Protein residue linking in a single spectrum for magic-angle spinning NMR assignment

Article Title

Protein residue linking in a single spectrum for magic-angle spinning NMR assignment

Journal Name

Journal of Biomolecular NMR

Authors

Loren B. Andreas,[1] Jan Stanek,[1] Tanguy Le Marchand,[1] Andrea Bertarello,[1] Diane Cala-De Paepe,[1] Daniela Lalli,[1] Magdaléna Krejčíková,[1,2] Camille Doyen,[1] Carl Öster,[1,3] Benno Knott,[4] Sebastian Wegner,[4] Frank Engelke,[4] Isabella C. Felli,[5] Roberta Pierattelli,[6] Nicholas E. Dixon,[6] Lyndon Emsley,[1,7] Torsten Herrmann,[1] Guido Pintacuda[1*]

Affiliations

[1] Centre de RMN à Très Hauts Champs, Institut des Sciences Analytiques (CNRS, ENS Lyon, UCB Lyon 1), Université de Lyon, 69100 Villeurbanne, France

[2] Centre for Structural Biology, Central European Institute of Technology (CEITEC), Masaryk University, 625 00 Brno, Czech Republic

[3] Department of Chemistry, University of Warwick, Coventry CV4 7AL, United Kingdom

[4] Bruker Biospin, 76287 Rheinstetten, Germany

[5] Department of Chemistry "Ugo Schiff" and Magnetic Resonance Center (CERM), University of Florence, 50019 Sesto Fiorentino, FI, Italy

[6] School of Chemistry, University of Wollongong, New South Wales 2522, Australia

[7] Department of Chemistry, Ecole Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland

Corresponding Author's email Address

*guido.pintacuda@ens-lyon.fr

Summary

Here we introduce a new pulse sequence for resonance assignment that halves the number of data sets required for sequential linking by directly correlating sequential amide resonances in a single diagonal-free spectrum. The method is demonstrated with both microcrystalline and sedimented deuterated proteins spinning at 60 and 111 kHz, and a fully protonated microcrystalline protein spinning at 111 kHz, with as little as 0.5 mg protein sample. We find that amide signals have a low chance of ambiguous linkage, which is further improved by linking in both forward and backward directions. The spectra obtained are amenable to automated resonance assignment using general-purpose software such as UNIO-MATCH.

Keywords

Magic-angle spinning; protein resonance assignment; proton detection; automation.

Nuclear magnetic resonance (NMR) accounts for a sizable number of all determined protein structures, with the majority of NMR structures determined from solution. For proteins in the solid state, determination of atomic resolution global structure by NMR often hinges on the ability to resolve protein resonances, which can be improved by narrowing lines and increasing polarization. With the recent development of both labeling schemes and of hardware capable of higher magic-angle spinning (MAS) frequencies now reaching 100 kHz (Kobayashi et al. 2013; Agarwal et al. 2014; Lamley et al. 2014; Nishiyama et al. 2014), proton-detected spectra (Reif and Griffin 2003; Chevelkov et al. 2006) can be recorded with high resolution and sensitivity analogous to the case in solution.

Sequence-specific assignment of resonances is a critical, and often time consuming step on the path to determination of site-specific protein structural and dynamical information. Failure to assign spectra is often the bottleneck to structure determination. Under MAS, sequence-specific resonance assignment is accomplished primarily by inter-residue matching with ^{13}C shifts, either using ^{13}C detection (Hong

1999; Rienstra et al. 2000; Böckmann et al. 2003) or more recently, ^{1}H detection (van Rossum et al. 2003; Knight et al. 2011; Linser et al. 2011; Ward et al. 2011; Marchetti et al. 2012; Zhou et al. 2012; Barbet-Massin et al. 2014; Xiang et al. 2014).

