C. A. Schneider, W. S. Rasband, K. W. Eliceiri, NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**, 671–675 (2012).

26. F. Wang, S. K. Ramakrishna, P. Sun, R. Fu, Triple-pulse excitation: An efficient way for suppressing background signals and eliminating radio-frequency acoustic ringing in direct polarization NMR experiments. *J. Magn. Reson.* **332**, 107067 (2021).

27. E. E. Najbauer, L. B. Andreas, Correcting for magnetic field drift in magic-angle spinning NMR datasets. *J. Magn. Reson.* **305**, 1–4 (2019).

Acknowledgments:
Special thanks to Daniel Eduardo Gutierrez Ortegon and Valentina Gomez Osorio for their assistance with protein expression and purification and mammalian cell culture. We also thank Caitlin M. Quinn and Roman Zadorozhnyi for helping with the sample insertion for room temperature experiments.

Funding:
This work was supported by the National Science Foundation (NSF Grant CHE-1708773 to A. M. G. and T. P.), by the National Institutes of Health (NIH Grant 1U54AI170791 to A. M. G., T. P., NMR Core)

Author contributions:

Conceptualization: TP and AMG
Methodology: TP, AMG, KTM
Resources: TP, DB, KTM, BR (protein expression)
Investigation: KTM, TP, DB, WZ
Data interpretation: KTM and TP
Supervision: TP, AMG, JK
Writing—original draft: TP and KTM
Writing—review & editing: TP, AMG, KTM, DB, JK
Project administration: TP, AMG, JK
Funding acquisition: TP and AMG
Visualization: TP, AMG, KTM

Competing interests: The authors declare no competing interests.

Data and materials availability: All data discussed in this paper are available in the main text or the supplementary materials."

Supplementary Materials

The Supplementary Materials contain information about the chemicals used and additional figures (figs. S1 to S9) showing ^{19}F *in-cell* DNP spectra with microwave ON and OFF (fig. S1A), ^{19}F solution NMR spectra from 5F-Trp,U-^{15}N-N^{NTD} (fig. S1B), ^{13}C *in-cell* DNP spectra with microwave ON and OFF (fig. S6), mass spectrometry data for the N^{NTD} mutants (W14F, W70F and W94F, fig. S2), 2D ^{1}H-^{15}N HSQC from all the proteins used in this study (fig. S3), electrophoretic mobility shift assays for all the protein constructs used (WT, W14F, W70F and W94F, fig. S4) and the DNP build up times of ^{1}H and ^{19}F (figs. S7 and S8) as well as protein quantification data using sodium dodecyl sulfate-polyacrylamide gel electrophoresis (fig. S5), and 2D spin diffusion ^{19}F spectrum using 1 second mixing time (fig. S9). In addition, the tables in the Supplementary Materials provide information on protein quantification, ^{19}F chemical shifts, the NMR acquisition parameters and the interfluorine distances for 5F-Trp-N^{NTD}.

Figure captions

Fig. 1. Domain delineation, amino acid sequence and ribbon diagram structure of SARS-CoV-2 N^{NTD}. (A) Top: domain organization of SARS-CoV-2 nucleocapsid (N) protein; N-terminal domain (N^{NTD}), C-terminal domain (N^{CTD}). Bottom: N^{NTD} amino acid sequence with Trp residues shown in magenta. Residue numbering 2-136 in the current N^{NTD} construct corresponds to 40-174 in the full-length N protein.

(B) Ribbon representation of the lowest-energy conformer of the 10-conformer MAS NMR structure ensemble (PDB: 7SD4) of N^{NTD}. W14, W70, and W94 side chains are in stick representation. The fluorine atoms at the 5 positions of the indole rings are shown by magenta spheres. Interfluorine distances are indicated by dashed lines.

Fig. 2. Sensitivity of *in-cell* ^{19}F DNP and MAS NMR spectra of SARS-CoV-2 5F-Trp-N^{NTD}. (A) *In-cell* DNP-enhanced spectrum of a sample containing ~1.2-1.5 million cells, containing a total of ~0.3 nanomoles of protein and 13 mM AMUPol. The spectrum was recorded with 64 scans and a recycle delay of 15 s in 16 min; the MAS frequency was 40 kHz. **(B)** DNP-enhanced (magenta trace) and control microwave-off (black trace) spectra of a sample containing ~1.2-1.5 million cells containing a total of ~0.8 nanomoles of protein and 6.3 mM AMUPol. The spectra were recorded with 1024 scans and a recycle delay of 5 s in 85 min; the MAS frequency was 30 kHz. **(C)** MAS NMR spectrum of a sample containing ~4-5 million cells, with a total of ~1.9 nanomoles of protein. The spectrum was acquired at 20.0 T (850 MHz ^{1}H Larmor frequency) and MAS frequency of 20 kHz. **(D)** MAS NMR spectrum of a sample containing ~1.2-1.5 million cells containing a total of ~1 nanomole of protein. The spectrum was acquired at 11.7 T (500 MHz ^{1}H Larmor frequency) and MAS frequency of 30 kHz. **(E)** DNP signal buildup time constants, T_b, for *in-cell* 5F-Trp-N^{NTD} samples, plotted against AMUPol concentrations, for ^{19}F (black) and ^{1}H (red). **(F)** Sensitivity (normalized I/I_{max}) of ^{19}F *in-cell* DNP-enhanced signals as a function of MAS frequency for a sample containing 6.3 mM AMUPol. For comparison, the dependence of DNP signal enhancement on MAS frequency is shown for ^{13}C signals of the carbonyl groups detected in DNP-enhanced ^{13}C CPMAS spectrum of the same sample. All DNP data were recorded at 100 K and 9.4 T, with a microwave power of <5 W. The NMR acquisition parameters are detailed in Supplementary **Tables S8** and **S9**.

