Structure of fully protonated proteins by proton-detected magic-angle spinning NMR

Loren B. Andreas[a], Kristaps Jaudzems[a], Jan Stanek[a], Daniela Lalli[a], Andrea Bertarello[a], Tanguy Le Marchand[a], Diane Cala-De Paepe[a], Svetlana Kotelovica[b], Inara Akopjana[b], Benno Knott[c], Sebastian Wegner[c], Frank Engelke[c], Anne Lesage[a], Lyndon Emsley[a,d], Kaspars Tars[b], Torsten Herrmann[a], and Guido Pintacuda[a,1]

[a]Centre de Résonance Magnétique Nucléaire à Très Hauts Champs, Institut des Sciences Analytiques (UMR 5280 – CNRS, Ecole Normale Supérieure de Lyon, Université Claude Bernard Lyon 1), Université de Lyon, 69100 Villeurbanne, France; [b]Biomedical Research and Study Centre, LV-1067 Riga, Latvia; [c]Bruker Biospin, 76287 Rheinstetten, Germany; and [d]Institut des Sciences et Ingénierie Chimiques, Ecole Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland

Edited by Adriaan Bax, National Institutes of Health, Bethesda, MD, and approved June 21, 2016 (received for review February 18, 2016)

Protein structure determination by proton-detected magic-angle spinning (MAS) NMR has focused on highly deuterated samples, in which only a small number of protons are introduced and observation of signals from side chains is extremely limited. Here, we show in two fully protonated proteins that, at 100-kHz MAS and above, spectral resolution is high enough to detect resolved correlations from amide and side-chain protons of all residue types, and to reliably measure a dense network of ^{1}H-^{1}H proximities that define a protein structure. The high data quality allowed the correct identification of internuclear distance restraints encoded in 3D spectra with automated data analysis, resulting in accurate, unbiased, and fast structure determination. Additionally, we find that narrower proton resonance lines, longer coherence lifetimes, and improved magnetization transfer offset the reduced sample size at 100-kHz spinning and above. Less than 2 weeks of experiment time and a single 0.5-mg sample was sufficient for the acquisition of all data necessary for backbone and side-chain resonance assignment and unsupervised structure determination. We expect the technique to pave the way for atomic-resolution structure analysis applicable to a wide range of proteins.

NMR spectroscopy | magic-angle spinning | protein structures | proton detection | viral nucleocapsids

Despite tremendous progress in the analysis of biomolecular samples over the last two decades (1–7), routine application of magic-angle spinning (MAS) NMR in biology is still limited by the inherently low sensitivity. The direct detection of proton resonances is a straightforward way to counter this problem, but entails a trade-off with resolution due to the strong homonuclear dipolar interactions among proton nuclei. High-resolution proton-detected methods were first demonstrated with modest spinning frequencies by today's standards (~10 kHz) and relied on a reduction of ^{1}H-^{1}H couplings by high levels of dilution with deuterium, typically perdeuteration, and complete (8, 9) or partial (10–12) protonation at exchangeable sites. The need for narrow proton resonances without such extreme levels of deuteration has motivated a continuous technological development, resulting in a dramatic increase in the available spinning frequency (13–20).

At MAS frequencies of 40–60 kHz, deuteration and 100% reprotonation at exchangeable sites, primarily amide protons, result in resolved and sensitive spectra, similar in quality to the case of higher dilution levels and lower spinning frequencies (21–23). This opens the way to rapid sequential assignment of backbone resonances (24–27), as well as to the unambiguous measurement of detailed structural and dynamical parameters (28–32). A further increase in the MAS frequency to 100 kHz allows resonance assignment (20), a structure determination of a model protein (16), and interaction studies (15) with as little as 0.5 mg of sample.

However, a high deuteration level severely limits observation of side-chain signals, which are essential for the determination of a protein structure at high resolution. The redundancy of information intrinsic to spectra of side-chain protonated proteins, which leads to a mutually supportive network of distance restraints, is also crucial to an unbiased and robust spectral analysis by unsupervised algorithms (33–36).

One approach to access side-chain proton resonances is the fractional labeling of side chains with ^{2}H and ^{1}H, which has the advantage of protonation at many sites, but with a dramatic reduction in sensitivity and potential loss of resolution due to the presence of multiple isotopomers (37–39). These problems have limited the application of these strategies for protein structure determination.

Alternatively, ILV-methyl labeling with suitable precursors has been used for structure determinations (16, 28, 40), and other tailored labeling approaches have also been proposed (41). However, these methods still introduce only a limited set of side-chain protons. Furthermore, they are implemented with deuterated growth media, which can reduce or even eliminate protein expression. Moreover, complete amide reprotonation in the interior of the protein can be problematic for systems that lack a refolding protocol (42), such as investigated in the present study. Extremely narrow aliphatic lines were reported for leucine residues in a stereo-array isotope-labeled (SAIL) protein at 80-kHz MAS (43). Unfortunately, the high cost of SAIL has prevented its widespread application.

All of these drawbacks are overcome if resolved side-chain proton resonances are available from a fully protonated sample. Fully protonated samples are by far the simplest to produce, and

Significance

Protein structure determination is key to the detailed description of many biological processes. The critical factor that would allow general application of magic-angle spinning (MAS) solid-state NMR to this end is improvement in sensitivity and resolution for as many nuclear spins as possible. This is achieved here with detection of resolved ^{1}H resonances in protonated proteins by increasing MAS rates to frequencies of 100 kHz and above. For large proteins and assemblies, ultrafast spinning narrows spectral resonances better than Brownian motion on which solution NMR relies, removing a fundamental barrier to the NMR study of large systems. This is exploited here to determine the de novo structure of a 28-kDa protein dimer in a 2.5-MDa viral capsid assembly.

Author contributions: L.B.A., K.J., J.S., D.L., A.L., L.E., K.T., T.H., and G.P. designed research; L.B.A., K.J., J.S., D.L., A.B., T.L.M., D.C.-D.P., S.K., I.A., K.T., T.H., and G.P. performed research; D.C.-D.P., S.K., I.A., B.K., S.W., and F.E. contributed new reagents/analytic tools; L.B.A., K.J., J.S., D.L., A.B., T.L.M., T.H., and G.P. analyzed data; and L.B.A., K.J., J.S., D.L., A.B., A.L., L.E., K.T., T.H., and G.P. wrote the paper.

The authors declare no conflict of interest.

Data deposition: The atomic coordinates and restraints have been deposited in the Protein Data Bank, www.wwpdb.org/ (PDB ID codes 5JXV and 5JZR), and chemical shifts have been deposited in the Biological Magnetic Resonance Data Bank, www.bmrb.wisc.edu/ (accession codes 30094 and 30088).