Resonance assignment with ^{1}H detection has matured into a methodology that significantly reduces the acquisition time while also improving confidence in the result. This approach requires pairs of spectra to establish inter-residue correlations, ideally making use of six spectra simultaneously correlating CO, CA and CB resonances to resolve degeneracies for sequential connectivities. Here we introduce a concept for resonance assignment that is enabled by very fast MAS, based on direct matching of amide resonances. It can halve the number of experiments required for linking residues and improves the process by providing double linking in both the forward and backward directions. The approach is demonstrated at up to 111 kHz MAS, and is used for rapid unambiguous sequential assignment on perdeuterated and 100% ^{1}H back-exchanged proteins using as little as 0.5 mg of sample in a 0.7 mm rotor.

In solution, introduction of multiple ^{15}N frequencies has been shown to greatly reduce the ambiguity in crowded spectra of disordered proteins (Liu et al. 2000; Panchal et al. 2001; Yoshimura et al. 2015), an approach that built upon other amide matching strategies (Weisemann et al. 1993; Matsuo et al. 1996; Bracken et al. 1997; Frueh et al. 2006). These pulse sequences rely on weak $^{3}J_{cc}$ scalar couplings, or require many scalar transfer steps. The efficiency is high for small molecules or disordered proteins, but generally prohibitively low for larger molecules in solution or in biological solids. For proteins in the solid state, through-space magnetization transfers have been applied at both slow (van Rossum et al. 2003) and fast spinning conditions (Nishiyama et al. 2014), but the resulting spectra contain both sequential and long-range cross-peaks and are therefore of limited use for resonance assignment.

Figure 1 shows two variants of a new 3D pulse sequence, (H)N(COCA)NH and (H)N(CACO)NH that correlate the amide ^{15}N-^{1}H pair of a central residue *i* with the amide ^{15}N of residue *i+1* and *i-1*, respectively. The sequence is enabled by high transfer efficiencies of typically 30 to 60% at each step

due to long coherence lifetimes under fast MAS and low power ^{1}H decoupling. These conditions also favor short recycle times (~1 s) with low risk of sample degradation. Five coherence transfers are used: one ^{13}C-^{13}C scalar based transfer, and 4 cross polarization (CP) transfers between ^{1}H and ^{15}N and between ^{13}C and ^{15}N. Without modifying the transfer pathway, the pulse sequence can evolve three additional periods, namely the chemical shifts of a starting H^N and of the bridging CO and CA nuclei enabling acquisition of high dimensional spectra.

