Integrating ^{19}F Distance Restraints for Accurate Protein Structure Determination by Magic Angle Spinning NMR Spectroscopy

Brent R. Runge,[1,2,#] Roman Zadorozhnyi,[1,2,#] Caitlin M. Quinn,[1] Ryan W. Russell,[1,2] Manman Lu,[2,3] Santiago Antolínez,[1] Jochem Struppe,[4] Charles D. Schwieters,[5] In-Ja L. Byeon,[2,3] Jodi A. Hadden-Perilla,[1] Angela M. Gronenborn,[2,3*] Tatyana Polenova[1,2*]

[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]Office of Intramural Research, Center for Information Technology, National Institutes of Health, Building 12A, Bethesda, MD 20892, United States.

***Corresponding authors:** Tatyana Polenova, Department of Chemistry and Biochemistry, University of Delaware, Newark, DE, USA, Tel.: (302) 831-1968; Email: tpolenov@udel.edu; Angela M. Gronenborn, Department of Structural Biology, University of Pittsburgh School of Medicine, 3501 Fifth Ave., Pittsburgh, PA 15260, USA, Tel.: (412) 648-9959; Email: amg100@pitt.edu.

[#]These two authors contributed equally to this work

Keywords: ^{19}F NMR, *Oscillatoria agardhii* agglutinin (OAA), magic angle spinning (MAS) NMR, atomic-resolution structure

ABSTRACT

Traditional protein structure determination by magic angle spinning (MAS) solid-state NMR spectroscopy primarily relies on interatomic distances up to 8 Å, extracted from ^{13}C-, ^{15}N-, and ^{1}H-based dipolar-based correlation experiments. Here, we show that ^{19}F fast (60 kHz) MAS NMR spectroscopy can supply additional, longer distances. Using 4F-Trp,U-^{13}C,^{15}N crystalline *Oscillatoria agardhii* agglutinin (OAA), we demonstrate that judiciously designed 2D and 3D ^{19}F-based dipolar correlation experiments such as (H)CF, (H)CHF, and FF, can yield interatomic distances in the 8-16 Å range. Incorporation of fluorine-based restraints into structure calculation improved the precision of Trp side chain conformations as well as regions in the protein around the fluorine containing residues, with notable improvements observed for residues in proximity to the Trp pairs (W10/W17 and W77/W84) in the carbohydrate-binding loops, which lacked sufficient long-range ^{13}C-^{13}C distance restraints. Our work highlights the use of fluorine and ^{19}F fast MAS NMR spectroscopy as a powerful structural biology tool.

INTRODUCTION

Protein structure determination by magic angle spinning (MAS) NMR spectroscopy relies primarily on ^{13}C-, ^{15}N-, and ^{1}H-based distance restraints. While this approach has proven effective for solving structures of compact single-chain proteins,[1] the local nature of distance restraints derived from experiments involving these nuclei, with the upper limit being 6-8 Å, poses challenges in structural studies of large complex systems. For instance, when proteins are composed of multiple domains, possess quaternary structure or form supramolecular assemblies, the relative subunit orientations and intermolecular interfaces need to be defined.[2,3] This information can be obtained by integrating data from MAS NMR and other methods, such as cryo electron microscopy (cryo-EM),[4-6] small-angle X-ray scattering (SAXS),[7,8] X-ray crystallography,[9,10] electron paramagnetic resonance (EPR),[11] and computation,[12,13] or a combination thereof.[7] In addition, it would be desirable to also have access to distances greater than 8 Å, which, in principle, can be experimentally obtained via ^{1}H-^{1}H dipolar correlations, given the high proton gyromagnetic ratio.[14] However, dipolar truncation and relayed transfer phenomena present in molecules with dense ^{1}H networks make the measurement of long-range ^{1}H-based distances difficult,[15] except in extensively deuterated systems. Therefore, alternative NMR reporter nuclei possessing suitable magnetic properties that can be exploited for extracting longer interatomic distances are desirable.

Such a nucleus is the 100% abundant spin -½ ^{19}F isotope. Due to its high gyromagnetic ratio, its sensitivity is 83% of ^{1}H, the chemical shifts cover a range larger than 300 ppm, and, most importantly, fluorine is virtually absent from most biological systems.[16] Furthermore, fluorine can readily be incorporated into proteins, without inducing major structural perturbations.[17-19] Given these favorable properties of fluorine, ^{19}F MAS NMR has emerged as a versatile and unique method for studying proteins, providing interatomic distances of up to 23 Å and 15 Å from ^{19}F-^{19}F and ^{19}F-^{1}H correlation experiments, respectively.[20,21] Notably, high MAS frequencies of 40-111 kHz have proven advantageous for applications to a broad range of systems investigated to date, ranging from small-molecule pharmaceuticals to biological assemblies.[21-26] These fast MAS conditions offer large gains in resolution and sensitivity, permitting in-depth characterization of nanomole quantities of sample through 2D and 3D fluorine-based homonuclear and heteronuclear (^{19}F-^{13}C and ^{19}F-^{1}H) correlations.[20,21,27-30] In addition, large, over 100-fold signal enhancements are observed in ^{19}F dynamic nuclear polarization (DNP) MAS NMR experiments[31,32] and ^{19}F-^{13}C correlations corresponding to distances of up to 12 Å could be readily observed, as reported for HIV-1 CA capsid protein assemblies.[31]

While the above studies demonstrated the suitability of ^{19}F fast MAS NMR for atomic-level characterization of biological systems, fluorine-based distance restraints have not been implemented in protein structure calculation protocols to date. Here, we report on the incorporation of ^{19}F-^{19}F and ^{13}C-^{19}F distance restraints obtained from MAS NMR experiments into protein structure calculation, using as a model system the 132-residue protein, *Oscillatoria agardhii* agglutinin (OAA) that has been extensively structurally characterized by X-ray crystallography, solution and MAS NMR spectroscopy.[33-36] OAA exhibits a unique β-barrel structure (Figure 1a) and two carbohydrate-binding sites.[34] The amino acid sequence comprises six Trp residues (Figure 1a), four close to each carbohydrate binding site, W10/W17 (loop1) and W77/W84 (loop 2), and two in the β-barrel core (W23/W90). We incorporated fluorine atoms at the 4-, 5-, or 7- positions of each Trp and structurally characterized 4F-Trp,U-^{13}C,^{15}N-labeled OAA by MAS NMR. 123 fluorine-based correlations were observed in dipolar-based 2D (H)CF HETCOR and 3D (H)CHF HETCOR spectra, and were converted into F-C distances of up to 16.7 Å. The inclusion of these fluorine-based distances resulted in an overall notable improvement in the accuracy and precision of the ensemble, specifically for Trp residues and for residues in proximity to the Trp pairs (W10/W17 and W77/W84) in the carbohydrate-binding loops, which lacked sufficient long-range ^{13}C-^{13}C distance restraints. Our results establish the proof of concept for ^{19}F fast MAS NMR based protein structure determination. The methodology introduced herein is envisioned to be particularly beneficial to large proteins and protein assemblies.

RESULTS AND DISCUSSION

^{19}F NMR chemical shifts for 4F-Trp, 5F-Trp, and 7F-Trp labeled OAA

Incorporation of 4F-Trp, 5F-Trp, and 7F-Trp into OAA was assessed by mass spectrometry (Figure S1, Supporting Information) and solution ^{19}F NMR spectra as well as ^{1}H-^{15}N HSQC spectra of 4F-Trp,U-^{13}C,^{15}N-OAA, 5F-Trp,U-^{13}C,^{15}N-OAA, and 7F-Trp,U-^{13}C,^{15}N-OAA. The ^{19}F NMR spectra are well dispersed, exhibiting six resolved fluorine resonances for 7F-Trp,U-^{13}C,^{15}N-OAA, five for 4F-Trp,U-^{13}C,^{15}N-OAA with double intensity for one of these, and four for 5F-Trp,U-^{13}C,^{15}N-OAA with two resonances exhibiting double intensity (Figure S2a, Supporting Information). The ^{1}H-^{15}N HSQC spectra of the three fluorotryptophan-containing OAA samples exhibit noticeable chemical shift changes compared to the non-fluorinated protein and in all of them the indole N$^{\varepsilon}$ resonances are missing (Figures S2b-d, Supporting Information), as expected, since ^{14}N fluoroindoles were used for labeling the proteins. ^{1}H-^{15}N amide backbone assignments were transferred from the WT OAA spectrum.[37] Overall, all NMR spectra are of excellent quality

indicating that fluorotryptophan incorporation at all six sites did not negatively impact the structure of these proteins.

X-ray structure of 4F-Trp OAA

4F-Trp OAA crystalized in the P1 space group with two monomers in the asymmetric unit and two bound CAPS molecules per monomer, one in each sugar binding site, similar to WT OAA, whose X-ray crystal structure was determined by us previously.[33] X-ray diffraction data for 4F-Trp OAA crystals were collected to compare the structures of WT OAA and the fluorinated analog. The structure of a single chain of 4F-Trp OAA is shown in Figure 1a. Complete statistics for X-ray data collection, phasing, and refinement are provided in Table 1. The atomic RMSD between chain A and chain B in the asymmetric unit is 0.1 Å for backbone and 0.3 Å for all heavy atoms. Incorporation of 4F-Trp had no discernible effect on the structure, with backbone and heavy atom RMSDs of 0.2 Å and 0.5 Å, respectively, between WT and 4F-Trp OAA for both molecules in the asymmetric unit cell (Figure S3, Supporting Information). As shown in Figure 1b, the Trp side chain conformations are also not altered from the WT conformations and the presence of the fluorine atom at the 4 position is readily discerned from the increased electron density. The only clear differences in the X-ray structure of 4F-Trp OAA are seen for the side chains of S14 and D80 located in the more flexible carbohydrate binding loops.