Fig. 3. Resonance assignments and interfluorine correlations in *in-cell* ^{19}F DNP-enhanced MAS spectra of SARS-CoV-2 5F-Trp-N^{NTD}. (A, B) ^{19}F *in-cell* DNP-enhanced MAS NMR and solution NMR spectra of 5F-Trp-N^{NTD} WT and W94F-N^{NTD}, W70F-N^{NTD}, and W14F-N^{NTD} variants. **(C)** ^{19}F *in-cell* DNP-enhanced MAS NMR spectra of 5F-Trp-N^{NTD} recorded with selective DANTE magnetization inversion pulses, followed by non-selective excitation and signal detection. The frequencies of DANTE inversion pulses are indicated with colored arrows, and the spectra are colored accordingly. **(D)** ^{19}F-^{19}F *in-cell* DNP-enhanced spin diffusion spectra of 5F-Trp-N^{NTD} recorded with a mixing time of 2 s (left panel) and with no mixing (control, right panel). The cross peaks between fluorine signal of 5F-Trp-70 and 5F-Trp-94 are labeled. The corresponding 1D traces are shown on the right of each spectrum. Each spectrum was recorded with 256 scans, a recycle delay of 5 seconds, and a total experimental time of 32 and 28 hours for the spectra acquired with and without mixing, respectively. The MAS frequency was 40 kHz.

562 **Table 1.** Summary of MAS frequency dependence and buildup times of ^{19}F and ^{1}H signals in DNP-enhanced MAS NMR experiments on 5F-Trp,U-$^{15}N(^{13}C)$-N^{NTD} delivered in
563 human A2780 cells under different experimental conditions.

564

Sample			5F-Trp,U-^{13}C,^{15}N-N^{NTD}	5F-Trp,U-^{13}C,^{15}N-N^{NTD}	5F-Trp,U-^{15}N-N^{NTD}	5F-Trp,U-^{13}C,^{15}N-N^{NTD}
cells/rotor (x10^6)			~1.2-1.5	~1.2-1.5	~1.2-1.5	~1.1
Amount of N^{NTD} in the NMR sample (nmoles)			0.61	0.83	0.29	0.15
[AMUPol] by EPR (mM)			<3	6.3	13	est. 13**
Buffer			10% DMSO, 10% FBS, 80% RPMI	10% DMSO, 10% FBS, 80% RPMI	10% DMSO-d6, 20% D_2O, 20% FBS, 50% RPMI	10% DMSO-d6, 20% D_2O, 20% FBS, 50% RPMI
Cell viability (%)			>90	>90	>80	>90
^{19}F T_b (s)			14.6	13.3	16.2	
^{1}H T_b (s)					1.7	
^{19}F Sensitivity*	ω_r (kHz)	20	0.32	0.23	1.52	1.22
	ω_r (kHz)	30	0.31	0.23	1.42	1.26
	ω_r (kHz)	40	0.35	0.32	3.17	3.33

565

566 *Sensitivity is the SNR per nanomole of protein in the rotor per square root of the number of scans, measured with recycle delays corresponding to ~0.3*T_b.

567 **The AMUPol concentration was estimated based on the protocol used and the ^{19}F DNP enhancements.

568

Table 2. Sensitivity of ^{19}F MAS NMR and DNP-enhanced MAS NMR experiments on 5F-Trp,U-^{15}N(^{13}C)-N^{NTD} delivered in human A2780 cells.

Sample [AMUPol] (mM)	NTD in rotor (nmol)	NTD in rotor (µg)	Rotor size (mm)	Field strength (T)	Temperature (K)	Recycle delay (s)	MAS frequency (kHz)	SNR_{norm}[1]
5F-Trp,U-^{15}N-N^{NTD}	1.86	27.6	1.9	19.97	271[3]	5	20	0.0004[2]
5F-Trp,U-^{15}N-N^{NTD}	1.02	15.2	1.3	11.7	266[3]	5	30	0.002
5F-Trp,U-^{15}N-N^{NTD}	1.02	15.2	1.3	11.7	273[3]	5	40	0.002
5F-Trp,U-^{13}C,^{15}N-N^{NTD} (<3)	0.61	9.0	1.3	9.4	100	5	20	0.010
							30	0.009
							40	0.011
5F-Trp,U-^{13}C,^{15}N-N^{NTD} (6.3)	0.83	12.3	1.3	9.4	100	5	20	0.007
							30	0.007
							40	0.010
5F-Trp,U ^{15}N-N^{NTD} (13)	0.29	4.4	1.3	9.4	100	5	20	0.046
							30	0.043
							40	0.095
5F-Trp,U-^{13}C,^{15}N-N^{NTD} (est. 13[3])	0.15	2.3	1.3	9.4	100	5	20	0.037
							30	0.038
							40	0.100
5F-Trp,U ^{15}N-N^{NTD} W14F	0.58	8.7	1.3	9.4	100	5	40	0.011
5F-Trp,U ^{15}N-N^{NTD} W70F	0.44	6.5	1.3	9.4	100	5	40	0.008
5F-Trp,U-^{15}N-N^{NTD} W94F	0.49	7.3	1.3	9.4	100	5	40	0.016
5F-Trp,U ^{15}N-N^{NTD} W70F	0.28	4.1	1.3	9.4	100	5	40	0.008

[1]SNR_{norm} - the signal-to-noise ratio estimated from the most intense peak, per nanomole of protein in the rotor per square root of experimental time.

[2]The low SNR_{norm} is associated with the 1.9 mm HX probe used in the measurements.

[3]The AMUPol concentration was estimated based on the protocol used and the ^{19}F DNP enhancements.