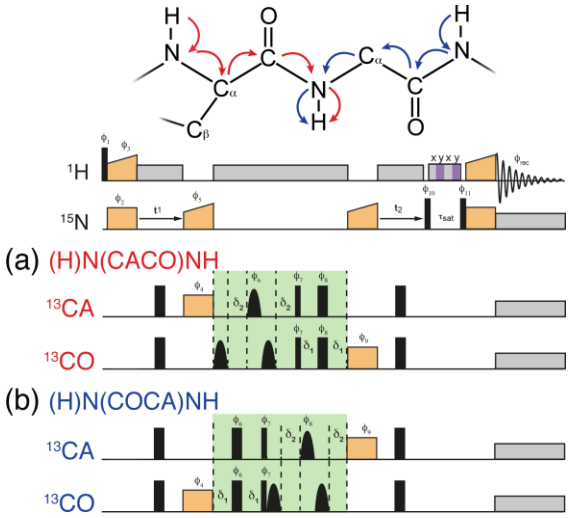

Fig. 1 Pulse sequence variants corresponding to the two 3D correlation spectra. Each is used to correlate an amide ^{1}H-^{15}N pair with either the (a) preceding or (b) subsequent amide ^{15}N. Heteronuclear transfers via cross polarization (CP) are shown in yellow, while the homonuclear ^{13}CA-^{13}CO transfer is highlighted in green. We set the carrier frequency of the ^{13}C channel to about 100 ppm, and set the offset of ^{13}C on resonance (with either ^{13}CO or ^{13}CA) for CP. The phase alignment was set at the end of the pulse for the CP before the ^{13}C-^{13}C transfer, and at the beginning of the pulse for the CP that follows the ^{13}C-^{13}C transfer. Bell shapes represent selective inversion pulses, here 350 μs Q3 pulses (Emsley and Bodenhausen 1992) with the frequency offset applied at the center of the range for either ^{13}CO or ^{13}CA (note that partial refocusing of deleterious CA-CB passive coupling can be obtained with selective CA refocusing pulses). Excitation and inversion pulses are shown as narrow and broad black rectangles, respectively. Grey rectangles represent heteronuclear decoupling. Water suppression (Zhou and Rienstra 2008) is shown as purple and grey rectangles on the ^{1}H channel, and was typically applied for 100 ms. Delays δ_1 and δ_2 were set to 3.5 and 3.0 ms, respectively. The phase cycle is ϕ_1=13, ϕ_2=1, ϕ_3=2, ϕ_4=1, ϕ_5=1133, ϕ_6=0000 0000 1111 1111, ϕ_7=1, ϕ_8=0000 1111, ϕ_9=1133, ϕ_{10}=1, ϕ_{11}=1, ϕ_{rec}=1313 3131 3131 1313

With five transfer steps, the overall experimental efficiency is the product of five individual transfer efficiencies and minimizing losses at each step becomes critically important. We measured the ^{1}H-^{15}N transfer efficiency by comparing the signal after two transfers (^{1}H-^{15}N-^{1}H) and after four (^{1}H-^{15}N-^{1}H-^{15}N-^{1}H) and assuming equal efficiency in each step. The ^{13}C-^{15}N transfer efficiency was similarly determined with an ^{1}H-^{15}N-^{13}C-^{15}N-^{1}H pulse sequence. The remaining efficiency was attributed to the ^{13}C-^{13}C transfer. The transfer efficiencies for microcrystalline proteins were 45% for ^{1}H-^{15}N and ^{15}N-^{1}H CP, 25 to 35% for ^{15}N-^{13}C and ^{13}C-^{15}N CP, and above 60% for the ^{13}C-^{13}C scalar transfer (Nieuwkoop et al. 2015) (see Figure S1a in the Electronic Supplementary Material). Such efficiency compares favorably with scalar based methods applied in solution, where magnetization is split during the N-CA and CA-N transfers leading to considerable signal losses and autocorrelation peaks.

In contrast to the case with scalar transfer, the new sequence results in spectra that are diagonal free, which decreases spectral crowding; they correlate the chemical shifts of a residue (*i*) and the ^{15}N shift of *only* the preceding (*i*-1) or the following (*i*+1) residue.

Where sensitivity is a limiting factor, we expect carbon based matching with CA and CO to remain the primary method of assignment due to a significant sensitivity advantage of about 2 to 4-fold. However, we found similar efficiency for the nitrogen matching experiments and the (HCA)CB(CACO)NH sequence used for CB matching. We therefore expect that when CA and CO matching are not sufficient for assignment, nitrogen matching will be a useful approach comparable or even better than CB matching.

We recorded amide correlation spectra of five proteins. In Table 1, we report a summary of the datasets acquired for the four proteins, as well as details of the resulting spectra. As a measure of the experimental efficiency, we compared the first point of each spectrum with the first point of the (H)NH spectrum (CP-HSQC), in the same way as was recently reported for carbon based matching spectra (Barbet-Massin et al. 2014).

In contrast to ^{13}C matching approaches, here each individual signal encodes a pair of nitrogen shifts, and a single spectrum is thus sufficient for linking amide resonances along the protein backbone. This is illustrated in Figure 2a, where sequential correlations are shown for microcrystalline SH3 (62 amino acids) (van Rossum et al. 2003; Barbet-Massin et al. 2014; Nieuwkoop et al. 2015). Here 48 cross-signals were observed, corresponding to 52 residues (98% of expected non-proline residues observed in MAS spectra).