Chemical shift assignments of microcrystalline 4F-Trp OAA

We devised and employed a protein batch crystallization protocol and produced samples of homogenous microcrystalline WT U-^{13}C,^{15}N-OAA and 4F-Trp,U-^{13}C,^{15}N-OAA for MAS NMR. Excellent MAS NMR spectra with ^{13}C line widths as narrow as 20 Hz (Figure S4, Supporting Information) were obtained. Interestingly, the spectrum of the 4F-Trp,U-^{13}C,^{15}N-OAA appears to exhibit better resolution, compared to that of WT U-^{13}C,^{15}N-OAA (Figure 1c). Almost complete ^{13}C, ^{15}N, and ^{19}F backbone and side chain chemical shift assignments were obtained from a total of five 2D and three 3D ^{19}F- and ^{13}C- detected MAS NMR experiments and a representative sequential backbone walk for residues T50-I59 is shown in Figure 1d. ^{13}C and ^{15}N chemical shifts were determined for all 132 amino acids, with complete shift assignments accomplished for 94 residues. This is a notable improvement over WT OAA, where ^{13}C and ^{15}N resonance assignments for a large number of residues remained ambiguous due to resonance overlap associated with the sequence repeats.[36] In general, chemical shift changes of ~1 ppm are observed for ^{13}C and ^{15}N resonances associated with amino acids in close proximity to ^{19}F (Figure 1c,d), alleviating chemical shift degeneracy and, in turn, permitting unambiguous assignment of most ^{13}C and ^{15}N resonances. Examples of the improved dispersion in the ^{13}C-^{13}C combined R-

symmetry spin diffusion (CORD)[38] spectra for several overlapping resonances of WT U-^{13}C,^{15}N-OAA (gray) compared to two resolved resonances of 4F-Trp,U-^{13}C,^{15}N-OAA (purple) are provided in Figure 1c; the full overlay of the aliphatic region of the CORD spectrum is shown in Figure S5, Supporting Information. Notably, all spin systems (132) have been observed and sequentially assigned in the 2D CORD spectra with 25-50 ms mixing times. Of these, 28 spin systems associated with pairs of residues in the sequence repeats were assigned with 2-fold ambiguity. Despite the improvement in resolution upon 4F-Trp incorporation, the stretches of residues V33-T46 and V100-Y113 could not be unambiguously assigned and, therefore, the respective chemical shifts were handled with 2-fold ambiguity (Figure S6, Supporting Information). Proton chemical shifts for selected residues were assigned via a combination of 2D (H)NH and (H)CH MAS NMR spectra at 60 kHz using the published solution chemical shifts as a guide (Figure S7, Supporting Information).[37]

The fluorine chemical shifts were assigned by a combination of heuristic knowledge taken from the structure and ^{19}F-detected 2D (H)CF and 3D (H)CHF experiments. As shown in Figure 2, numerous correlations were observed between ^{19}F signals and ^{13}C and ^{1}H backbone and side chain resonances, allowing for their unambiguous assignments. The 2D ^{19}F-^{19}F radio frequency driven recoupling (RFDR) spectrum recorded with the RFDR mixing time of 10 ms at 50 kHz MAS frequency is shown in Figure 1e, with the 1D spectrum displayed on the top. Similar to the solution ^{19}F NMR spectrum, five distinct resonances are present, with the most shielded peak clearly comprising two partially resolved resonances. The three Trp pairs could be readily distinguished in the spectrum based on the three sets of strong cross peaks observed between the Trp residues belonging to each pair and the much weaker (if present at all) cross peaks between Trp residues of different pairs. The resonances at -42.3 ppm and -42.9 ppm were unequivocally assigned to the core pair W23 and W90 on the basis of the high-intensity cross peaks consistent with the corresponding interfluorine distance being the shortest amongst all pairs of Trp residues, 5.1 Å. By comparison, the other two pairs of cross peaks are approximately half as intense, as expected, given that the interfluorine distance between the Trp residues within each carbohydrate loop pair is considerably longer, 6.8 Å (Figure S8, Supporting Information). Notably, cross peaks corresponding to longer distances are also present in the 2D RFDR spectrum shown in Figure 1e: an intra-chain correlation (W17–W23,12.6 Å, Table S1, Supporting Information) as well as inter-chain correlations (W17–W84 and W10–W77, 16.9 and 22.9 Å, respectively, Table S1, Supporting Information). This result is consistent with our prior observations of ^{19}F-^{19}F interatomic distances of up to 23 Å in 2D RFDR spectra.[21] Finally, the ^{19}F chemical shifts for each Trp residue were assigned on the basis of ^{19}F-detected 2D (H)CF and 3D (H)CHF experiments acquired at a

MAS frequency of 60 kHz. As shown in Figure 2, numerous correlations were observed between ^{19}F signals with ^{13}C and ^{1}H backbone and side chain resonances, allowing for their unambiguous assignments. These 2D and 3D spectra were also used to determine ^{13}C-^{19}F interatomic distance restraints as discussed below.

Long-range distance restraints

^{13}C-based distance restraints for microcrystalline 4F-Trp,U-^{13}C,^{15}N-OAA were determined using 2D CORD spectra acquired with mixing times of 50, 100, 250 and 500 ms. Given the high spectral resolution, a significant number of well-resolved cross peaks were uniquely assigned to inter-residue correlations (Figure S9, Supporting Information), and the complete list of experimental ^{13}C-^{13}C distance restraints, comprising 976 medium range ($1<|i-j|<5$) and 2632 ($|i-j|\geq5$) long-range distance restraints, including ambiguous restraints up to 4-fold is summarized in Table 2. Using a fully protonated 4F-Trp,U-^{13}C,^{15}N-OAA sample, we experimentally determined an average of 43.2 ^{13}C-^{13}C restraints per residue (56.7% completeness), with approximately 7.4 medium-range restraints per residue (61.6% completeness) and 19.8 long-restraints per residue (43.2% completeness), including ≤4-fold ambiguous restraints. The total number of restraints per residue (including ^{19}F restraints discussed below) is plotted versus residue number in Figure S9, Supporting Information.

Traditionally, protein structure determination by ^{13}C, ^{15}N, and ^{1}H MAS NMR involves correlation-derived distance restraints, with the largest distances typically not exceeding 8 Å. We posited that fluorine-based distances would augment the number of long-range distances. We therefore recorded heteronuclear ^{19}F-detected 2D and 3D spectra at a MAS frequency of 60 kHz, and ^{13}C-^{19}F distance restraints were extracted from a 2D (H)CF HETCOR data set (Figure 2a). Overall, a total of 118 ^{13}C-^{19}F correlations were assigned, including 108 unambiguous assignments and the rest of 2-fold ambiguity. About half of the observed correlations in the 2D (H)CF HETCOR spectrum correspond to internuclear pairs separated by less than 5 Å, such as intra-residue 4F-Trp C^{β}, C^{α}, and C' correlations. Interestingly, in the 2D (H)CF HETCOR spectrum we observed 37 cross peaks corresponding to internuclear ^{13}C-^{19}F distances of $5\ \text{Å} \leq r \leq 8\ \text{Å}$, as well as 23 correlations corresponding to $r \geq 8$ Å. Remarkably, three distinct correlations were present that correspond to much longer distances: C^{γ}(L66)-F(W84) (14.7 Å), $C^{\gamma2}$(I25)-F(W84) (15.0 Å), and L92 C^{γ}/L132 C^{γ}-W17 4F (16.3/16.7 Å).

In addition, we extracted ^{13}C-^{19}F distance restraints from a 3D (H)CHF HETCOR spectrum (Figure 2b). To the best of our knowledge, this is the first report of a ^{19}F-detected 3D (H)CHF HETCOR experiment. The magnetization transfer pathway from ^{13}C to ^{19}F through an additional ^{1}H dimension employed in this experiment, was indispensable for extracting unique ^{13}C-^{19}F

correlations. Remarkably, even with limited ^{1}H resolution, as expected for spectra of a fully protonated protein at 60 kHz MAS frequency, unambiguous assignments were obtained for 28 ^{13}C-^{19}F correlations (plus two 2-fold ambiguous cross peaks), corresponding to interatomic distances of up to 9 Å. Moreover, 11 nonredundant cross peaks each, corresponding to interatomic distances 5 Å ≤ r ≤ 8 Å and r < 5 Å were identified, together with F(W17)-C$^{\beta}$(P16)-H$^{\zeta 3}$ (W10) and F(W17)-C$^{\beta}$(A15)-H(A15) correlations in the 3D spectrum corresponding to cross peaks arising from F and C atoms separated by 8 Å and 9 Å, respectively. Additionally, 5 cross peaks in the 3D (H)CHF HETCOR spectrum gave rise to unique distance restraints inaccessible in other spectra. Interestingly, inter-residue C-H-F correlations corresponding to C-F distances between 3.4-7.5 Å are observed at several ^{19}F frequencies. These correlations are primarily between backbone C$^{\alpha}$ or side chain C$^{\beta}$/C$^{\gamma 2}$/C$^{\delta 1}$ and aromatic 4F-Trp side chain protons (H$^{\varepsilon 1}$, H$^{\eta 2}$, and H$^{\zeta 3}$) to ^{19}F of the 4F-Trp.

Overall, homonuclear (^{19}F-^{19}F) and heteronuclear (^{13}C-^{19}F, ^{1}H-^{19}F) ^{19}F-detected experiments together yielded a total of 126 distance restraints with no more than 2-fold ambiguity, with 95 unambiguously assigned distances. The corresponding fluorine-based interatomic distances range from 2.9 to 16.7 Å, comprising both local restraints around the fluorine atoms of the 4F-Trp side chains and additional invaluable long-range inter-residue restraints with distances exceeding 8 Å. A summary of all ^{19}F-based restraints is provided in Table 2.

Structure calculation

To include ^{19}F-based distance restraints in protein structure calculation protocols, force field parameters for 4F-Trp were incorporated into XPLOR-NIH.[39] To this end, the recently developed AMBER ff15iqp protein force field parameters of fluorinated aromatic amino acids were converted into a series of pseudo potentials.[40,41] The structure of a single 4F-Trp,U-^{13}C,^{15}N-OAA chain was calculated using all available distance and dihedral angle restraints (Table 2). A superposition of the 10 lowest energy structures is provided in Figure 3a (colored purple). The average pairwise backbone and heavy atom RMSD between individual members of this ensemble and the 4F-Trp,U-^{13}C,^{15}N-OAA X-ray structure are 1.1±0.1 Å and 1.4±0.1 Å, respectively (PDB: 9AUQ and Figure 3a, colored blue). The average precision of the ensemble as measured by pairwise RMSD, is 0.4±0.1 Å and 0.9±0.1 Å for backbone and heavy atoms, respectively.

