Backbone and side-chain proton NMR assignment in fully protonated proteins: microcrystals, sedimented assemblies, and amyloid fibrils
Jan Stanek, Loren Andreas, Kristaps Jaudzems, Diane Cala, Daniela Lalli, Andrea Bertarello, Tobias Schubeis, Inara Akopjana, Svetlana Kotelovica, Kaspars Tars, et al.

To cite this version:

HAL Id: hal-01567784
https://hal.archives-ouvertes.fr/hal-01567784
Submitted on 24 Jul 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives| 4.0 International License
Backbone and side-chain proton NMR assignment in fully protonated proteins: microcrystals, sedimented assemblies, and amyloid fibrils


Abstract: We demonstrate sensitive detection of alpha protons of fully protonated proteins by solid-state NMR with 100-111 kHz magic-angle spinning (MAS). The excellent resolution in the Cα-Hα plane is demonstrated for 5 proteins, including microcrystals, a sedimented complex, a capsid and amyloid fibrils. A set of 3D spectra based on a Cα-Hα detection block was developed and applied for the sequence-specific backbone and aliphatic side-chain resonance assignment using only 500 µg of sample. These developments accelerate structural studies of biomolecular assemblies available in submilligram quantities without the need of protein deuteration.

Proton detection[1] at MAS rates up to 60 kHz has revolutionized solid-state NMR investigations of microcrystalline samples,[2] insoluble proteins,[3] and membrane proteins[4, 5] by improving sensitivity and reducing the amount of sample required. Up to 60 kHz MAS, optimal proton resolution requires extensive dilution of the proton dipolar interaction network. This can be achieved by complete[6a, 6b, 6c] or partial[6d, 6e, 6f] deuteration, where 1H can be reintroduced by chemical exchange with water.

Due to this dilution requirement, only amide protons have been used for detection in sequence specific protein assignment strategies. Analogous to the case in solution, resonance assignment proceeds efficiently using a series of 3D spectra with common 15N and 1H dimensions, and with a third dimension encoding 13C, 13C’ or 13Cβ of either the same or the preceding residue, allowing linking of sequential 15N-1H pairs.[6a, 6b, 6c]

The upper limit of MAS frequency was recently increased to 111 kHz with the introduction of 0.7 and 0.8 mm probes,[6a, 6h] allowing higher sensitivity in multidimensional experiments[6i] and a further reduction of sample volume.[6j] Most importantly, this extends the methodology to fully protonated proteins, allowing the detection of alpha and side-chain protons at a resolution comparable with those of the amide groups.[6k, 6l, 6m]

Here we show a strategy for 1H, 13C, and 15N resonance assignment for fully protonated biological samples based on a suite of 3D spectra sharing Hα-Cα pairs as sensitive and resolved probes at >100 kHz MAS and high magnetic field. We additionally use the same Hα-Cα resonances as a starting point for propagation of assignments throughout side-chains. We expect this approach to be of critical importance for the determination of protein structures by solid-state NMR, since it removes the requirement for deuteration, which previously limited 1H detection-based strategies due to problems with expression yields, protein refolding and amide proton exchange.[6o] Also it provides a reliable method for sequence specific assignment of aliphatic protons, which are ideal reporters of tertiary and quaternary structure.[6p]

Less than 500 µg of five fully protonated 13C,15N-labelled protein samples was packed into a 0.7 mm MAS rotor. The cross-polarization heteronuclear single quantum correlation (CP-HSQC) spectra acquired at 100-111 kHz MAS show resolved 13Cα-1Hα resonances (Fig. 1) for a range of proteins in various aggregation states. They include small crystalline proteins GB1[7a] and MNEI[7b] and biological assemblies with a significantly higher degree of complexity, such as the Acinetobacter phage 205 nucleocapsid (AP205),[7c] and the prion domain of HET-s in its amyloid fibrillar form. In AP205, 90 nearly equivalent dimers of the coat protein (AP205CP) assemble into a 2.5 MDA particle of icosahedral symmetry.[7d] ρβ sliding clamp is a 81 kDa circular dimer which was immobilized by sedimentation.[7e] Lastly, the prion domain of HET-s (218-289) aggregates into fibrils with a parallel in-register amyloid core showing a triangular β-solenoid fold.[7f]