[1]To whom correspondence should be addressed. Email: Guido.Pintacuda@ens-lyon.fr.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1602248113/-/DCSupplemental.

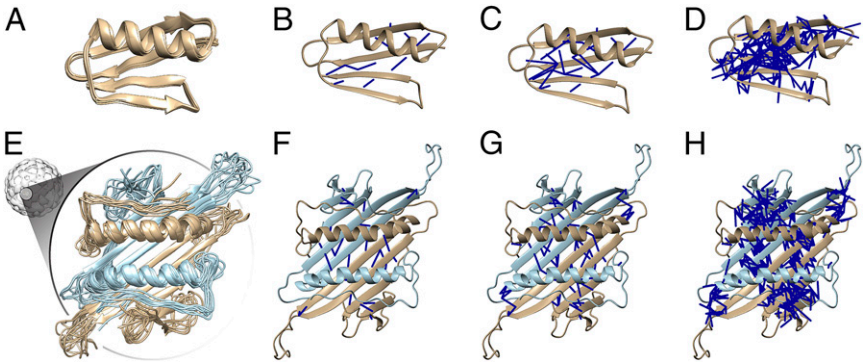

Fig. 1. MAS NMR structures of GB1 (PDB ID code 5JXV; *Top*) and AP205CP (PDB ID code 5JZR; *Bottom*). Ribbon diagram of the 10 lowest-energy conformers for GB1 (*A*) and dimeric AP205CP (*E*), with monomers colored in cyan and tan. The approximate location of 1 of the 90 dimers in the cryo-EM electron density is also illustrated (53). The long-range proton–proton contacts measured in this study are depicted as dark blue lines onto the lowest-energy conformers of the two proteins: contacts between amide protons (*B* and *F*), contacts among amide and ILV-methyl–labeled protons (*C* and *G*), and all proton–proton contacts (*D* and *H*).

their effective use could be applied to a much wider array of molecules including biomolecules labeled in mammalian systems (44). However, although amide proton assignment is possible for fully protonated proteins above 40-kHz MAS (45–48), proton resonances remain significantly dipolar broadened at 40–60 kHz, limiting the applicability of this spinning regime for side-chain assignment and structure determination (41, 45, 49–51).

Here, we investigate the resolution at an increased MAS frequency of 100 kHz and above, and find that it enables rapid structure determination in fully protonated proteins. To our knowledge, we present the first two examples of such structure determinations, the small model protein GB1 (52) and the de novo determination of *Acinetobacter* phage 205 (AP205) coat protein (AP205CP), a 28-kDa dimer in a 2.5-MDa viral capsid assembly (53). In both cases, we used a single 0.5-mg sample of uniformly ^{13}C,^{15}N–labeled protein for backbone and side-chain resonance assignment, and for collection of a dense network of proton–proton distance restraints. Each structure was calculated from data acquired in less than 2 weeks, and making use of unsupervised analysis algorithms as implemented in UNIO (54).

Results and Discussion

NMR Structures. Fig. 1 shows structural ensembles for microcrystalline GB1 and sedimented microcrystalline nucleocapsids of AP205CP, determined by solid-state NMR with MAS at frequencies of 100 kHz and above.

GB1 is a small and well-characterized globular protein of 56 residues. The bundle of NMR conformers calculated here (Fig. 1*A*) has a backbone heavy-atom root-mean-square deviation (rmsd) of 0.48 Å, an all-heavy-atom rmsd of 1.04 Å, and deviates from the X-ray structure (PDB ID code 2QMT) by 1.45 Å (see full statistics in Table S1).

AP205CP forms a homodimer of 2 × 130 residues, which is the basic subunit of the icosahedral AP205 capsid comprised of 90 dimers (53). The bacteriophage encodes for four genes, which are packed as single-stranded RNA (ssRNA) in a 28- to 30-nm icosahedral protein capsid. Because of its size, the AP205 capsid