To evaluate the contribution of adding fluorine-based restraints to the calculated 4F-Trp,U-^{13}C,^{15}N-OAA structure, parallel calculations were performed without these restraints, and the resulting 10 lowest energy structures are shown in Figure 3a (colored grey). Without the inclusion of fluorine-based restraints, the average pairwise backbone and heavy atom RMSD between individual members of this ensemble and the 4F-Trp,U-^{15}N-OAA X-ray structure are 1.1±0.1 Å

and 2.1±0.1 Å, respectively (PDB: 9AUQ). The average precision of the ensemble as measured by pairwise RMSD, is 0.7±0.1 Å and 1.3±0.2 Å for backbone and heavy atoms, respectively.

The above results indicate that inclusion of fluorine-based restraints led to noticeably reduced pairwise and target RMSDs (Table 2). Considerable improvements in local RMSD values were observed in key regions of OAA outside of the tightly folded β-barrel core. As expected, these improvements are associated with more restricted side chain conformations for all six Trp residues (Figure 3b) as well as other residues throughout the protein (Figure S9, Supporting Information). For example, in the carbohydrate-binding loops comprising residues N8-E19 and N75-S86 few ^{13}C-^{13}C inter-residue distance restraints were present, when compared to the β-barrel core. As a consequence, in these regions the addition of fluorine-based distance restraints leads to a marked improvement in RMSD values for several residues, especially around the Trp residue pairs, W10/W17 and W77/W84 and more restricted side chain conformations (Figure S10, Supporting Information). Specifically, the target backbone and heavy atom RMSD values for loop 1 (residues 8-19) of 4F-Trp,U-^{13}C,^{15}N-OAA showed improvements from 0.6±0.3 Å and 3.0±0.3 Å to 0.7±0.1 Å and 1.0±0.1 Å upon the addition of fluorine-based distance restraints, respectively. Similarly, the average precision RMSD values for loop 1 improved from 0.6±0.3 Å and 1.4±0.6 Å to 0.1±0.1 Å and 0.4±0.2 Å for backbone and heavy atoms, respectively. Similar improvements were observed for loop 2 (residues 75-86) with target backbone and heavy atom RMSD values changing from 0.6±0.2 Å and 3.0±0.4 Å to 0.7±0.1 Å and 1.2±0.1 Å upon the addition of fluorine-based distance restraints, respectively, and average precision RMSD values improving from 0.4±0.1 Å and 1.4±0.5 Å to 0.3±0.1 Å and 0.9±0.4 Å for backbone and heavy atoms, respectively. Furthermore, side chain conformations of residues N8, Q9, P16, I25, I36, F61, N75, Q76, and F128 improved through the inclusion of ^{13}C-^{19}F restraints. Interestingly, several other residues close to the carbohydrate-binding loops, such as V38, D42, and T131 also improved with respect to target and pairwise RMSD values, upon restraining the side chain conformations of the above set of residues.

Molecular dynamics

To further assess the structures calculated using fluorine-based pseudo potentials in XPLOR-NIH, a secondary refinement of the 10 lowest energy 4F-Trp,U-^{13}C,^{15}N-OAA models was carried out applying the ff15ipq protein force field and its compatible 4F-Trp parameters[40,41] in the AMBER molecular dynamics (MD) simulation engine.[42] Although a routine simulated annealing protocol enforcing equivalent NMR-derived restraints did not improve the target RMSD values beyond those already achieved by XPLOR-NIH, the AMBER refinement increased the frequency of the secondary structure match to the target by 8±3% on average. Additionally, the models

gained 10±2.5 Ramachandran-allowed dihedrals, and the average clash score was reduced from 28.3±2.9 to zero, with the average MolProbity score[43,44] improving from 3.2±0.1 to 2.0±0.1. Importantly, the structural improvements in the carbohydrate binding loops in the initial structure calculations (Figure 11, Supporting Information) were maintained. While secondary refinement using a rigorously parameterized and fluorine-aware AMBER force field led to modest improvements in the calculated models, the structures produced using XPLOR-NIH with newly implemented fluorine pseudo potentials demonstrate excellent quality and satisfy high standards of accuracy and precision relative to the crystallographic target.

CONCLUSIONS

We examined whether and how inclusion of fluorine-based distance restraints affects the accuracy and precision of protein structures determined by MAS NMR using the microcrystalline 4F-Trp,U-^{13}C,^{15}N-labeled OAA as a benchmark protein. In the present work we provide a general approach for ^{19}F fast MAS NMR spectroscopy based on 2D and 3D homo- and heteronuclear dipolar correlation experiments and introduce a new 3D (H)CHF HETCOR experiment. These ^{19}F-based experiments permitted unambiguous assignments of ^{19}F chemical shifts and the extraction of valuable fluorine-based long-distance restraints. The inclusion of such long-range fluorine-based distances improved the accuracy and precision of the calculated structures, particularly for residues in proximity to the Trp pairs (W10/W17 and W77/W84) in the carbohydrate-binding loops and helped to define the local conformations in the carbohydrate binding loops, for which few long-range ^{13}C-^{13}C MAS NMR-derived restraints were available. Finally, the implementation of specific parameterized force fields for aromatic fluorinated amino acids in XPLOR-NIH and AMBER permitted the inclusion of ^{19}F-based distance restraints in NMR protein structure calculation and refinement protocols. This opens doors for the application of our approach to a wide range of biological systems including protein assemblies, protein-ligand complexes and signaling receptors.

EXPERIMENTAL

Protein expression

4F-Trp,U-^{13}C,^{15}N-OAA, 4F-Trp,U-^{15}N-OAA, 5F-Trp,U-^{13}C,^{15}N-OAA, and 7F-Trp,U-^{13}C,^{15}N-OAA were expressed and purified as reported previously with modifications.[33,36] In brief, transformed *E. coli* Rosetta2 (DE3) cells were cultured in 5-10 mL of Luria-Bertani (LB) medium supplemented with 50 µg/mL kanamycin and 30 µg/mL chloramphenicol. The LB pre-culture was incubated at 37 °C until the OD_{600} reached 1.2-1.5 and 1 mL of this pre-culture was used to

inoculate 50 mL of modified M9 medium, supplemented with 1 g/L $^{15}NH_4Cl$ and 2 g/L $^{13}C_6$-glucose. After overnight growth at 37 °C, the 50 mL culture was added to 1 L of isotope-containing modified M9 medium, and grown at 37 °C up to an OD_{600} of 0.7-0.75. At that point, the temperature was lowered to 18 °C, 20 mg of 4-, 5-, 6- or 7-fluroindole was added, and cells were grown for ~1 hour. After ~1 hour, protein expression was induced with isopropylthio-β-galactoside (IPTG), at a final concentration of 500 μM. Cells were grown at 18 °C for 16-18 hours and harvested at 4000 x g for 10 min at 4 °C. Cell pellets were resuspended in lysis buffer (20 mM Tris-HCl, 3 mM NaN_3 pH 7.5) and stored at -80 °C until further use.

Protein purification

Cells were ruptured by sonication (Branson Digital Sonifier 450) at 70% power for 2 minutes (10 s pulse on and 50 s pulse off) on ice. The cellular lysate was clarified by centrifugation at 14,000 x g for 1 hour at 4 °C, and the supernatant was buffer exchanged overnight into 20 mM Tris-HCl, NaN_3 (pH 7.5) and subsequently loaded onto an anion exchange column (HiTrap Q FF 5 mL, GE Healthcare). Proteins were eluted using a 0-1 M NaCl gradient in 20 mM Tris-HCl (pH 7.5) and protein-containing fractions were pooled and concentrated (2-3 mL) prior to additional purification by size-exclusion chromatography (HiLoad 26/600 Superdex 75 prep grade, GE Healthcare) in 20 mM NaOAc buffer, 100 mM NaCl, 3 mM NaN_3 (pH 5.0). For X-ray diffraction and solid-state NMR, purified protein was buffer exchanged into 20 mM Tris-HCl, 100 mM NaCl (pH 8.0) and concentrated up to 30 mg/mL for crystallization. For solution NMR studies, proteins were buffer exchanged into 20 mM NaOAc buffer, 20 mM NaCl, 3 mM NaN_3 (pH 6.0).

Crystallization of 4F-Trp OAA

Small-scale crystallization was carried out at room temperature (~20 °C) using sitting drop vapor diffusion. 2 μL of 4F-Trp,U-^{15}N-OAA in 20 mM Tris-HCl, 100 mM NaCl, pH 8.0 were mixed with 2 μL of 1.2 M NaH_2PO_4/0.8 M K_2HPO_4 (pH 5.5), 0.2 M Li_2SO_4, and 0.1 M CAPS (pH 10.5), according to a previously published crystallization condition.[33,36]

Batch crystallization of 4F-Trp,U-^{13}C,^{15}N-OAA for MAS NMR experiments was performed by a large scale sitting drop method using volumetric proportions of 500 μL sitting drop crystallization wells, as previously reported.[45] A series of pre-sterilized Petri dishes were used to generate a concentric sitting drop vessel with a 25-75 mL reservoir and three droplet wells with an optimal volume of 300-1000 μL. The sitting drop vessel was sealed using vacuum grease and left undisturbed at 20 °C for 7-9 days. Once crystallization was complete, protein crystals were harvested and packed into Bruker 3.2 and 1.3 mm rotors by centrifugation at 4 °C at 7-10,000 x g and 10-15,000 x g, respectively.

Diffraction data collection and structure determination

X-ray diffraction data were collected in the X-ray crystallography facility of the Department of Structural Biology, University of Pittsburgh, on 4F-Trp,U-^{15}N-OAA crystals grown in 1.2 mM NaH_2PO_4/0.8 mM K_2HPO_4 (pH 5.5), 0.2 mM Li_2SO_4, and 0.1 mM CAPS (pH 10.5), pH 6.3. Data were collected up to 2.37 Å resolution on flash-cooled crystals (-180°C) using a Rigaku FR-E generator with a R-AXIS IV image plate detector at a wavelength corresponding to the copper edge (1.54 Å). All diffraction data were processed, integrated, and scaled using d*TREK software[46], and eventually converted to mtz format using the CCP4 package.[47] The crystal contained two polypeptide chains per asymmetric unit. Phases were determined by molecular replacement using the structure of WT OAA (PDB:3OBL), comprising residues 2–133 as the structural probe. All pertinent statistics are summarized in Table 1. The unit-cell dimensions of the P_1 crystals were *a* = 42.48 Å, *b* = 47.30 Å, and *c* = 47.28 Å with an estimated solvent content of ~29.0% (V_m = 1.94 Å^3/Da). Modeling of the processed MTZ maps was carried out in PHENIX[48] to automatically generate the initial structure. The generated models were refined using Coot.[49]

Solution NMR spectroscopy

2D ^{1}H-^{15}N HSQC and 1D ^{19}F solution NMR spectra were collected at 25 °C for 4F-, 5F-, and 7F-Trp,U-^{13}C,^{15}N-OAA on a 14.1 T Bruker AVIII spectrometer equipped with a triple-resonance inverse detection (TXI) probe; the ^{1}H Larmor frequency was 600.1 MHz.