Each of these resolved 13C-1H dipolar-based correlations was obtained in approx. 30 minutes, providing a sensitive fingerprint at site-specific resolution without the need for sample deuteration. The 13Cα-1Hα region shows similar dispersion in chemical shift for proton and heteronuclei (approx. 3 and 15 ppm) as that observed in 1H-1H correlation spectra (Fig. S1), and in contrast to 1H-15N spectra, allows detection of prolines. In the case of GB1, we found that 1Hα and 1Hβ line widths are the same within experimental error, averaging to about 105 Hz

[a] Dr. J. Stanek, Dr. L. B. Andreas, Dr. K. Jaudzems, Dr. D. Cala, Dr. D. Lalli, A. Bertarello, T. Schubeis, Dr. G. Pintacuda Centre de RMN à Très Hauts Champs, Institut des Sciences Analytiques (UMR 5280 - CNRS, ENS Lyon, UCB Lyon 1), Université de Lyon, 5 rue de la Doua, 69100 Villeurbanne, France E-mail: guido.pintacuda@ens-lyon.fr
[b] Dr. I. Akopjana, Dr. S. Kotelovica, Pr. K. Tars Biomedical Research and Study Centre, Rātsupilē 1, LV1067, Riga, Latvia
[c] Dr. A. Pica, Dr. S. Leone, Pr. D. Picone Department of Chemical Sciences, University of Naples Federico II, Via Cintia, I-80126, Naples, Italy
[d] Dr. Z.-Q. Xu, Pr. N. E. Dixon Institute of Chemistry & Biology of Membranes & Nanoobjects (UMR 5248 CBMN - CNRS, University of Bordeaux, Bordeaux INP), All. Geoffroy Saint-Hilaire, 33600 Pessac, France
[e] Dr. S. Sauge Institut de Biochimie et de Génétique Cellulaire (UMR 5095, CNRS - Université de Bordeaux), 33077 Bordeaux, France Supporting information for this article is given via a link at the end of the document.
Figure 1. Hα-Cα regions of CP 13C-HSQC spectra of microcrystalline proteins GB1 (A) and MNEI (B), AP205 coat protein dimer assembling a nucleocapsid (C), sedimented β2 sliding clamp (D) and HET-s construct (218-289) forming amyloid fibrils (E). 1D traces with representative proton line widths are reported for isolated peaks. Spectra were acquired on a 1 GHz NMR spectrometer at MAS rates of 111 (A, B, D) and 100 kHz (C, E).

(Fig. S2), and comparable line widths were also observed for many isolated signals in the other four samples. This is nearly as narrow as the best-case 70 Hz line width that corresponds to the inherent heterogeneity of microcrystalline GB1 (see SI).

Notably, a significant improvement in resolution is observed for these correlations with respect to the previously accessible MAS standard of 60 kHz due to a significant increase in the bulk 1H and 13C coherence lifetime T2. For the β2 clamp and HET-s samples, lifetimes of 3.0 and 2.5 (for 1H) and 22 and 28 ms (for 13C) were reached at 100 kHz MAS, respectively, which are noticeably shorter than for microcrystalline proteins. Nevertheless, estimated residual homogeneous 1H line broadening is only 70 Hz (for GB1), 110 Hz (for AP205CP), 105 Hz (for β2 clamp) and 130 Hz (for HET-s) in the fastest MAS regime currently available, and enables one to observe resolved Hα-Cα correlations for all proteins studied here.

To exploit the favorable resolution and sensitivity of the Hα resonances, we developed a suite of 3D experiments for sequential backbone and side-chain assignment that employ Cα-Hα correlation as the detection scheme. We used two complementary 3D pulse sequences, (H)N(CO)CAH and (H)NCAH, to record inter- and intra-residue correlations, respectively, as shown schematically in Fig. 2A-B (also more detailed in Fig. S3). These allow to sequentially link Hα-Cα resonance pairs by matching a common 15N chemical shift.

An additional pair of 3D spectra, (H)CO(N)CAH and (H)COCAH was also developed to link sequential Cα–Hα pairs based on correlations to a common 13C′ chemical shift.