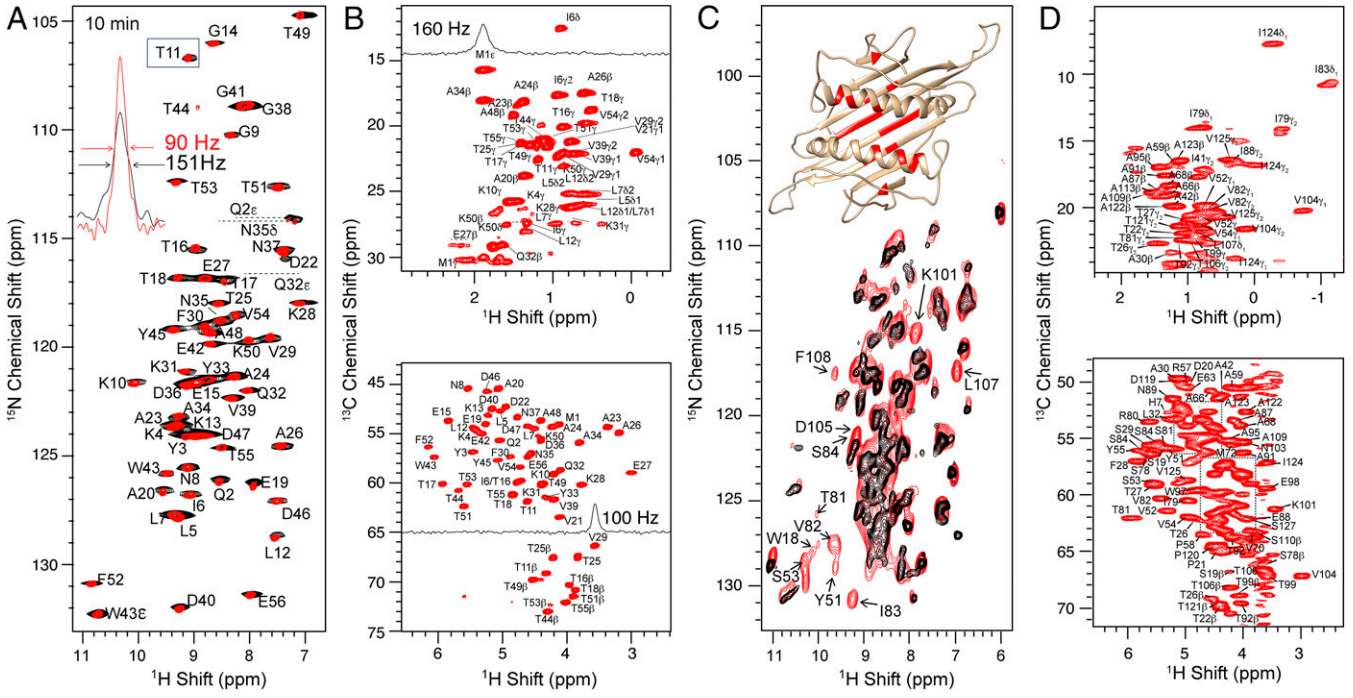

Fig. 2. High-resolution ^{1}H-detected spectra recorded on fully protonated microcrystalline GB1 and AP205CP. (*A*) ^{15}N-^{1}H CP-HSQC of GB1 at 60 kHz (black), and at 111.111-kHz MAS (red). (*B*) Hα-Cα (*Bottom*) and methyl regions (*Top*) of a ^{13}C-^{1}H CP-HSQC of GB1 at 111.111-kHz MAS. (*C*) ^{15}N-^{1}H CP-HSQC of AP205CP acquired on a fully protonated sample at 100-kHz MAS (red), and on a perdeuterated sample, exchanged in 100% H_2O, at 60-kHz MAS (black). The exchange-protected residues are labeled in the spectrum and colored on the NMR structure in the *Inset*. (*D*) Hα-Cα (*Bottom*) and methyl regions (*Top*) of a ^{13}C-^{1}H CP-HSQC of AP205CP at 100-kHz MAS.

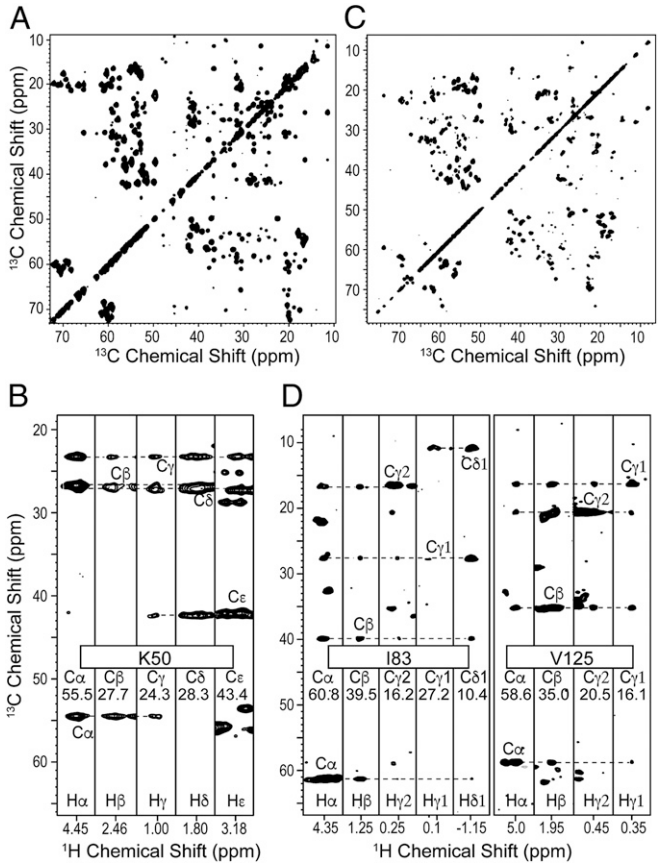

Fig. 3. (*A* and *C*) ^{13}C-^{13}C projection and (*B* and *D*) selected strips of the (H)CCH spectrum of GB1 (*A* and *B*) and AP205CP (*C* and *D*). GB1 and AP205CP spectra were run at a MAS frequency of 111 or 100 kHz, respectively, with WALTZ-16 mixing applied for 15 ms at 27.8 kHz or 14.4 ms at 25 kHz of radiofrequency field.

is inaccessible to solution NMR. Its crystals can be obtained easily in many different conditions, but despite our best efforts they do not diffract beyond 15-Å resolution, and similarly the cryo-EM structure (53) is resolved to only about 20 Å. However, sedimented microcrystals of AP205 nucleocapsids provide resolved MAS NMR spectra (27). The bundle of NMR conformers of the AP205CP dimer calculated here (Fig. 1*E*) has a backbone heavy-atom rmsd of 1.23 Å and an all-heavy-atom rmsd of 1.84 Å over structured regions (Table S1). The structure reveals a chain-intertwined dimer, typical for all ssRNA phages, with an extended β-sheet formed from both monomers and a pair of long helices engaged in a multitude of intermolecular contacts. The NMR structure is in agreement with the X-ray structure of an assembly-deficient AP205CP mutant dimer (PDB ID code 5FS4) to a backbone heavy-atom rmsd of 2.35 Å.

Each of the two structures is held together by a dense network of long-range (between residues i and j, $|i-j|>4$) ^{1}H-^{1}H contacts, also shown in Fig. 1, which encode distances up to about 5.5 Å. This amounts to 236 and 410 meaningful long-range contacts for GB1 and AP205CP, respectively, including 104 intermolecular contacts that define the dimer interface of AP205CP (Table S1). For these two structures, the total information content is described by ∼13.7 and 5.3 meaningful, nonredundant distance restraints per residue, approaching the criteria established for solution NMR determinations at high resolution (55). For AP205CP, the restraints per residue increase to 6.2 by considering only the assigned residues, and to 7.7 by considering only residues in structured regions.

Previous structural approaches in ^{1}H-detected MAS NMR relied on extensive deuteration, which resulted in the observation of contacts only among amide protons and ILV-methyl sites (16, 23, 28, 31, 56). However, amide–amide contacts (Fig. 1 *B* and *F*), or even contacts between ILV-methyl and amide protons (Fig. 1 *C* and *G*), represent a small number of the total, the large majority of which involve side-chain protons (Fig. 1 *D* and *H*). Although in principle a small set of long-range restraints could define a high-quality structure, a large set indicates a high-quality determination in the presence of limited precision and assignment ambiguity, which are typical of many systems of interest. As shown in Fig. 1, a large set of proton–proton contacts can be effectively acquired and automatically assigned in fully protonated proteins. This set includes side-chain protons, which are essential to constrain secondary structure elements (helix–sheet contacts in GB1, and helix–sheet and helix–helix contacts in AP205CP). The key for obtaining these sets of distance constraints lies in the possibility of directly recording and assigning resolved resonances for proton sites throughout the backbone and side chains of the two proteins.

Resolution of ^{1}H-Detected Spectra. GB1 is a 6.2-kDa model microcrystalline system known for its high spectral quality in solid-state NMR (52). Fig. 2 *A* and *B* shows its amide and aliphatic 2D correlation spectra [cross-polarization heteronuclear single-quantum correlation (CP-HSQC)] at 60 and 111 kHz, which at the faster spinning condition demonstrate amide proton linewidths averaging about 100 Hz (0.1 ppm at 1 GHz) and aliphatic and aromatic proton linewidths of 100–200 Hz (0.1–0.2 ppm).