MAS NMR spectroscopy

For MAS NMR experiments, 30 mg of microcrystalline 4F-Trp,U-^{13}C,^{15}N-OAA were packed into a 3.2 mm thin-wall Bruker rotor, and 2.6 mg of protein were center-packed into a 1.3 mm Bruker rotor using silicone spacers.

2D CORD, NCACX, and NCOCX MAS NMR spectra were recorded on a Bruker 20.0 T AVIII spectrometer, outfitted with a 3.2 mm E-free HCN probe. The MAS frequency was 14 kHz controlled to within ±10 Hz. The actual sample temperature was calibrated using KBr as an external temperature sensor (by recording the ^{79}Br T_1 relaxation time)[50] and was maintained at 4±1 °C. The Larmor frequencies were 850.4 MHz (^{1}H), 213.8 MHz (^{13}C), and 86.2 MHz (^{15}N). The typical 90° pulse lengths were 2.7 µs (^{1}H), 4.3 µs (^{13}C), and 4.5 µs (^{15}N). For ^{1}H-^{13}C and ^{1}H-^{15}N cross-polarization (CP) a 20% linear amplitude ramp was applied on ^{1}H, with the center of the ramp matched to the Hartmann–Hahn condition at the first spinning sideband, and contact times of 0.9–1.1 ms. The radio frequency (rf) fields during ^{1}H-^{13}C and ^{1}H-^{15}N CP were 65 kHz (^{1}H), 53 kHz (^{13}C), and 48 kHz (^{15}N). Band-selective ^{15}N-^{13}C SPECIFIC-CP employed a 20% tangent amplitude ramp on ^{15}N, the center of the ramp was matched to the Hartmann-Hahn condition at

the first spinning sideband, and the contact time was 5.0–6.0 ms. 90 kHz SPINAL-64[51] decoupling was applied during the evolution (t_1) and acquisition (t_2) times. The rf fields were 32/34 kHz (^{13}C) and 21/49 kHz (^{15}N) for NCA/NCO transfers, respectively. The CORD mixing time was 50 ms.[38]

Additional MAS NMR experiments were also carried out on a Magnex/Bruker 14.1 T AVIII HD spectrometer. The Larmor frequencies were 599.8 MHz (^{1}H), 150.8 MHz (^{13}C), and 60.8 MHz (^{15}N). The sample temperature was calibrated with KBr as described above and maintained at 37±1 °C. ^{1}H detected 2D (H)NH and (H)CH HETCOR spectra were recorded using a 1.3 mm HCN MAS probe at a MAS frequency of 60 kHz, controlled to within ±5 Hz. The typical 90° pulse lengths were 1.4 µs (^{1}H), 2.7 µs (^{13}C), and 3.2 µs (^{15}N). For ^{1}H-^{13}C and ^{1}H-^{15}N CP a linear 20% amplitude ramp was applied on ^{1}H, with the center of the ramp matched to the Hartmann–Hahn condition at the first spinning sideband, and ^{1}H-^{15}N and ^{1}H-^{13}C CP contact times were 600 µs and 800 µs, respectively. The ^{13}C-^{1}H and ^{15}N-^{1}H CP contact times were 200 µs. The rf fields were 150 kHz (^{1}H) and 86 kHz (^{13}C) during ^{1}H-^{13}C CP, and 105 kHz (^{1}H) and 45 kHz (^{15}N) during ^{1}H-^{15}N CP. 2D CORD spectra were recorded using a 3.2 mm E-free HCN probe at a MAS frequency of 14 kHz, controlled to within ±5 Hz. The CORD mixing times were 25 ms, 100 ms, 250 ms, and 500 ms. Typical pulse lengths were similar to those for the experiments performed at 20.0 T (see above).

Heteronuclear ^{19}F-detected 2D (H)CF and 3D (H)CHF HETCOR experiments on 4F-Trp,U-^{13}C,^{15}N-OAA crystals were performed on an 11.7 T wide-bore Bruker AVIII spectrometer outfitted with a 1.3 mm HFX MAS probe. The Larmor frequencies were 500.1 MHz (^{1}H), 470.6 MHz (^{19}F), and 125.8 MHz (^{13}C). All MAS NMR spectra were acquired at a MAS frequency of 60 kHz, controlled to within ±10 Hz. The sample temperature was calibrated with KBr as described above and maintained at 37±1 °C. Typical pulse lengths were 2.3 µs (^{1}H), 2.4 µs (^{19}F), and 2.9 µs (^{13}C). For the 2D (H)CF HETCOR experiment, ^{13}C-^{19}F CP was performed with a 30% linear amplitude ramp on ^{13}C, with the center of the ramp Hartmann-Hahn-matched to the first spinning sideband; the rf fields were 16 kHz (^{19}F) and 42 kHz (^{13}C). Swept-frequency two-pulse phase modulation (SWf-TPPM)[52] ^{1}H decoupling (15 kHz) was applied during acquisition. ^{19}F decoupling was performed with a rotor-synchronized π-pulse during evolution in the ^{13}C dimension. The ^{13}C-^{19}F CP contact time was 3.0 ms. The recycle delay was 2 s. For the 3D (H)CHF HETCOR experiment, ^{1}H-^{19}F CP was carried out with a 20% linear amplitude ramp on ^{1}H, with the center of the ramp Hartmann-Hahn-matched to the first spinning sideband; the rf fields were 108 kHz (^{1}H) and 49 kHz (^{19}F). ^{13}C-^{1}H and ^{1}H-^{19}F CP contact times were 1.0 ms and 1.5 ms, respectively. 10 kHz WALTZ-16[53] ^{13}C decoupling was applied during the ^{1}H evolution and acquisition periods. 15 kHz SWf-TPPM ^{1}H decoupling was applied during the ^{13}C evolution and acquisition periods. The recycle delay was 1.4 s.

The 2D ^{19}F-^{19}F RFDR spectrum was recorded at Bruker Billerica, on a 11.7 T wide bore Bruker AVIII spectrometer, outfitted with a 1.3 mm HFX MAS probe. The Larmor frequencies were 500.1 MHz (^{1}H), 470.6 MHz (^{19}F), and 125.8 MHz (^{13}C). The MAS frequency was 50 kHz, controlled to within ±10 Hz. The sample temperature was calibrated with KBr as described above and maintained at 24±1 °C. The ^{19}F pulse length was 2.0 µs. RFDR $(XY8)^{1}_{4}$ phase cycle was used; the RFDR mixing time was 10 ms. SWf-TPPM[52] ^{1}H decoupling (15 kHz) was applied during the acquisition.

Additional details on the experimental settings are summarized in Table S2, Supporting Information.

Data processing

All MAS NMR data were processed using Bruker TopSpin and/or NMRPipe.[54] ^{1}H chemical shifts were indirectly referenced to the external standard adamantane ^{13}C chemical shift.[55] ^{13}C and ^{15}N chemical shifts were referenced to the external standards adamantane[56] and ammonium chloride,[57] respectively. ^{19}F chemical shifts were referenced to trifluoroacetic acid (100 µM solution in 25 mM phosphate buffer, pH 6.5), used as an external standard (0 ppm). The 2D and 3D data sets were processed by applying 30°, 45°, 60°, and 90° shifted sine bell apodization, followed by a Lorentzian-to-Gaussian transformation in both dimensions. 2D and 3D ^{1}H-detected data sets were processed with Gaussian and/or square sine window multiplication and quadrature baseline correction.

Resonance assignments

All solution and MAS NMR spectra were analyzed using NMRFAM-Sparky[58,59] and CCPN.[60] ^{1}H and ^{15}N solution chemical shifts were assigned based on the published WT OAA chemical shifts (BMRB:25306).[37] ^{13}C and ^{15}N chemical shift assignments of microcrystalline 4F-Trp,U-^{13}C,^{15}N-OAA were transferred from the chemical shifts of microcrystalline WT OAA reported by us previously,[36] and verified by *de novo* backbone and side chain assignments, using 2D ^{13}C-^{13}C CORD (50 ms mixing time), 2D/3D NCACX (50 ms mixing time) and NCOCX (25 ms mixing time) experiments. ^{1}H backbone and side chain resonances were assigned using ^{1}H-deteced 2D (H)NH HETCOR and (H)CH HETCOR spectra. ^{19}F chemical shifts were assigned based on ^{19}F-deteced 2D RFDR, (H)CF, and 3D (H)CHF spectra.

Parameterization of fluorinated amino acids in XPLOR-NIH

XPLOR-NIH structure calculations employed pseudo potentials for the fluorinated amino acids. A total of 8 pseudo potentials were constructed, one for each of the fluorinated amino acid variants from the AMBER ff15ipq_fl force field.[41] Topology and parameters were

taken from the AMBER force field and used to design the pseudo potentials. Bond lengths were set to their equilibrium values, while bond angles and planar improper torsions were set to their geometrically idealized values. For consistency with standard XPLOR-NIH calculations, which do not use explicit electrostatic interactions, no realistic charge information was included in the pseudo potentials.

Structure calculation

The complete set of experimental ^{13}C-^{13}C distance restraints obtained from 2D ^{13}C-^{13}C CORD spectra recorded with the mixing times of 100 ms, 250 ms, and 500 ms was cross checked against the complete set of ^{13}C-^{13}C distances up to 7 Å generated from the 4F-Trp OAA X-ray structure using scripts in Python 2.7 Biopython[61] and Bio-PDB[62] modules. From this approach, a final set of medium and long-range C-C distances was obtained with up to 4-fold ambiguity.^{19}F-^{19}F and ^{19}F-^{13}C distance restraints were extracted from the following spectra: 2D RFDR (mixing time of 10 ms), 2D (H)CF HETCOR (^{13}C-^{19}F contact time of 3.0 ms), and 3D (H)CHF HETCOR (contact times of 1.0 ms (^{13}C-^{1}H) and 1.5 ms (^{1}H-^{19}F), respectively).