Protein	Ubiquitin	SOD	SSB	GB1	SH3
Field (MHz ^{1}H)	800	700	1000	1000	700
Size (Residues)	76	153	178	56	62
Automatically Assigned Residues	38	81	not performed	55	not performed
(H)N(CACO)NH					
Sensitivity w.r.t. CP-HSQC	7%	6.3%	-	4.4%	-
Experiment time	36h	46h	-	7.5h	-
Peaks identified	41	99	-	54	-
Maximum ^{15}N indirect evolution time	8/10.2 ms	12.8 ms	-	8.3 ms	-
(H)N(COCA)NH					
Sensitivity w.r.t. CP-HSQC	6%	5.0%	>3%	4.5%	4.7%
Experiment time	36h	46h	92h	7.5h	69h
Peaks identified	43	90	not performed	54	48
Maximum ^{15}N indirect evolution time	8/10.2 ms	12.8 ms	5.1 ms	8.3 ms	17.9 ms

Table 1. Experimental parameters and summary of results for each of the 5 proteins investigated.

When both 3D datasets are recorded, the simultaneous analysis of signals from the pair builds triplets of ^{15}N chemical shifts. As illustrated in Figure 2b-c, the simultaneous matching of these triplets establishes connections spanning three residues. The alignment of stretches relies on two ^{15}N frequencies that are directly correlated in each spectrum. Comparable reliability of sequential

assignment would require two pairs of ^{13}C-based spectra, or double the set of experiments required here for resonance assignment. Additionally, matching ^{15}N shifts results in low peak overlap due to the excellent resolution ^{15}N. Of particular relevance is also the low correlation of sequential ^{15}N chemical shifts as compared with intra-residue correlation of CO, CA, and CB (see discussion below, and Figure S2 in the Electronic Supplementary Material).

Once the signals are identified by peak-picking, the sequence-specific assignment of resonances can be quickly obtained with automated protocols, such as those developed for solution NMR data and implemented in UNIO-MATCH (Volk et al. 2008; Dutta et al. 2015). In the case of microcrystalline ubiquitin (76 amino acids) (Igumenova et al. 2004), 41 and 43 peaks were identified in the (H)N(CACO)NH and (H)N(COCA)NH spectra, respectively. This corresponds to all of the 54 amide ^{15}N resonances that could be assigned using ^{13}C matching and ^{1}H-detected spectra (data not shown). Inter-residue linking was determined by UNIO-MATCH for 38 ^{15}N-^{1}H pairs. The resulting assignment shows little deviations with respect to that of non His-6 tagged microcrystals (Schanda et al. 2010). For the 153 amino acid microcrystalline human Cu^{I},Zn^{II}-superoxide dismutase (SOD) (Knight et al. 2012), 99 and 90 peaks were identified in the (H)N(CACO)NH and (H)N(COCA)NH spectra, respectively, each acquired in only 46 hours. Inter-residue linking was determined for 81 amide ^{15}N-^{1}H pairs using UNIO-MATCH out of 136 that are possible to assign with the extensive application of ^{13}C-based matching. The ^{15}N-^{15}N projection of the (H)N(CACO)NH spectrum of SOD is shown in Figure 3.