The 28-kDa AP205CP dimer represents a challenging case of biological interest. The sample exhibits increased heterogeneity, manifested in linewidths of about 150–200 Hz (0.15–0.20 ppm) for amide, Hα, and methyl protons, at 100-kHz MAS (Fig. 2 *C* and *D*). A successful structure calculation in the presence of this inhomogeneity suggests that the method will be applicable to a range of biological targets of similar spectral quality, such as membrane-embedded proteins, fibrils, and macromolecular assemblies (27, 57, 58).

Fig. 2*C* also compares the amide spectrum of AP205CP to that of a perdeuterated sample purified in H_2O. Although most amide protons exchange during purification, as indicated by the observation of a cross-peak in both spectra, several weak or missing signals for the perdeuterated sample indicate incomplete exchange, a phenomenon observed for a variety of proteins (24, 59). For AP205CP, amide groups engaged in hydrogen bonds in the middle of the two α-helices, in the innermost β-strands, or at the dimer interface, are inaccessible to proton exchange. In general, nonexchangeable protons lie in the most structured regions of a protein, where the density of potential internuclear contacts is higher. The possibility of studying samples directly expressed in 100% H_2O is an advancement that extends ^{1}H-detected structural determinations to key exchange-protected regions.

Backbone and Side-Chain Resonance Assignment. ^{1}H, ^{13}C, and ^{15}N backbone and side-chain resonances were assigned using a suite of proton detected 3D spectra correlating backbone and side-chain chemical shifts.

For GB1, $^{1}H^{N}$, ^{13}Cα, ^{13}C′, and ^{15}N backbone resonances were first assigned using both ^{13}C-based interresidue matching (27), amide ^{15}N matching (48), and automated analysis using the UNIO-MATCH algorithm (60, 61). These assignments were then extended by measurement of the Hα shifts using a (H)NCAHA spectrum (46). For AP205CP, previous backbone $^{1}H^{N}$, ^{13}Cα, ^{13}Cβ, ^{13}C′, and ^{15}N assignments (27) were manually extended to 78% of the sequence and to the assignment of Hα resonances with a set of experiments linking Cα/Hα pairs in sequential residues through ^{15}N or ^{13}C′ resonances.

Next, Hα and Cα chemical shifts were used as anchors to propagate resonance assignment from backbone to side chains with a (H)CCH spectrum (62). This was implemented here with a CP-based sequence and WALTZ-16 mixing (63), applied at a nutation frequency of one-quarter of the rotor frequency (25 or 27.8 kHz), to induce isotropic mixing of ^{13}C magnetization. This spectrum correlates multiple ^{13}C resonances of an aliphatic side chain to each of its side-chain ^{1}H resonances, resulting in the robust identification of the ^{1}H and ^{13}C shifts. This is in contrast with ILV-labeled samples, in which it is difficult to unambiguously

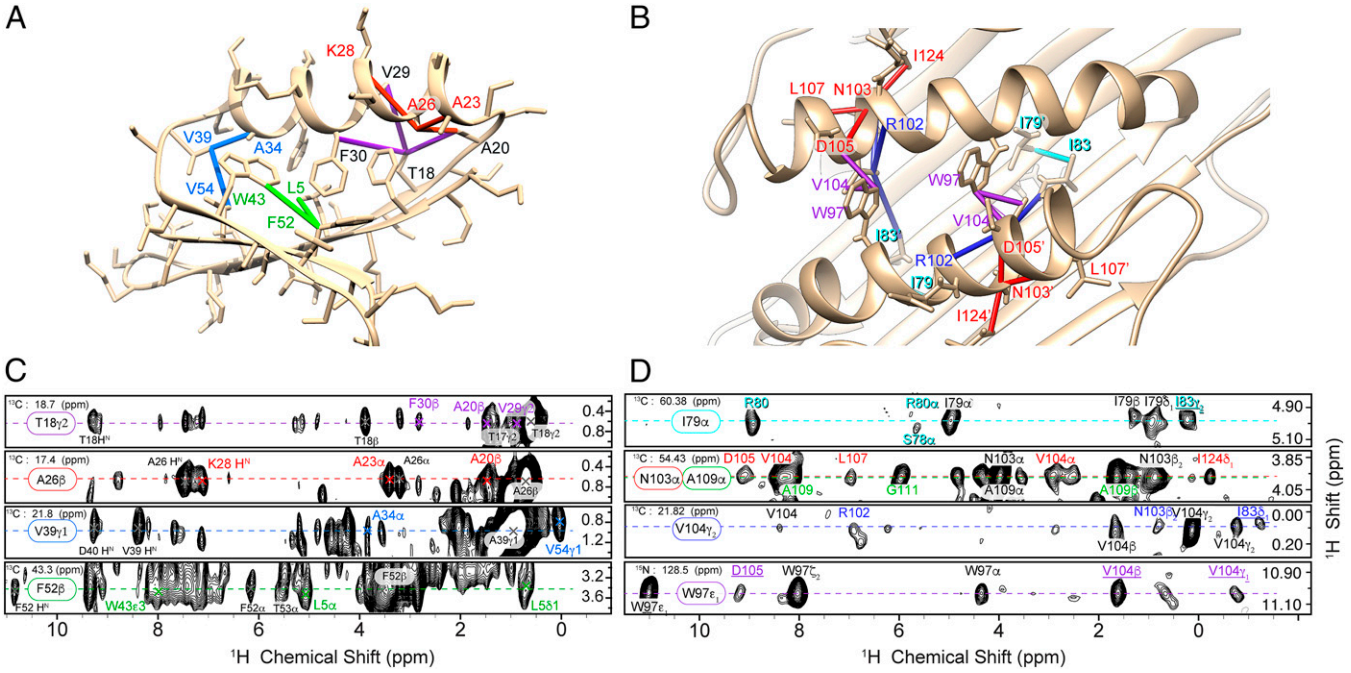

Fig. 4. (*A* and *B*) Representative restraints from the (H)CHH spectra are displayed on the lowest-energy NMR structures of GB1 (*A*) and AP205CP (*B*). The color of these restraints indicates the corresponding strip from the spectra shown in *C* and *D*. In the case of AP205CP (*D*), the labels of intermolecular cross-peaks are underlined. The RFDR mixing time was 0.5 ms for GB1 and 1.0 ms for AP205CP.

connect a single methyl proton at the end of a side chain to an amide ^{1}H resonance (in a ^{1}H-detected approach) or to a backbone ^{13}C resonance (in a ^{13}C-detected approach). Fig. 3 shows the ^{13}C-^{13}C planes from the 3D (H)CCH spectrum for the two proteins, and strips corresponding to the ^{13}C-^{1}H resonances of selected residues.