The final MAS NMR structure was calculated in XPLOR-NIH version 2.53[39,63,64] using 5671 distance restraints. Distances were supplemented with backbone phi (φ) and psi (ψ) dihedral angles predicted by TALOS+[65] from experimental ^{13}C and ^{15}N chemical shifts. The sidechain dihedral angle xi (χ) was restrained to 90±20 degrees for each Trp residue based on solution NMR chemical shifts.[66] ^{13}C-^{13}C distance restraints were set to 1.5-6.5 Å (4.0±2.5 Å) and 2.0-7.2 Å (4.6±2.6 Å) for intra- and inter-residue correlations, respectively.[2,3,45] The minimum and maximum XPLOR bounds for ^{19}F-based distance restraints were set as follows: 4.0-8.0 Å (6.0±2.0 Å) for ^{19}F-^{19}F inter-residue restraints, 2.0-4.2 Å (3.1±1.1 Å) to 2.0-8.0 Å (5.0±3.0 Å) for lower- and upper-bound ^{13}C-^{19}F intra-residue restraints, and 2.0-3.8 Å (2.9±0.9 Å) to 2.0-15.0 Å (10.0±5.0 Å) for lower- and upper-bound ^{13}C-^{19}F inter-residue restraints, respectively. The ^{13}C-^{19}F distance restraint bounds were estimated from Trp C^{β}-F cross peak intensities used as a ruler (C-F interatomic distance of 3.3 Å), to set the initial values. The bounds for other restraints were then estimated from the corresponding cross peak intensities and by taking into account the approximate $1/r^3$ dependence of intensity on interatomic distance. The lower bound was set to 2 Å, while the upper bound was set to the estimated distance +1 Å.

Single-chain calculations were seeded from an extended strand of the primary sequence and 1000 structures were generated using torsion angle dynamics with two consecutive annealing schedules and a final energy minimization in Cartesian space, as described previously.[2,3,45] The first annealing calculation was performed at 3500 K for 800 ps or 8000 steps, keeping the energy constant. This was followed by a second phase of simulated annealing, applied to the 10 lowest

energy structures, with lowering the temperature from 3000 K to 25 K in 12.5 K steps. During the initial simulated annealing force constraints for distance restraints were ramped from 10 to 50 kcal/mol/Å^2; in the second phase they were ramped from 2 to 30 kcal/mol/Å^2. Dihedral restraints were excluded from high temperature dynamics run at 3500 K, and set to 200 kcal/mol/rad^2 during simulated annealing. During both annealing schedules the force constant for the radius of gyration was scaled from 0.002 to 1. Final structures were minimized using a Powell energy minimization scheme and the 10 lowest energy structures were used for refinement.

Structure refinement with the ff15ipq force field

The 10 lowest energy structures obtained from XPLOR-NIH were refined using a simulated annealing protocol, applying the ff15ipq and ff15ipq_fl force fields.[40,41] The refinement was performed using *pmemd.cuda*[67] from the AMBER 22 package.[42] Distance restraints used for structure calculations in XPLOR-NIH were converted to AMBER readable format, such that an identical restraint set was used in both programs. Structures were subjected to 10,000 steps of energy minimization using the steepest descent algorithm for the first 1,000 steps, followed by conjugate gradient minimization for the remainder. Simulated annealing was carried out in implicit solvent[68] with a time step of 1 fs. Non-bonded interactions were computed in direct space. Temperature was controlled using the Berendsen thermostat with a coupling constant of 1 ps. Structures were heated from 0 to 500 K over 250 ps, maintained at high temperature for 500 ps, then cooled from 500 to 0 K over an additional 250 ps. The annealing protocol was run in triplicate for each structure from the XPLOR-NIH ensemble, and the lowest energy structure out of three was selected for the AMBER-refined ensemble. Secondary structure assignment used the STRIDE[69,70] implementation in VMD 1.9.3.[71] Structural quality statistics were calculated using MolProbity.[43]

Structure analysis and visualization

The atomic RMSD values were calculated for each of 10 lowest energy structures for backbone atoms (N, C^{α}, $C^{'}$) and heavy atoms with respect to the X-ray crystal structure (PDB:9AUQ) in XPLOR-NIH (version 2.53).[39,63,64] Rendering and visualization of the structures was performed in PyMOL[72].

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation (NSF grant CHE1708773) and the National Institutes of Health (NIH Grant 1U54AI170791 to A.M.G. and T.P., NMR Core). We acknowledge the support of the National Science Foundation (NSF Grant CHE0959496) for the acquisition of the 850 MHz NMR spectrometer at the University of Delaware. CDS was supported by the Intramural Program of the National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health.

AUTHOR CONTRIBUTIONS

R.Z. and B.R.R. contributed equally to this work.

DATA AVAILABILITY

Coordinates of MAS NMR and X-ray structures have been deposited at the Protein Data Bank with accession codes PDB ID 9B4V, PDB ID 9B4W and PDB ID 9AUQ, respectively. The MAS NMR chemical shifts are deposited at the Biological Magnetic Resonance Bank (BMRB) under BMRB entry IDs 31152 and 31153.

DECLARATION OF INTERESTS

The authors declare no conflict of interest.

SUPPORTING INFORMATION AVAILABLE

Mass spectra, ^{19}F and ^{1}H-^{15}N HSQC spectra; ^{13}C CPMAS, 2D (H)NH HETCOR and (H)CH HETCOR spectra; 2D CORD spectra; graphical summary of resonance assignments; distance dependences of ^{19}F-^{19}F RFDR signal intensities and ^{19}F-^{19}F dipolar coupling; ^{13}C-based distance restraints; comparison of side chain conformations and RMSDs of MAS NMR structural ensembles calculated with and without ^{19}F-based distance restraints; shortest ^{19}F-^{19}F intrachain and interchain distances in crystalline 4F-Trp; summary of MAS NMR experiments; author contributions. This material is available free of charge via the internet at http://pubs.acs.org.

Table 1 | X-ray data collection and refinement statistics

4F-Trp,U-^{15}N OAA	
Data collection	
Wavelength (Å)	1.54
Space group	P 1
Cell dimensions	
a, *b*, *c* (Å)	42.48, 47.30, 47.28
α, β, γ (°)	78.44, 63.36, 63.37
Resolution range (Å)	22.64 - 2.37 (2.455 - 2.37)
Completeness (%)	95.07 (87.24)
Wilson B-factor, Å^2	15.32
Refinement	
Refinement program	Coot
Resolution range (Å)	29-2.37
No. reflections	11298
R_{work}/R_{free}	16.8/23.5
No. of nonhydrogen atoms	
Protein	1984
Ligand	56
Solvent (water)	100
B-factors	
Protein	14.13
Ligand	17.80
Solvent (water)	15.30
r.m.s.d deviations	
Bond lengths (Å)	0.01
Bond angles (°)	1.20
PDB ID	9AUQ

*Values in parentheses are for highest-resolution shell.

Table 2 | Summary of MAS NMR restraints and structure statistics

^{13}C-^{13}C MAS NMR distance restraints **up to 4-fold ambiguity, # unambiguous indicated in ()*				
Intra-residue	673 (220)			
Sequential (|i-j|=1)	1264 (381)			
Medium range (1<|i-j|<5)	976 (276)			
Long range (|i-j|≥5)	2632 (681)			
^{19}F-(^{13}C, ^{19}F) MAS NMR distance restraints **up to 2-fold ambiguity, # unambiguous indicated in ()*	***Total***	***F-F RFDR***	***CF 2D***	***CHF 3D***
r < 5 Å	58 (55)	0	58 (55)	0
5 Å ≤ r < 8 Å	44 (43)	3 (3)	37 (36)	4 (4)
r ≥ 8 Å	24 (18)	0	23 (17)	1 (1)
^{19}F Total	126 (95)	3 (3)	118 (108)	5 (5)
Total number of restraints assigned	***5671***			
TALOS-derived dihedral angle restraints				
Φ	99			
Ψ	99			
X (for Trp)	6			
Structure statistics for 10 lowest energy structures				
Violations (mean ± s.d.)	***w ^{19}F***		***w/o ^{19}F***	
Distance restraints (Å)	0.2 ± 0.1		0.2 ± 0.1	
Dihedral angle restraints (°)	6.0 ± 3.6		6.0 ± 3.5	
Max. distance restraint violation (Å)	0.9		1.0	
Max. dihedral angle restraint violation (°)	28.9		27.6	
Deviations from idealized geometry	***w ^{19}F***		***w/o ^{19}F***	
Bond lengths (Å)	0.0 ± 0.0		0.0 ± 0.0	
Bond angles (°)	0.8 ± 0.0		0.7 ± 0.0	
Impropers (°)	0.8 ± 0.0		0.7 ± 0.0	
Average pairwise r.m.s.d. (Å)	***w ^{19}F***		***w/o ^{19}F***	
Backbone (N, C^{α}, C')	0.4 ± 0.1		0.7 ± 0.1	
Heavy	0.9 ± 0.1		1.3 ± 0.2	
Average Target (X-ray structure) r.m.s.d. (Å)	***w ^{19}F***		***w/o ^{19}F***	
Backbone (N, C^{α}, C')	1.1 ± 0.1		1.1 ± 0.1	
Heavy	1.4 ± 0.1		2.1 ± 0.1	