Figure 2. Simplified pulse sequence diagrams with indirect (13C and 15N) chemical shift evolution periods denoted with arrows (A). The irradiation schemes employed for these five experiments incorporate from 2 to 4 heteronuclear dipolar transfers (between H-N, H-C, N-C), and 13C-15N and 13C-13C′ J-based transfers. The long lifetimes at above 100 kHz MAS are thus the key to high sensitivity in these sophisticated experiments involving multiple dimensions and transfer steps (Fig. S5). Schematic illustration of magnetization flow in the proposed 5 correlation experiments (B). Representative 2D cross-sections from 15N-(blue and red), 13C′-linked (cyan and orange) Hα-detected spectra for HET-s (C), GB1 (D) and AP205CP (E), along with the respective intra-residue Cβ-Cα-Hα correlations for AP205 (F).
Simultaneous matching of $^{15}$N and $^{13}$C chemical shifts resolves ambiguities and ensures high reliability of the sequential assignment. Among backbone nuclei, $^{15}$N and $^{13}$C were found to have the most favorable resolution (chemical shift dispersion with respect to line width), and thus provide the least ambiguous sequential matching. In contrast to the sibling strategy developed for solution-state NMR, residual water signal is efficiently suppressed and has no adverse impact on sensitivity of the H$\alpha$-detected experiments. A fifth pulse sequence, (H)CA(C)BCAH (bottom panel in Fig. 2A), was designed to provide access to $^{13}$C$\beta$ chemical shifts, which are of great relevance for determination of residue type and thus essential for the alignment of the linked stretches of H$\alpha$-C$\alpha$ pairs with the primary structure. Additionally, the simultaneous presence of two $^{13}$C dimensions provides great precision for the difference ($^{13}$C$\alpha$-$^{13}$C$\beta$), a parameter directly correlated to secondary structure and important during structure modeling.

For fully-protonated proteins, the experiments described above play an equivalent role as the suite of H$^0$-detected sequences for deuterated systems. As a particular advantage here, H$\alpha$-detection provides access to intraresidue $^{13}$C$\beta$ and $^{13}$C$\gamma$ shifts with significantly higher sensitivity with respect to H$^0$-detection, due to shorter coherence transfer pathways. Also, unlike H$^0$-detection, the proposed H$\alpha$-detected experiments enable the sequential walk over prolines, making them a method of choice for sequential assignment of proline-rich proteins. Proline resonances usually show higher intensity in the $^{13}$C-linking pair of spectra (Fig. 2E).

Representative strip plots showing the sequential walk for HET-s, GB1 and AP205 are shown in Fig. 2C-E. For the case of GB1, we obtained all expected correlations (Fig. S6) and for AP205CP we found 90, 98, 103 and 109 correlations in the 3D spectra (Fig. S7) and assigned in total 111 (85%) H$\alpha$-C$\alpha$ pairs of resonances. Missing peaks belong mostly to N-terminal residues, an internal loop (R34-N46) and the C-terminus that show increased dynamics and cannot be efficiently detected in dipolar-coupling based transfers. For HET-s amyloid fibrils, we observed 41 out of 44 intraresidue correlations in the $\beta$-strands forming the rigid fibril core, i.e. N226-A248 and T261-Y281 (Fig. S8) and 40 of 42 expected interresidue correlations in the (H)N(CO)CAH spectrum. The flexible loop, as reported for $^{13}$C-detected spectra, is not visible in CP-based experiments. The assigned CP-HSQC spectra of HET-s are shown in Fig. S9.

An alternative approach for H$\alpha$ assignment would employ H$\beta$-detection for sequential backbone assignment of $^{13}$C$\alpha$ and $^{15}$N shifts, followed by the use of an extra (H)NCAH to transfer the assignment to $\alpha$-protons. However, this relies on the uniqueness of $^{13}$C$\alpha$ and $^{15}$N shift pairs, the condition which is often met only for systems as simple as GB1. In AP205, for example, all five H$\alpha$-detected spectra were necessary to obtain assignment of $\alpha$-protons. Therefore, the set of spectra shown here represents the first general and stand-alone strategy for assignment of H$\alpha$-proton shifts in the solid-state.