Overall, for GB1, the completeness of assignment of proton resonances was 94.4%, and the all-atom completeness was 85.9%, with the missing resonances being primarily aromatics. For AP205CP, the completeness of assignment of proton resonances was 64.9%, and the all-atom completeness was 62.6%.

^{1}H-^{1}H Contacts and Structure Calculation. For protein fold determination, 3D radiofrequency-driven recoupling (RFDR) spectra (64) were recorded that directly probe ^{1}H-^{1}H proximities. These spectra contain two ^{1}H dimensions and therefore benefit doubly from the narrow backbone and side-chain proton resonances. In this specific implementation, ^{1}H-^{1}H contacts were resolved using the shift of ^{15}N, aliphatic ^{13}C, or aromatic ^{13}C, and acquired in 72 h for GB1 and 103 h for AP205CP.

For each of these spectra, signals were manually identified and converted into unassigned peak lists, containing the frequency coordinates and the intensity of each signal. For GB1, the NMR structure was calculated with the UNIO software package using as inputs these unrefined peak lists, the assigned backbone and side-chain chemical shifts, and backbone dihedral angles predicted from the chemical shifts using TALOS+ (65). The standard unsupervised protocol of iterative cross-peak assignment, conversion into distance restraints, and structure calculation was applied as implemented in the program UNIO-CANDID (66). The chemical-shift–based assignment tolerances were set to the corresponding experimental linewidths, namely 0.15 and 0.4 ppm for proton and heavy-atom dimensions, respectively. Despite the fivefold increase in the proton tolerances compared with typically used solution NMR parameters, the average initial assignment ambiguity per peak of ∼16 remained modest. Not surprisingly, due to the high quality of the input data, an efficient network-anchored assignment of the contacts (66) resulted in a defined fold of the protein already in the first cycle of the iterative protocol. This is a crucial criterion certifying the reliability of a result in an unsupervised data analysis run (35).

AP205CP represents a particular challenge for structure determination due to the larger size and the dimeric protein topology. Using assignment tolerances corresponding to the experimental linewidths, the initial chemical-shift–based assignment ambiguity per peak was ∼56, which corresponds to a computational ambiguity of 112, due to the need to distinguish between intrasubunit and intersubunit contacts in the dimer. To improve convergence, the UNIO-CANDID protocol described for GB1 was supplemented with four ^{1}H-^{1}H distance restraints and 27 hydrogen bond restraints. Namely, two intermolecular helix–helix restraints (Val104 Hγ1-Trp97 Hε1 and Asp105 Hα-Trp97 Hε1), one intermolecular helix–strand restraint (Val104 Hγ1-Ile83 Hδ1), and one intramolecular helix–loop restraint (Asn103 Hα-Ile124 Hδ1) were applied. These spectrally unambiguous distance restraints were manually identified based on unique chemical shifts of involved side-chain ^{1}H nuclei. Two helix–helix contacts were entered as intermolecular because they cannot be satisfied for a monomer structure, whereas the calculation converges to the correct fold when the other two manually entered restraints are defined as ambiguous, i.e., as intermolecular or intramolecular (*Supporting Information*). Based on the observed chemical shifts that indicated β-sheet secondary structure, and on the observation of cross-strand Hα-Hα, H^{N}-H^{N}, and Hα-H^{N} RFDR contacts, 6 intermolecular and 21 intramolecular hydrogen bonds were imposed in the calculation between β-strands as detailed in Fig. S1. By using these additional restraints, a defined fold for the dimer was found after the first iteration of the automated UNIO-CANDID cross-peak assignment ensuring a reliable final result (*Supporting Information*).

It should be emphasized that relative alignment of two AP205CP monomers could be determined from the NMR data without the need for mixed isotope labeling that is often applied to filter intramolecular contacts (67). This is particularly important because there is no protocol for disassembly and reassembly of the AP205CP complex that would be needed for such an approach.

Statistics of the experimental restraints and structure calculation are summarized in Tables S1–S3, and further details can be found in *Supporting Information*. Although it was not possible to

unambiguously identify any interdimer contacts from the present 3D RFDR spectra, the determination of the AP205CP dimer structure paves the way to a future joint analysis of the NMR data with the cryo-EM density maps (53) to assemble 90 dimers into an atomic-level description of the global capsid architecture.

Sensitivity of ^{1}H-Detected Spectra. The high-resolution spectra demonstrated in Figs. 2–4 were acquired with no compromise in sensitivity compared with the previous state of the art at 60-kHz MAS in 1.3-mm rotors. An increase in the spinning rate requires reduced sample dimensions. This provides increased detection sensitivity associated with improved inductive coupling of smaller coils (almost a factor of 2 with respect to 1.3-mm rotors) but invariably entails a reduction in the sample volume (here by a factor of 4–5 with respect to 1.3-mm rotors) (18, 68). Combined, we therefore expect a theoretical 2.2- to 2.7-fold loss in the sensitivity of a single-pulse spectrum. Surprisingly, we measured approximately equal sensitivity for ^{15}N-^{1}H and ^{13}C-^{1}H CP-HSQC spectra in a 0.7- and a 1.3-mm probe, as shown by the comparison of the total area in the 1D proton spectra of Figs. S2 and S3. This observation points to an improvement in other factors, such as radiofrequency (RF) homogeneity, matching of RF field profiles along the rotor axis, and probe electronics.

Additionally, loss in sample volume is offset by narrower proton lines, longer coherence lifetimes, and improved transfer efficiencies (69). For GB1, the amide ^{1}H refocused transverse coherence lifetime (T_2') at ∼111-kHz MAS increased to 4.6 ms compared with 1.8 ms at the previous state-of-the-art of 60-kHz MAS. This corresponds to a nonrefocusable linewidth of about 70 Hz, indicating that about one-half of the contribution to the linewidths is still homogeneous. We observed improvement in ^{1}H linewidths by about a factor of 1.5–1.8 for GB1 and dramatic improvements in T_2' of ^{13}C′, ^{13}Cα, and ^{15}N, when increasing the MAS rates from 60 to ∼111 kHz (*Supporting Information*). This results in improved *J*-transfer efficiency and therefore higher sensitivity or reduced experiment time for sequences that depend on the T_2', producing for GB1 an additional sensitivity gain of about 2.1 for an out-and-back Cα → Cβ transfer, and 1.6 for a full C′ → Cα transfer (48). Finally, for the RFDR-based 3Ds that encode two proton dimensions, there is also a sensitivity benefit that depends on the indirect sampling. For example, when sampling extends to $1.5T_2^*$ at 111-kHz MAS, as in the case of GB1, this amounts to a factor of about 1.3–1.5, as detailed in Fig. S4. Gains due to linewidth and T_2' are modest for AP205CP for backbone resonances due to significant inhomogeneity in this sample. The exact sensitivity gain depends on the details of the sample and the pulse sequence, but is more significant for experiments with many transfers on which the majority of the data acquisition time is spent. This bodes well for the extension to higher dimensional spectra for fully protonated samples once higher sensitivity is available from cryoprobes or higher field magnets.