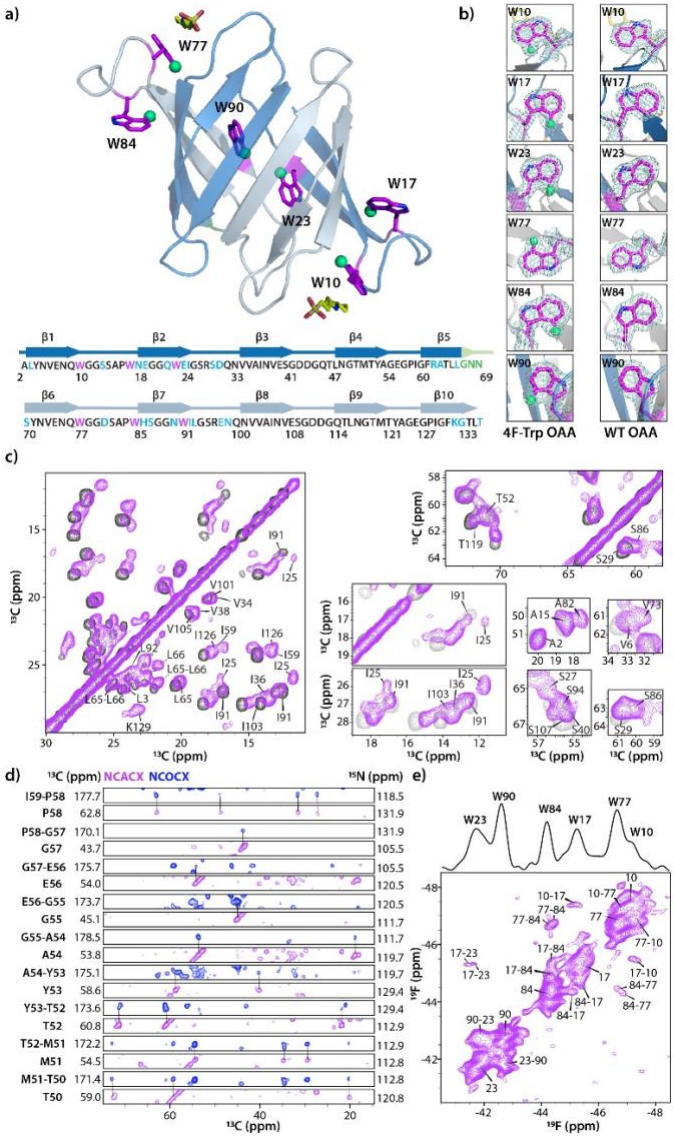


Figure 1 | a) Cartoon representation of the X-ray structure (top) and amino acid sequence and secondary structure (bottom) of 4F-Trp *O. agardhii* agglutinin (OAA). The 4F-Trp residues and bound CAPS molecules are shown in stick representation and colored magenta and yellow, respectively. The fluorine atoms are represented by green spheres. Amino acid sequence repeat 1 (residues 1-66) and repeat 2 (residues 70-133) in the structure are colored in blue and gray, respectively; residues between the repeats (G67-N69) are colored in green. In the sequence, conserved amino acids are in black type, non-conserved ones in cyan, residues between the repeats (G67-N69) are in green, and the Trp residues are in magenta. **b)** Electron density for the 4F-Trp and Trp side chains, contoured at 2.0 Å (blue mesh) in 4F-Trp OAA (left) and WT OAA (right). **c)** Superposition of representative regions in the 2D CORD spectra of crystalline WT U-^{13}C,^{15}N-OAA (gray) and 4F-Trp,U-^{13}C,^{15}N-OAA (purple). Cross peaks for residues that possess different chemical shifts in WT U-^{13}C,^{15}N-OAA and 4F-Trp,U-^{13}C,^{15}N-OAA are labeled by residue name, number and atom type. **d)** Backbone walk for the residue stretch T50-I59 in the 3D NCACX (purple) and 3D NCOCX (blue) spectra of crystalline 4F-Trp,U-^{13}C,^{15}N-OAA. **e)** 2D ^{19}F-^{19}F RFDR spectrum of 4F-Trp,U-^{13}C,^{15}N-OAA acquired with a mixing time of 10 ms and a MAS frequency of 50 kHz. The spectrum was processed with linear prediction of 256 points in the t_1 dimension, 60º and 90º squared sinebell apodization and zero filling to 4096 and 512 points in t_2 and t_1 dimensions, respectively. The corresponding 1D ^{19}F MAS spectrum is shown on the top.

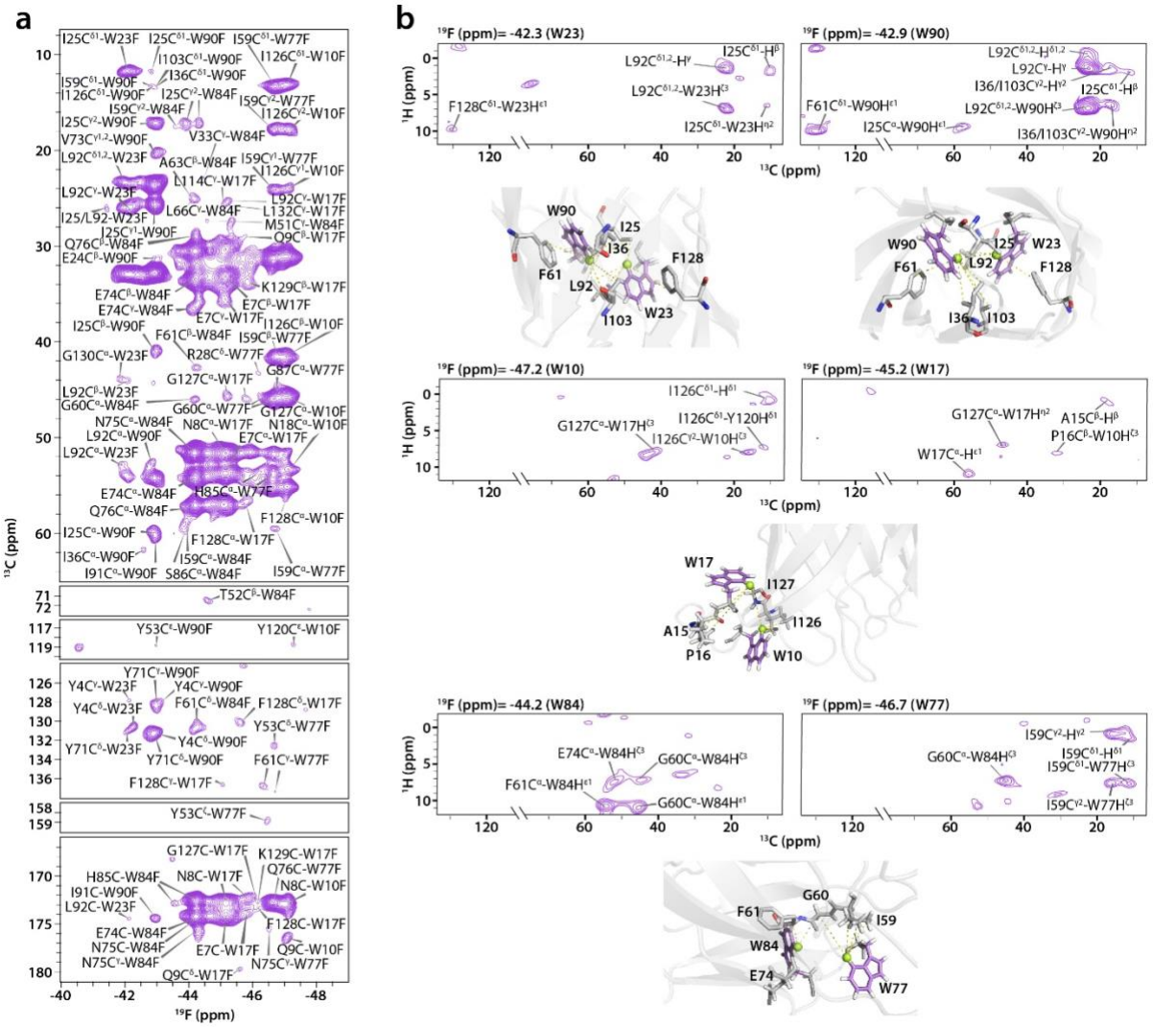


Figure 2 | a) Aliphatic, aromatic, and carbonyl regions of a 2D (H)CF HETCOR spectrum of 4F-Trp,U-^{13}C,^{15}N-OAA. Selected assignment labels are shown, full assignments are presented in Figure S11, Supporting Information. **b)** Selected 2D ^{1}H-^{13}C planes from a 3D (H)CHF spectrum of 4F-Trp,U-^{13}C,^{15}N-OAA, with representative ^{13}C-^{19}F interatomic distances extracted from 3D (H)CHF HETCOR spectra, shown on the 3D structures below each spectrum. The planes are shown for W23 (δ_{iso}(^{19}F)= -42.3 ppm, top left), W90 (δ_{iso}(^{19}F)= -42.9 ppm, top right), W10 (δ_{iso}(^{19}F)= -47.2 ppm, middle left), W17 (δ_{iso}(^{19}F)= -45.2 ppm, middle right), W84 (δ_{iso}(^{19}F)= -44.2 ppm, bottom left), and W77 (δ_{iso}(^{19}F)= -46.7 ppm, bottom right). A total of 123 ^{13}C-^{19}F inter-residue correlations were experimentally determined for W23 (16 correlations), W90 (23 correlations), W84 (26 correlations), W17 (27 correlations), W77 (17 correlations), and W10 (14 correlations).

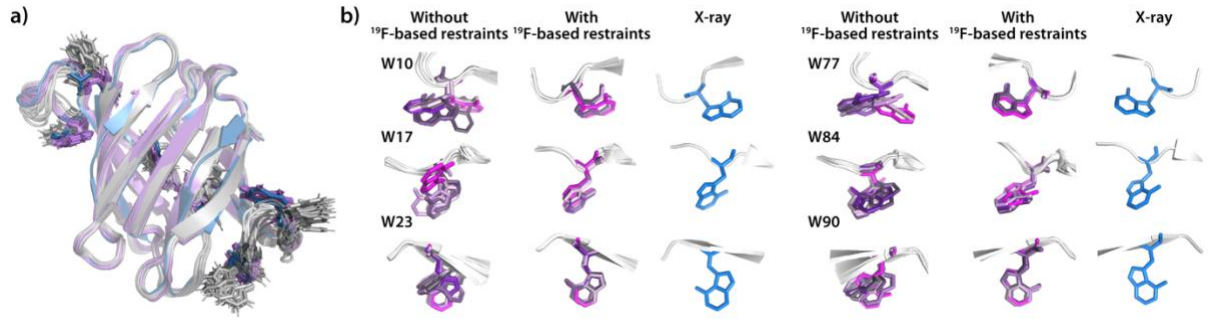


Figure 3 | a) Superposition of 10 lowest energy MAS NMR structures of 4F-Trp OAA calculated without (grey) and with (purple) the inclusion of ^{19}F-based interatomic distance restraints, together with the X-ray crystal structure (blue). **b)** Sidechain conformations of Trp residues in the 10 lowest energy MAS NMR structures of 4F-Trp OAA calculated without (left panel) and with (middle panel) ^{19}F-based interatomic distance restraints, and in the X-ray structure (right panel).