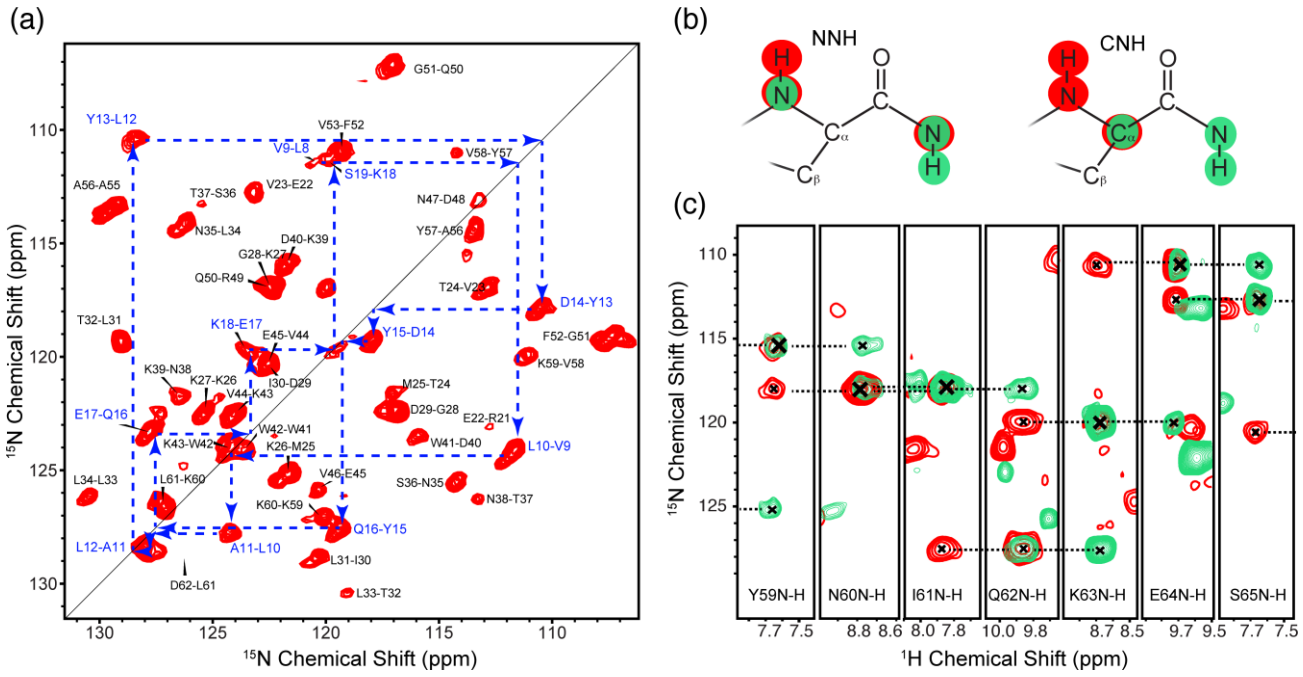

Fig. 2 Sequential linking using ^{15}N-^{15}N-^{1}H correlations in microcrystalline proteins at 60 kHz MAS. In (a), sequential connections are shown in the ^{15}N-^{15}N projection of the (H)N(COCA)NH recorded on chicken α spectrin SH3 microcrystals. Inter-residue correlations from residue L8 to residue S19 are connected with blue arrows. Recorded at $\omega_{0H}/2\pi$ of 700 MHz. In (b), the double linking in the pair of ^{15}N-^{15}N-^{1}H spectra is illustrated, as compared with single linking in a pair of ^{13}C-^{15}N-^{1}H spectra. In (c), doubling linking is shown with a superposition of (H)N(COCA)NH (red) and (H)N(CACO)NH (green) spectra of ubiquitin microcrystals at $\omega_{0H}/2\pi$ of 800 MHz.

At a spinning frequency of 111.111 kHz, we performed the assignment of the 56 amino acid protein GB1 (Franks et al. 2005) using a fully protonated sample and the (H)N(CACO)NH and (H)N(COCA)NH spectra (Figure 3b). Each of the amide correlation spectra were recorded in 7.5 hours using approximately 0.5 mg of sample packed in a 0.7 mm rotor, and all 54 expected cross-peaks were observed in each spectrum (only M1 was missing, while the number of cross-peaks is one less than the number of residues in the stretch of assignments due to the inter-residue transfer). In this case, a single (H)N(CACO)NH spectrum was sufficient to complete the assignment using UNIO-MATCH, however, there were several ambiguous sequential links, and the result therefore relied somewhat on chemical shift statistics. A reliable assignment was made with the two spectra.