Assignment of $\alpha$-protons opens the way to assign side-chain protons, which are sensitive probes of protein environment and interactions. In GB1 at 111 kHz MAS, side-chain $^1$H resonances demonstrate 100-150 Hz line widths, which is sufficient for their use in structural investigations. H$\alpha$-detected schemes can correlate $^1$H and $^{13}$C side-chain with backbone resonances using a shorter and thus more sensitive coherence transfer pathway compared to H$^0$-detected schemes. We incorporated $^{1}J_{\alpha\alpha}$-mediated $^{13}$C-$^{13}$C isotropic mixing into the H$\rightarrow$C$\rightarrow$H CP-HSQC. The resulting irradiation schemes, H(C)CH and (H)CH, are shown in Fig. S4, and incorporate WALTZ-16 $^{13}$C mixing analogous to established TOBSY schemes. High efficiency of $^{13}$C mixing based on relatively weak J-couplings results from long rotating-frame coherence lifetimes at 100-111 kHz spinning. For example, for GB1 T$_{1p}$ ($^{13}$C$\alpha$) = 130 ms is almost an order of magnitude longer compared to the time necessary to obtain a reasonably uniform exchange of magnetization between side-chain $^{13}$C nuclei (typically 10-15 ms). The two spectra encode aliphatic side-chain $^1$H or $^{13}$C spins of a residue to each directly bonded pair such as Hx-C$\alpha$, H$\beta$-C$\beta$, etc. (Fig. 3). Given that assignment of $\alpha$-resonances is available, these two spectra together enable to identify $^1$H and $^{13}$C chemical shifts further in the side-chains. The obtained values can be recursively used in the same way as the starting Hx-C$\alpha$ assignment to complete and/or confirm the resonance assignment in case of ambiguity or low intensity of some cross-peaks. In Fig. 3, six cross-sections are displayed at the positions of all H-C resonance pairs of isoleucine-124 of AP205CP and show a shared set of correlations to all aliphatic $^{13}$C or $^1$H resonances in this residue. One can appreciate the redundancy of information and mutual support for identification of side-chain signals, which stems from the fact that in principle, all intraresidue aliphatic $^{13}$C or $^1$H resonances are correlated to each directly bonded $^{13}$C-$^1$H spin pair. In AP205CP, the (H)CH (55h acquisition time) and H(C)CH (24h) resulted in assignment of 87% of side-chain $^1$H and $^{13}$C resonances for residues with assigned $^1$H shifts (see SI). In GB1, a single (H)CH (32.5h) resulted in 99.1% complete assignment of side-chain $^1$H and $^{13}$C resonances.

![Figure 3](image-url)
In conclusion, we demonstrated rapid side-chain $^1$H and $^{13}$C resonance assignment in fully-protonated proteins at very fast MAS rates (100-111 kHz). For proteins of about 20 kDa one week of experimental time and less than a milligram of sample are required, without the need of deuteration. The method can be widely applied not only to microcrystalline proteins, but also to viral nucleocapsids, amyloid fibrils and other protein assemblies or sediments. The accessibility of aliphatic $^1$H chemical shifts is of great importance for structure determination based on $^1$H-$^1$H distance constraints, accelerating solid-state NMR analysis for structural biology.

Acknowledgements

We acknowledge financial support from CNRS (IR-RMN FR3050 and Fondation pour la Chime des Substances Naturelles), from the People Programme of the European Union’s FP7 (FP7-PEOPLE-2012-ITN “n°317127 *pNMR*”), and from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (ERC-2015-CoG GA n°648974 to GP and ERC-2015-StG GA n°105945 to ALBA). KJ and JS are supported by three MC incoming fellowships (REA grant agreements n°624918 "MEM-MAS", n°661175 "virus-DNP-NMR", n°661799 "COMPLEX-FAST-MAS"). AP and SL are supported by Fondazione con il Sud (grant 2011-PDR-19).

Keywords: magic-angle spinning • proton detection • resonance assignment • solid-state NMR spectroscopy

Side-chain protons are all back: with new 111 kHz magic-angle spinning probes, high-resolution $^1$H-detected NMR spectroscopy of insoluble, crystalline or self-assembled protein aggregates is now feasible without replacing side-chain protons with deuterons. $\alpha$-Protons become sensitive and spectrally resolved NMR probes, which allow backbone and side-chain resonance assignment in about one week of experimental time for proteins of about 20 kDa.