REFERENCES

(1) Kraus, J.; Sarkar, S.; Quinn, C. M.; Polenova, T. Solid-state NMR spectroscopy of microcrystalline proteins, In *Annual Reports on NMR Spectroscopy*; Webb, G. A., Ed.; Academic Press: **2021**; Vol. 102, p 81.

(2) Lu, M.; Russell, R. W.; Bryer, A. J.; Quinn, C. M.; Hou, G.; Zhang, H.; Schwieters, C. D.; Perilla, J. R.; Gronenborn, A. M.; Polenova, T. Atomic-resolution structure of HIV-1 capsid tubes by magic-angle spinning NMR, *Nat. Struct. Mol. Biol.* **2020**, *27*, 863.

(3) Russell, R. W.; Fritz, M. P.; Kraus, J.; Quinn, C. M.; Polenova, T.; Gronenborn, A. M. Accuracy and precision of protein structures determined by magic angle spinning NMR spectroscopy: for some 'with a little help from a friend', *J. Biomol. NMR* **2019**, *73*, 333.

(4) Perilla, J. R.; Zhao, G.; Lu, M.; Ning, J.; Hou, G.; Byeon, I.-J. L.; Gronenborn, A. M.; Polenova, T.; Zhang, P. Cryo-EM structure refinement by integrating NMR chemical shifts with molecular dynamics simulations, *J. Phys. Chem. B* **2017**, *121*, 3853.

(5) Gauto, D. F.; Estrozi, L. F.; Schwieters, C. D.; Effantin, G.; Macek, P.; Sounier, R.; Sivertsen, A. C.; Schmidt, E.; Kerfah, R.; Mas, G.; Colletier, J.-P.; Güntert, P.; Favier, A.; Schoehn, G.; Schanda, P.; Boisbouvier, J. Integrated NMR and cryo-EM atomic-resolution structure determination of a half-megadalton enzyme complex, *Nat. Commun.* **2019**, *10*, 2697.

(6) Demers, J.-P.; Habenstein, B.; Loquet, A.; Kumar Vasa, S.; Giller, K.; Becker, S.; Baker, D.; Lange, A.; Sgourakis, N. G. High-resolution structure of the Shigella type-III secretion needle by solid-state NMR and cryo-electron microscopy, *Nat. Commun.* **2014**, *5*, 4976.

(7) Jehle, S.; Rajagopal, P.; Bardiaux, B.; Markovic, S.; Kühne, R.; Stout, J. R.; Higman, V. A.; Klevit, R. E.; Van Rossum, B.-J.; Oschkinat, H. Solid-state NMR and SAXS studies provide a structural basis for the activation of αB-crystallin oligomers, *Nat. Struct. Mol. Biol.* **2010**, *17*, 1037.

(8) Lattanzi, V.; André, I.; Gasser, U.; Dubackic, M.; Olsson, U.; Linse, S. Amyloid β 42 fibril structure based on small-angle scattering, *Proc. Natl. Acad. Sci. U. S. A.* **2021**, *118*.

(9) Schmidt, H. L.; Sperling, L. J.; Gao, Y. G.; Wylie, B. J.; Boettcher, J. M.; Wilson, S. R.; Rienstra, C. M. Crystal polymorphism of protein GB1 examined by solid-state NMR spectroscopy and X-ray diffraction, *J. Phys. Chem. B* **2007**, *111*, 14362.

(10) Bardiaux, B.; Favier, A.; Etzkorn, M.; Baldus, M.; Böckmann, A.; Nilges, M.; Malliavin, T. E. Simultaneous use of solution, solid-state NMR and X-ray crystallography to study the conformational landscape of the Crh protein during oligomerization and crystallization, *Adv. Appl. Bioinform. Chem.* **2010**, *3*, 25.

(11) Bogdanov, A.; Goldfarb, D. Electron Paramagnetic Resonance in Structural Biology, In *Integrated Structural Biology*; Polenova, T., Quinn, C. M., Gronenborn, A. M., Eds.; Royal Society of Chemistry: **2023**; Vol. 30, p 0.

(12) Kandasamy, S. K.; Lee, D.-K.; Nanga, R. P. R.; Xu, J.; Santos, J. S.; Larson, R. G.; Ramamoorthy, A. Solid-state NMR and molecular dynamics simulations reveal the oligomeric ion-channels of TM2-$GABA_A$ stabilized by intermolecular hydrogen bonding, *Biochim. Biophys. Acta. Biomembr.* **2009**, *1788*, 686.

(13) Lange, A.; Gattin, Z.; Van Melckebeke, H.; Wasmer, C.; Soragni, A.; van Gunsteren, W. F.; Meier, B. H. A combined solid-state NMR and MD characterization of the stability and dynamics of the HET-s(218-289) prion in its amyloid conformation, *Chembiochem* **2009**, *10*, 1657.

(14) Linser, R.; Bardiaux, B.; Higman, V.; Fink, U.; Reif, B. Structure calculation from unambiguous long-range amide and methyl ^1H-^1H distance restraints for a microcrystalline protein with MAS solid-state NMR spectroscopy, *J. Am. Chem. Soc.* **2011**, *133*, 5905.

(15) Bayro, M. J.; Huber, M.; Ramachandran, R.; Davenport, T. C.; Meier, B. H.; Ernst, M.; Griffin, R. G. Dipolar truncation in magic-angle spinning NMR recoupling experiments, *J. Chem. Phys.* **2009**, *130*, 114506.

(16) Danielson, M. A.; Falke, J. J. Use of ^{19}F NMR to probe protein structure and conformational changes, *Annu. Rev. Biophys.* **1996**, *25*, 163.

(17) Campos-Olivas, R.; Aziz, R.; Helms, G. L.; Evans, J. N. S.; Gronenborn, A. M. Placement of ^{19}F into the center of GB1: effects on structure and stability, *FEBS Lett.* **2002**, *517*, 55.

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

(19) Salwiczek, M.; Nyakatura, E. K.; Gerling, U. I. M.; Ye, S.; Koksch, B. Fluorinated amino acids: compatibility with native protein structures and effects on protein–protein interactions, *Chem. Soc. Rev.* **2012**, *41*, 2135.

(20) Shcherbakov, A. A.; Mandala, V. S.; Hong, M. High-sensitivity detection of nanometer ^1H–^{19}F distances for protein structure determination by ^1H-detected fast MAS NMR, *J. Phys. Chem. B* **2019**, *123*, 4387.

(21) Wang, M.; Lu, M.; Fritz, M. P.; Quinn, C. M.; Byeon, I. J. L.; Byeon, C. H.; Struppe, J.; Maas, W.; Gronenborn, A. M.; Polenova, T. Fast magic-angle spinning ^{19}F NMR spectroscopy of HIV-1 capsid protein assemblies, *Angew. Chem., Int. Ed.* **2018**, *57*, 16375.

(22) Quinn, C. M.; Zadorozhnyi, R.; Struppe, J.; Sergeyev, I. V.; Gronenborn, A. M.; Polenova, T. Fast ^{19}F magic-angle spinning nuclear magnetic resonance for the structural characterization of active pharmaceutical ingredients in blockbuster drugs, *Anal. Chem.* **2021**, *93*, 13029.

(23) Roos, M.; Wang, T.; Shcherbakov, A. A.; Hong, M. Fast magic-angle-spinning ^{19}F spin exchange NMR for determining nanometer ^{19}F–^{19}F distances in proteins and pharmaceutical compounds, *J. Phys. Chem. B* **2018**, *122*, 2900.

(24) Lu, M.; Sarkar, S.; Wang, M.; Kraus, J.; Fritz, M.; Quinn, C. M.; Bai, S.; Holmes, S. T.; Dybowski, C.; Yap, G. P. A.; Struppe, J.; Sergeyev, I. V.; Maas, W.; Gronenborn, A. M.; Polenova, T. ^{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* **2018**, *122*, 6148.

(25) Guo, C.; Fritz, M. P.; Struppe, J.; Wegner, S.; Stringer, J.; Sergeyev, I. V.; Quinn, C. M.; Gronenborn, A. M.; Polenova, T. Fast ^{19}F magic angle spinning NMR crystallography for structural characterization of fluorine-containing pharmaceutical compounds, *Anal. Chem.* **2021**, *93*, 8210.

(26) Kalabekova, R.; Quinn, C. M.; Movellan, K. T.; Gronenborn, A. M.; Akke, M.; Polenova, T. ^{19}F fast magic-angle spinning NMR spectroscopy on microcrystalline complexes of fluorinated ligands and the carbohydrate recognition domain of galectin-3, *Biochemistry* **2024**, *articles ASAP*.

(27) Shcherbakov, A. A.; Roos, M.; Kwon, B.; Hong, M. Two-dimensional ^{19}F–^{13}C correlation NMR for ^{19}F resonance assignment of fluorinated proteins, *J. Biomol. NMR* **2020**, *74*, 193.

(28) Lu, X.; Skomski, D.; Thompson, K. C.; McNevin, M. J.; Xu, W.; Su, Y. Three-dimensional NMR spectroscopy of fluorinated pharmaceutical solids under ultrafast magic angle spinning, *Anal. Chem.* **2019**, *91*, 6217.

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

(30) Porat-Dahlerbruch, G.; Struppe, J.; Quinn, C. M.; Gronenborn, A. M.; Polenova, T. Determination of accurate ^{19}F chemical shift tensors with R-symmetry recoupling at high MAS frequencies (60-100 kHz), *J. Magn. Reson.* **2022**, *340*, 107227.

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

(32) Viger-Gravel, J.; Avalos, C. E.; Kubicki, D. J.; Gajan, D.; Lelli, M.; Ouari, O.; Lesage, A.; Emsley, L. ^{19}F magic angle spinning dynamic nuclear polarization enhanced NMR spectroscopy, *Angew. Chem., Int. Ed. Engl.* **2019**, *58*, 7249.