For these three proteins, amide ^{15}N matching in two directions resulted in the unambiguous identification of sequential stretches of residues. For long stretches and highly complete data, correct assignment is possible, as determined using the GB1 data and synthetic and virtually complete input data for both ubiquitin and SOD. With less complete data, both correct residue placement and sequential linking can be done with the addition of datasets incorporating ^{13}CA or ^{13}CB shifts (e.g. (H)CANH or (HCA)CB(CA)NH spectra (Barbet-Massin et al. 2013)) because these shifts are particularly sensitive to residue type and secondary structure. This is particularly relevant when large portions of the protein (often flexible loops or tails) are not observed in the MAS spectra.

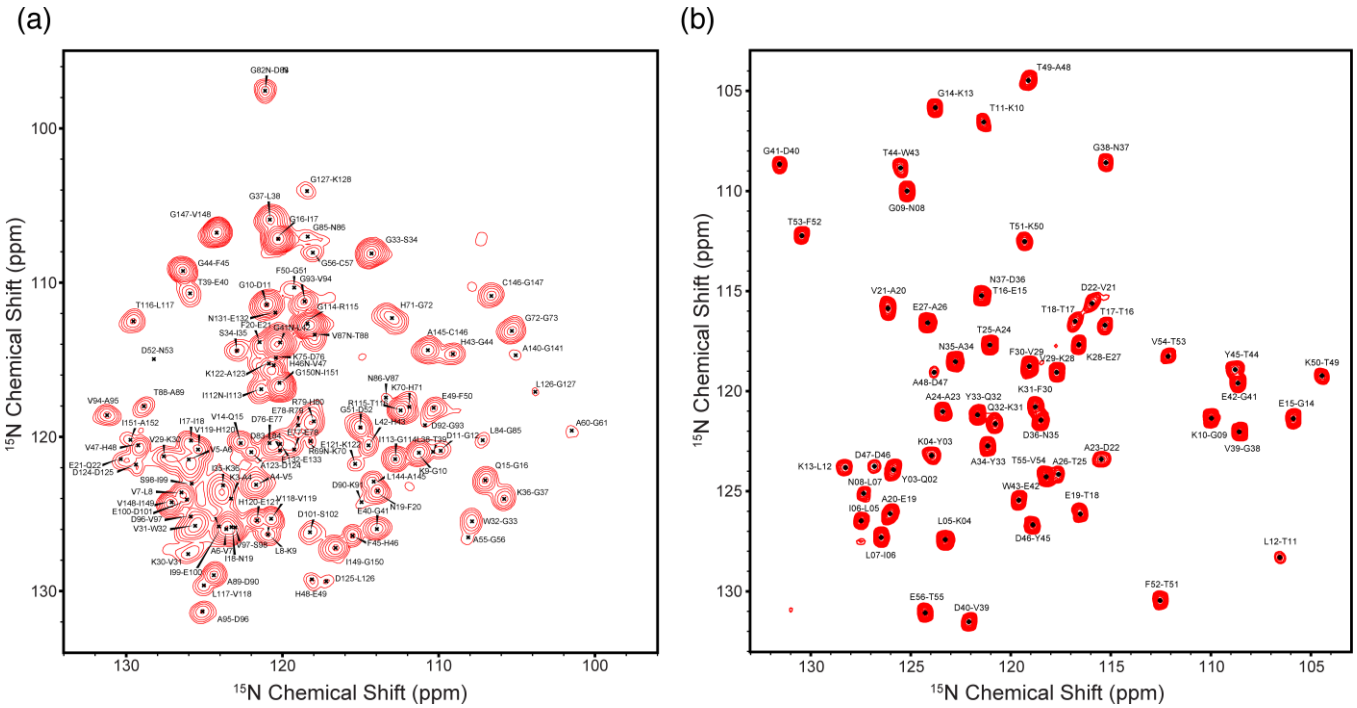

Fig. 3 Additional examples of (H)N(CACO)NH spectra showing the ^{15}N-^{15}N projection. In (a), sequential linking for microcrystalline SOD showing the 99 inter-residue cross-peaks that were identified. The spectrum was recorded at a MAS frequency of 60 kHz and $\omega_{0H}/2\pi$ =700 MHz. In (b) sequential linking in fully protonated GB1 showing all 55 expected cross-peaks (only M1 is missing). The spectrum was recorded at a MAS frequency of 111.111 kHz, and $\omega_{0H}/2\pi$ =1 GHz.