Conclusions

In conclusion, to our knowledge, we have presented here the first examples of protein structure determination by MAS NMR using fully protonated samples and ^{1}H detection. The narrowed ^{1}H linewidth, which arises due to fast sample spinning, allows the successful application of efficient protocols for resonance assignment and automated structure determination involving backbone and side-chain protons from all residue types. Notably, we have demonstrated that backbone and aliphatic side-chain assignment and high-resolution protein structure determination can be achieved using less than 2 weeks of instrument time, a single 0.5-mg sample, and rapid and unsupervised analysis of internuclear contacts.

The straightforward ^{1}H-based structure determination under fast sample spinning removes a fundamental barrier to the efficient NMR study of large systems, because the high resolution operated by MAS does not rely on stochastic Brownian motion, and therefore is independent of the size of the molecule, as demonstrated here in the de novo atomic-resolution structure of AP205CP in an intact 2.5-MDa capsid assembly. This approach overcomes the need for deuteration, and not only allows access to side-chain protons, but also circumvents potential problems of incomplete proton backexchange and difficulty of protein expression in deuterated media.

We expect the approach to enable structure determination for a wide range of molecules such as membrane proteins and macromolecular complexes.

Methods

Sample Preparation. Uniformly ^{13}C,^{15}N-labeled GB1 and AP205CP samples were expressed in *Escherichia coli*, purified, and precipitated as described previously (26, 52). A full description is provided in *Supporting Information*.

Structure Calculation. For both proteins, the standard UNIO protocol was used, which consists of seven cycles of cross-peak assignment, conversion into a meaningful, nonredundant set of distance restraints, and structure calculation by simulated annealing. In the second and subsequent cycles, the intermediate protein structures were used to guide the process of cross-peak assignment. Further details are discussed in *Supporting Information*, and calculation statistics are reported in Tables S2 and S3.

NMR Spectroscopy. All spectra were recorded at $\omega_{0H}/2\pi = 1$ GHz and a MAS rate of 100 kHz (AP205CP) or 111.111 kHz (GB1) using a Bruker 0.7-mm HCN probe. Spectrometer settings, as well as acquisition and processing parameters specific for each 2D and 3D spectrum, are discussed in *Supporting Information* and summarized in Tables S4 and S5. Simulations of the isotropic ^{13}C-^{13}C mixing were performed using the software package SIMPSON (70) and are reported in Fig. S5.

ACKNOWLEDGMENTS. We thank Lénaïc Leroux for technical assistance with the NMR spectrometers. We acknowledge financial support from CNRS (IR-RMN FR3050 and Fondation pour la Chimie des Substances Naturelles), from the People Programme of the European Union's Seventh Framework Programme (FP7) (FP7-PEOPLE-2012-ITN 317127 "pNMR"), and from the European Research Council under the European Union's Horizon 2020 Research and Innovation Programme (Grant 648974 "P-MEM-NMR"). L.B.A., K.J., and J.S. are supported by three Marie Curie incoming fellowships (Research Executive Agency Grant Agreements 624918 "MEM-MAS," 661175 "virus-DNP-NMR," and 661799 "COMPLEX-FAST-MAS"). J.S. received support from a European Molecular Biology Organization fellowship (ALTF 1506-2014, LTFCOFUND2013, and GA-2013-609409).

1. Castellani F, et al. (2002) Structure of a protein determined by solid-state magic-angle-spinning NMR spectroscopy. *Nature* 420(6911):98–102.
2. Lange A, et al. (2006) Toxin-induced conformational changes in a potassium channel revealed by solid-state NMR. *Nature* 440(7086):959–962.
3. Wasmer C, et al. (2008) Amyloid fibrils of the HET-s(218-289) prion form a beta solenoid with a triangular hydrophobic core. *Science* 319(5869):1523–1526.
4. Jehle S, et al. (2010) Solid-state NMR and SAXS studies provide a structural basis for the activation of alphaB-crystallin oligomers. *Nat Struct Mol Biol* 17(9):1037–1042.
5. Loquet A, et al. (2012) Atomic model of the type III secretion system needle. *Nature* 486(7402):276–279.
6. Wang S, et al. (2013) Solid-state NMR spectroscopy structure determination of a lipid-embedded heptahelical membrane protein. *Nat Methods* 10(10):1007–1012.
7. Xiao Y, et al. (2015) Aβ(1-42) fibril structure illuminates self-recognition and replication of amyloid in Alzheimer's disease. *Nat Struct Mol Biol* 22(6):499–505.
8. Chevelkov V, et al. (2003) ^{1}H detection in MAS solid-state NMR spectroscopy of biomacromolecules employing pulsed field gradients for residual solvent suppression. *J Am Chem Soc* 125(26):7788–7789.
9. Paulson EK, et al. (2003) Sensitive high resolution inverse detection NMR spectroscopy of proteins in the solid state. *J Am Chem Soc* 125(51):15831–15836.
10. Zheng L, Fishbein KW, Griffin RG, Herzfeld J (1993) 2-Dimensional solid-state ^{1}H-NMR and proton-exchange. *J Am Chem Soc* 115(14):6254–6261.
11. Chevelkov V, Rehbein K, Diehl A, Reif B (2006) Ultrahigh resolution in proton solid-state NMR spectroscopy at high levels of deuteration. *Angew Chem Int Ed Engl* 45(23):3878–3881.
12. Akbey U, et al. (2010) Optimum levels of exchangeable protons in perdeuterated proteins for proton detection in MAS solid-state NMR spectroscopy. *J Biomol NMR* 46(1):67–73.
13. Kobayashi T, et al. (2013) Study of intermolecular interactions in the corrole matrix by solid-state NMR under 100 kHz MAS and theoretical calculations. *Angew Chem Int Ed Engl* 52(52):14108–14111.