(33) Koharudin, L. M. I.; Furey, W.; Gronenborn, A. M. Novel fold and carbohydrate specificity of the potent anti-HIV cyanobacterial lectin from *Oscillatoria agardhii*, *J. Biol. Chem.* **2011**, *286*, 1588.

(34) Koharudin, L. M. I.; Kollipara, S.; Aiken, C.; Gronenborn, A. M. Structural insights into the anti-HIV activity of the *Oscillatoria agardhii* agglutinin homolog lectin family, *J. Biol. Chem.* **2012**, *287*, 33796.

(35) Fritz, M.; Quinn, C. M.; Wang, M.; Hou, G.; Lu, X.; Koharudin, L. M. I.; Polenova, T.; Gronenborn, A. M. Toward closing the gap: quantum mechanical calculations and experimentally measured chemical shifts of a microcrystalline lectin, *J Phys Chem B* **2017**, *121*, 3574.

(36) Fritz, M.; Quinn, C. M.; Wang, M.; Hou, G.; Lu, X.; Koharudin, L. M. I.; Struppe, J.; Case, D. A.; Polenova, T.; Gronenborn, A. M. Determination of accurate backbone chemical shift tensors in microcrystalline proteins by integrating MAS NMR and QM/MM, *Chem. Phys.* **2018**, *20*, 9543.

(37) Carneiro, M. G.; Koharudin, L. M. I.; Griesinger, C.; Gronenborn, A. M.; Lee, D. ^{1}H, ^{13}C and ^{15}N resonance assignment of the anti-HIV lectin from *Oscillatoria agardhii*, *Biomol. NMR. Assign.* **2015**, *9*, 317.

(38) Hou, G.; Yan, S.; Trébosc, J.; Amoureux, J.-P.; Polenova, T. Broadband homonuclear correlation spectroscopy driven by combined $R2_n^v$ sequences under fast magic angle spinning for NMR structural analysis of organic and biological solids, *J. Magn. Reson.* **2013**, *232*, 18.

(39) Schwieters, C. D.; Kuszewski, J. J.; Tjandra, N.; Marius Clore, G. The Xplor-NIH NMR molecular structure determination package, *J. Magn. Reson.* **2003**, *160*, 65.

(40) Debiec, K. T.; Cerutti, D. S.; Baker, L. R.; Gronenborn, A. M.; Case, D. A.; Chong, L. T. Further along the road less traveled: AMBER ff15ipq, an original protein force field built on a self-consistent physical model, *J. Chem. Theory Comput.* **2016**, *12*, 3926.

(41) Yang, D. T.; Gronenborn, A. M.; Chong, L. T. Development and validation of fluorinated, aromatic amino acid parameters for use with the AMBER ff15ipq protein force field, *J. Phys. Chem. A* **2022**, *126*, 2286.

(42) Case, D. A.; Aktulga, H. M.; Belfon, K., et al. AMBER 2022. University of California, San Francisco **2022**.

(43) Chen, V. B.; Arendall, W. B., 3rd; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C. MolProbity: all-atom structure validation for macromolecular crystallography, *Acta Crystallogr D Biol Crystallogr* **2010**, *66*, 12.

(44) Williams, C. J.; Headd, J. J.; Moriarty, N. W.; Prisant, M. G.; Videau, L. L.; Deis, L. N.; Verma, V.; Keedy, D. A.; Hintze, B. J.; Chen, V. B.; Jain, S.; Lewis, S. M.; Arendall, W. B., 3rd; Snoeyink, J.; Adams, P. D.; Lovell, S. C.; Richardson, J. S.; Richardson, D. C. MolProbity: More and better reference data for improved all-atom structure validation, *Protein Sci* **2018**, *27*, 293.

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

(46) Pflugrath, J. W. The finer things in X-ray diffraction data collection, *Acta Crystallogr.* **1999**, *55*, 1718.

(47) Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G. W.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S. Overview of the CCP4 suite and current developments, *Acta Crystallogr.* **2011**, *67*, 235.

(48) Adams, P. D.; Grosse-Kunstleve, R. W.; Hung, L.-W.; Ioerger, T. R.; McCoy, A. J.; Moriarty, N. W.; Read, R. J.; Sacchettini, J. C.; Sauter, N. K.; Terwilliger, T. C. PHENIX: building new software for automated crystallographic structure determination, *Acta Crystallogr.* **2002**, *58*, 1948.

(49) Emsley, P.; Cowtan, K. Coot: model-building tools for molecular graphics, *Acta Crystallogr.* **2004**, *60*, 2126.

(50) Thurber, K. R.; Tycko, R. Measurement of sample temperatures under magic-angle spinning from the chemical shift and spin-lattice relaxation rate of ^{79}Br in KBr powder, *J. Magn. Reson.* **2009**, *196*, 84.

(51) Fung, B. M.; Khitrin, A. K.; Ermolaev, K. An improved broadband decoupling sequence for liquid crystals and solids, *J. Magn. Reson.* **2000**, *142*, 97.

(52) Thakur, R. S.; Kurur, N. D.; Madhu, P. K. Swept-frequency two-pulse phase modulation for heteronuclear dipolar decoupling in solid-state NMR, *Chem. Phys. Lett.* **2006**, *426*, 459.

(53) Shaka, A. J.; Keeler, J.; Frenkiel, T.; Freeman, R. An improved sequence for broadband decoupling: WALTZ-16, *J. Magn. Reson. (1969)* **1983**, *52*, 335.

(54) Delaglio, F.; Grzesiek, S.; Vuister, G.; Zhu, G.; Pfeifer, J.; Bax, A. NMRPipe: a multidimensional spectral processing system based on UNIX pipes, *J. Biomol. NMR* **1995**, *6*.

(55) Markley, J. L.; Bax, A.; Arata, Y.; Hilbers, C. W.; Kaptein, R.; Sykes, B. D.; Wright, P. E.; Wüthrich, K. Recommendations for the presentation of NMR structures of proteins and nucleic acids, *J. Mol. Biol.* **1998**, *280*, 933.

(56) Morcombe, C. R.; Zilm, K. W. Chemical shift referencing in MAS solid state NMR, *J. Magn. Reson.* **2003**, *162*, 479.

(57) Bertani, P.; Raya, J.; Bechinger, B. ^{15}N chemical shift referencing in solid state NMR, *Solid State Nucl Magn Reson* **2014**, *61-62*, 15.

(58) Goddard, T. D., Kneller, D. G. SPARKY 3, Univ. of California, San Francisco, 2004.

(59) Lee, W.; Tonelli, M.; Markley, J. L. NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy, *Bioinform.* **2014**, *31*, 1325.

(60) Stevens, T. J.; Fogh, R. H.; Boucher, W.; Higman, V. A.; Eisenmenger, F.; Bardiaux, B.; Van Rossum, B.-J.; Oschkinat, H.; Laue, E. D. A software framework for analysing solid-state MAS NMR data, *J. Biomol. NMR* **2011**, *51*, 437.

(61) Cock, P. J. A.; Antao, T.; Chang, J. T.; Chapman, B. A.; Cox, C. J.; Dalke, A.; Friedberg, I.; Hamelryck, T.; Kauff, F.; Wilczynski, B.; De Hoon, M. J. L. Biopython: freely available Python tools for computational molecular biology and bioinformatics, *Bioinform.* **2009**, *25*, 1422.

(62) Hamelryck, T.; Manderick, B. PDB file parser and structure class implemented in Python, *Bioinform.* **2003**, *19*, 2308.

(63) Schwieters, C. D.; Bermejo, G. A.; Clore, G. M. Xplor-NIH for molecular structure determination from NMR and other data sources, *Protein Sci.* **2018**, *27*, 26.

(64) Schwieters, C. D.; Kuszewski, J. J.; Marius Clore, G. Using Xplor-NIH for NMR molecular structure determination, *Prog. Nucl. Magn. Reson. Spectrosc.* **2006**, *48*, 47.

(65) Shen, Y.; Delaglio, F.; Cornilescu, G.; Bax, A. TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts, *J. Biomol. NMR* **2009**, *44*, 213.

(66) Carneiro, M. G.; Koharudin, L. M. I.; Ban, D.; Sabo, T. M.; Trigo-Mourino, P.; Mazur, A.; Griesinger, C.; Gronenborn, A. M.; Lee, D. Sampling of glycan-bound conformers by the anti-HIV lectin Oscillatoria agardhii agglutinin in the absence of sugar, *Angew. Chem., Int. Ed.* **2015**, *54*, 6462.

(67) Götz, A. W.; Williamson, M. J.; Xu, D.; Poole, D.; Le Grand, S.; Walker, R. C. Routine microsecond molecular dynamics simulations with AMBER on GPUs. 1. Generalized Born, *J. Chem. Theory Comput.* **2012**, *8*, 1542.

(68) Nguyen, H.; Roe, D. R.; Simmerling, C. Improved generalized Born solvent model parameters for protein simulations, *J. Chem. Theory Comput.* **2013**, *9*, 2020.

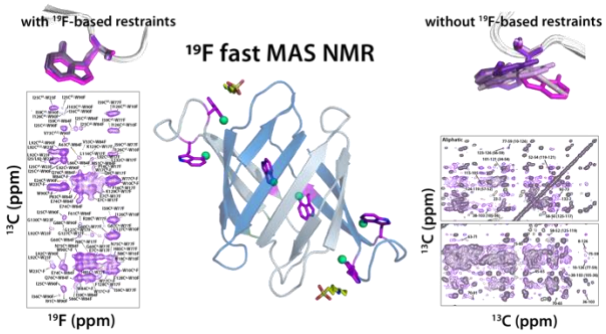
(69) Frishman, D.; Argos, P. Knowledge-based protein secondary structure assignment, *Proteins* **1995**, *23*, 566.

(70) Heinig, M.; Frishman, D. STRIDE: a web server for secondary structure assignment from known atomic coordinates of proteins, *Nucleic Acids Res.* **2004**, *32*, W500.

(71) Humphrey, W.; Dalke, A.; Schulten, K. VMD: visual molecular dynamics, *J Mol Graph* **1996**, *14*, 33.

(72) The PyMOL Molecular Graphics System v. 2.0 (Schrödinger, L., 2000).

TOC Graphic


with ¹⁹F-based restraints
¹⁹F fast MAS NMR
without ¹⁹F-based restraints
¹³C (ppm)
¹⁹F (ppm)
¹³C (ppm)
¹³C (ppm)