14. Nishiyama Y, Malon M, Ishii Y, Ramamoorthy A (2014) 3D $^{15}N/^{15}N/^{1}H$ chemical shift correlation experiment utilizing an RFDR-based $^{1}H/^{1}H$ mixing period at 100 kHz MAS. *J Magn Reson* 244:1–5.
15. Lamley JM, et al. (2014) Solid-state NMR of a protein in a precipitated complex with a full-length antibody. *J Am Chem Soc* 136(48):16800–16806.
16. Agarwal V, et al. (2014) De novo 3D structure determination from sub-milligram protein samples by solid-state 100 kHz MAS NMR spectroscopy. *Angew Chem Int Ed Engl* 53(45):12253–12256.
17. Zhang R, Pandey MK, Nishiyama Y, Ramamoorthy A (2015) A novel high-resolution and sensitivity-enhanced three-dimensional solid-state NMR experiment under ultrafast magic angle spinning conditions. *Sci Rep* 5:11810.
18. Andreas LB, Le Marchand T, Jaudzems K, Pintacuda G (2015) High-resolution proton-detected NMR of proteins at very fast MAS. *J Magn Reson* 253:36–49.
19. Böckmann A, Ernst M, Meier BH (2015) Spinning proteins, the faster, the better? *J Magn Reson* 253:71–79.
20. Penzel S, et al. (2015) Protein resonance assignment at MAS frequencies approaching 100 kHz: A quantitative comparison of J-coupling and dipolar-coupling-based transfer methods. *J Biomol NMR* 63(2):165–186.
21. Zhou DH, et al. (2007) Solid-state protein-structure determination with proton-detected triple-resonance 3D magic-angle-spinning NMR spectroscopy. *Angew Chem Int Ed Engl* 46(44):8380–8383.
22. Lewandowski JR, et al. (2011) Enhanced resolution and coherence lifetimes in the solid-state NMR spectroscopy of perdeuterated proteins under ultrafast magic-angle spinning. *J Chem Phys Lett.* 2(17):2205–2211.
23. Knight MJ, et al. (2011) Fast resonance assignment and fold determination of human superoxide dismutase by high-resolution proton-detected solid-state MAS NMR spectroscopy. *Angew Chem Int Ed Engl* 50(49):11697–11701.
24. Ward ME, et al. (2011) Proton-detected solid-state NMR reveals intramembrane polar networks in a seven-helical transmembrane protein proteorhodopsin. *J Am Chem Soc* 133(43):17434–17443.
25. Zhou DH, et al. (2012) Solid-state NMR analysis of membrane proteins and protein aggregates by proton detected spectroscopy. *J Biomol NMR* 54(3):291–305.
26. Barbet-Massin E, et al. (2013) Out-and-back ^{13}C-^{13}C scalar transfers in protein resonance assignment by proton-detected solid-state NMR under ultra-fast MAS. *J Biomol NMR* 56(4):379–386.
27. Barbet-Massin E, et al. (2014) Rapid proton-detected NMR assignment for proteins with fast magic angle spinning. *J Am Chem Soc* 136(35):12489–12497.
28. Huber M, et al. (2011) A proton-detected 4D solid-state NMR experiment for protein structure determination. *ChemPhysChem* 12(5):915–918.
29. Knight MJ, et al. (2012) Structure and backbone dynamics of a microcrystalline metalloprotein by solid-state NMR. *Proc Natl Acad Sci USA* 109(28):11095–11100.
30. Ma P, et al. (2014) Probing transient conformational states of proteins by solid-state $R_{1\rho}$ relaxation-dispersion NMR spectroscopy. *Angew Chem Int Ed Engl* 53(17):4312–4317.
31. Linser R, et al. (2014) Solid-state NMR structure determination from diagonal-compensated, sparsely nonuniform-sampled 4D proton-proton restraints. *J Am Chem Soc* 136(31): 11002–11010.
32. Good DB, et al. (2014) Conformational dynamics of a seven transmembrane helical protein *Anabaena* Sensory Rhodopsin probed by solid-state NMR. *J Am Chem Soc* 136(7):2833–2842.
33. Guerry P, Herrmann T (2011) Advances in automated NMR protein structure determination. *Q Rev Biophys* 44(3):257–309.
34. Rosato A, et al. (2012) Blind testing of routine, fully automated determination of protein structures from NMR data. *Structure* 20(2):227–236.
35. Guerry P, Duong VD, Herrmann T (2015) CASD-NMR 2: Robust and accurate unsupervised analysis of raw NOESY spectra and protein structure determination with UNIO. *J Biomol NMR* 62(4):473–480.
36. Buchner L, Güntert P (2015) Systematic evaluation of combined automated NOE assignment and structure calculation with CYANA. *J Biomol NMR* 62(1):81–95.
37. Asami S, Schmieder P, Reif B (2010) High resolution ^{1}H-detected solid-state NMR spectroscopy of protein aliphatic resonances: Access to tertiary structure information. *J Am Chem Soc* 132(43):15133–15135.
38. Asami S, Szekely K, Schanda P, Meier BH, Reif B (2012) Optimal degree of protonation for ^{1}H detection of aliphatic sites in randomly deuterated proteins as a function of the MAS frequency. *J Biomol NMR* 54(2):155–168.
39. Mance D, et al. (2015) An efficient labelling approach to harness backbone and side-chain protons in ^{1}H-detected solid-state NMR spectroscopy. *Angew Chem Int Ed Engl* 54(52):15799–15803.
40. Andreas LB, et al. (2015) Structure and mechanism of the influenza A M218-60 dimer of dimers. *J Am Chem Soc* 137(47):14877–14886.
41. Sinnige T, Daniëls M, Baldus M, Weingarth M (2014) Proton clouds to measure long-range contacts between nonexchangeable side chain protons in solid-state NMR. *J Am Chem Soc* 136(12):4452–4455.
42. Lian LY, Middleton DA (2001) Labelling approaches for protein structural studies by solution-state and solid-state NMR. *Prog Nucl Magn Reson Spectrosc* 39:171–190.
43. Wang S, et al. (2015) Nano-mole scale side-chain signal assignment by ^{1}H-detected protein solid-state NMR by ultra-fast magic-angle spinning and stereo-array isotope labeling. *PLoS One* 10(4):e0122714.
44. Chow WY, et al. (2014) NMR spectroscopy of native and in vitro tissues implicates polyADP ribose in biomineralization. *Science* 344(6185):742–746.
45. Marchetti A, et al. (2012) Backbone assignment of fully protonated solid proteins by ^{1}H detection and ultrafast magic-angle-spinning NMR spectroscopy. *Angew Chem Int Ed Engl* 51(43):10756–10759.
46. Zhou DH, et al. (2007) Proton-detected solid-state NMR spectroscopy of fully protonated proteins at 40 kHz magic-angle spinning. *J Am Chem Soc* 129(38): 11791–11801.
47. Wang S, et al. (2015) Nano-mole scale sequential signal assignment by ^{1}H-detected protein solid-state NMR. *Chem Commun (Camb)* 51(81):15055–15058.
48. Andreas LB, et al. (2015) Protein residue linking in a single spectrum for magic-angle spinning NMR assignment. *J Biomol NMR* 62(3):253–261.
49. Barbet-Massin E, et al. (2014) Insights into the structure and dynamics of measles virus nucleocapsids by ^{1}H-detected solid-state NMR. *Biophys J* 107(4):941–946.
50. Vasa SK, Rovo P, Giller K, Becker S, Linser R (2016) Access to aliphatic protons as reporters in non-deuterated proteins by solid-state NMR. *Phys Chem Chem Phys* 18:8359–8363.
51. Xiang S, Biernat J, Mandelkow E, Becker S, Linser R (2016) Backbone assignment for minimal protein amounts of low structural homogeneity in the absence of deuteration. *Chem Commun (Camb)* 52(21):4002–4005.
52. Franks WT, et al. (2005) Magic-angle spinning solid-state NMR spectroscopy of the β1 immunoglobulin binding domain of protein G (GB1): ^{15}N and ^{13}C chemical shift assignments and conformational analysis. *J Am Chem Soc* 127(35):12291–12305.
53. van den Worm SH, Koning RI, Warmenhoven HJ, Koerten HK, van Duin J (2006) Cryo electron microscopy reconstructions of the *Leviviridae* unveil the densest icosahedral RNA packing possible. *J Mol Biol* 363(4):858–865.
54. Guerry P, Herrmann T (2012) Comprehensive automation for NMR structure determination of proteins. *Methods Mol Biol* 831:429–451.
55. Montelione GT, et al. (2013) Recommendations of the wwPDB NMR Validation Task Force. *Structure* 21(9):1563–1570.
56. Linser R, Bardiaux B, Higman V, Fink U, Reif B (2011) Structure calculation from unambiguous long-range amide and methyl ^{1}H-^{1}H distance restraints for a microcrystalline protein with MAS solid-state NMR spectroscopy. *J Am Chem Soc* 133(15):5905–5912.
57. Linser R, et al. (2011) Proton-detected solid-state NMR spectroscopy of fibrillar and membrane proteins. *Angew Chem Int Ed Engl* 50(19):4508–4512.
58. Eddy MT, et al. (2015) Lipid bilayer-bound conformation of an integral membrane beta barrel protein by multidimensional MAS NMR. *J Biomol NMR* 61(3-4):299–310.
59. Chill JH, Louis JM, Miller C, Bax A (2006) NMR study of the tetrameric KcsA potassium channel in detergent micelles. *Protein Sci* 15:684–698.
60. Volk J, Herrmann T, Wüthrich K (2008) Automated sequence-specific protein NMR assignment using the memetic algorithm MATCH. *J Biomol NMR* 41(3):127–138.
61. Dutta SK, et al. (2015) APSY-NMR for protein backbone assignment in high-throughput structural biology. *J Biomol NMR* 61(1):47–53.
62. Agarwal V, Reif B (2008) Residual methyl protonation in perdeuterated proteins for multi-dimensional correlation experiments in MAS solid-state NMR spectroscopy. *J Magn Reson* 194(1):16–24.
63. Shaka AJ, Frenkiel T, Freeman R (1983) NMR broadband decoupling with low radio-frequency power. *J Magn Reson* 52:159–163.
64. Paulson EK, et al. (2003) High-sensitivity observation of dipolar exchange and NOEs between exchangeable protons in proteins by 3D solid-state NMR spectroscopy. *J Am Chem Soc* 125(47):14222–14223.
65. Shen Y, Delaglio F, Cornilescu G, Bax A (2009) TALOS+: A hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J Biomol NMR* 44(4):213–223.
66. Herrmann T, Güntert P, Wüthrich K (2002) Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. *J Mol Biol* 319(1):209–227.
67. Loquet A, Habenstein B, Lange A (2013) Structural investigations of molecular machines by solid-state NMR. *Acc Chem Res* 46(9):2070–2079.
68. Peck TL, Magin RL, Lauterbur PC (1995) Design and analysis of microcoils for NMR microscopy. *J Magn Reson B* 108(2):114–124.
69. Nieuwkoop AJ, et al. (2015) Sensitivity and resolution of proton detected spectra of a deuterated protein at 40 and 60 kHz magic-angle-spinning. *J Biomol NMR* 61(2):161–171.
70. Bak M, Rasmussen JT, Nielsen NC (2000) SIMPSON: A general simulation program for solid-state NMR spectroscopy. *J Magn Reson* 147(2):296–330.
71. Güntert P, Mumenthaler C, Wüthrich K (1997) Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J Mol Biol* 273(1):283–298.
72. O'Donoghue SI, King GF, Nilges M (1996) Calculation of symmetric multimer structures from NMR data using a priori knowledge of the monomer structure, co-monomer restraints, and interface mapping: The case of leucine zippers. *J Biomol NMR* 8(2):193–206.
73. Klovins J, Overbeek GP, van den Worm SHE, Ackermann HW, van Duin J (2002) Nucleotide sequence of a ssRNA phage from *Acinetobacter*: Kinship to coliphages. *J Gen Virol* 83:1523–1533.
74. Bertini I, et al. (2011) Solid-state NMR of proteins sedimented by ultracentrifugation. *Proc Natl Acad Sci USA* 108(26):10396–10399.
75. Böckmann A, et al. (2009) Characterization of different water pools in solid-state NMR protein samples. *J Biomol NMR* 45(3):319–327.
76. Zhou DH, Rienstra CM (2008) High-performance solvent suppression for proton detected solid-state NMR. *J Magn Reson* 192(1):167–172.
77. Bax A, Clore GM, Gronenborn AM (1990) ^{1}H-^{1}H correlation via isotropic mixing of ^{13}C magnetization, a new 3-dimensional approach for assigning ^{1}H and ^{13}C spectra of ^{13}C-enriched proteins. *J Magn Reson* 88:425–431.
78. Baldus M, Meier BH (1996) Total correlation spectroscopy in the solid state. The use of scalar couplings to determine the through-bond connectivity. *J Magn Reson A* 121(1):65–69.
79. Hardy EH, Verel R, Meier BH (2001) Fast MAS total through-bond correlation spectroscopy. *J Magn Reson* 148(2):459–464.
80. Bennett AE, Ok JH, Griffin RG, Vega S (1992) Chemical-shift correlation spectroscopy in rotating solids: Radio frequency-driven dipolar recoupling and longitudinal exchange. *J Chem Phys* 96(11):8624–8627.