To further investigate the relative utility of matching amide nitrogen frequency compared with other types of nuclei, we analyzed the likelihood of peak overlap according to a simple model. We calculated

the likelihood of peak overlap as a function of linewidth given the chemical shift distributions in the BioMagResBank (bmrb.wis.edu), and using the average distribution of amino acids found in vertebrates (Further details are presented in the Electronic Supplementary Material). The vertebrate weighted chemical shift distributions are shown in Fig. 4a. The chemical shift axis scale was adjusted to correspond to the observed linewidths of SH3 (in ppm at 23.5 T) allowing a direct visual comparison of chemical shift dispersion for this protein. Figure 4b shows approximately the fraction of overlapping peaks according to the following calculations.

It was assumed that (i) the primary sequence is random, i.e. consecutive aminoacid types are uncorrelated, (ii) amino acid frequencies are as used for construction of weighted histograms, (iii) proline residues are absent, (iv) all peaks have equal line widths. The database-derived weighted histograms were assumed to approximate the probability density function $p(\delta)$ for a peak to possess a chemical shift δ. The probability of observing other peaks with chemical shifts within a linewidth (LW) is given by the integral

$$P_d = \int_{d-LW/2}^{d+LW/2} p(d')dd' \quad (1)$$

Due to lack of sequential correlation, the probability of chemical shift overlap is simply a product of the probabilities $p(\delta)$ and P_δ:

$$P_{ovlp} = p(d')dd' \times P_d = p(d')dd' \int_{d-LW/2}^{d+LW/2} p(d'')dd'' \quad (2)$$

The average number of possible assignments of a sequential correlation peak, including the correct one, for a sequence of N amino acids, is as follows:

$$a = 1 + (N-2)\int_{-\infty}^{+\infty} P_{ovlp} \quad (3)$$

and is position independent. Thus, the overall complexity of assignment can be approximated by a product

$$A = \prod_{k=1}^{N-1} a_k = a^{N-1} \quad (4)$$

Clearly the assignment possibilities grow rapidly with an increase in the probability for ambiguous linking at each step.

The overlap index, i.e. the fraction of peaks that overlap on average with a given correlation peak was calculated as the global integral of P_{ovlp} from Eq. 2. The dependence of overlap index on line width for the 5 kinds of nuclei considered here is shown in Fig. 4b. As expected, it exhibits approximately linear correlation for reasonably small linewidths. A meaningful interpretation of these plots, however, requires a reliable estimation of the linewidths. For this purpose the experimental distributions of linewidths in spectra of SH3 protein recorded at $\omega_{0H}/2\pi$ of 1 GHz were superimposed onto the respective curves, and the median values displayed by arrows.

The results shown here suggest that ^{15}N chemical shift yields the least ambiguous matching for microcrystalline samples. The next best case appears to be ^{13}CO nuclei, and ^{13}CB, ^{13}CA and $^{1}H^{N}$ in particular are less favorable for this purpose. The relative order for other samples, however, will depend on the actual observed linewidths, which are dependent on many factors including sample homogeneity, internal dynamics, field homogeneity, MAS rate, and the availability of ^{13}C-^{2}H decoupling. The case for amide linking is further supported by a low correlation of peaks in the ^{15}N-^{15}N plane as shown in Figures 2-3 and compared with other pairs of nuclei in Figure S2 in the Electronic Supplementary Material.

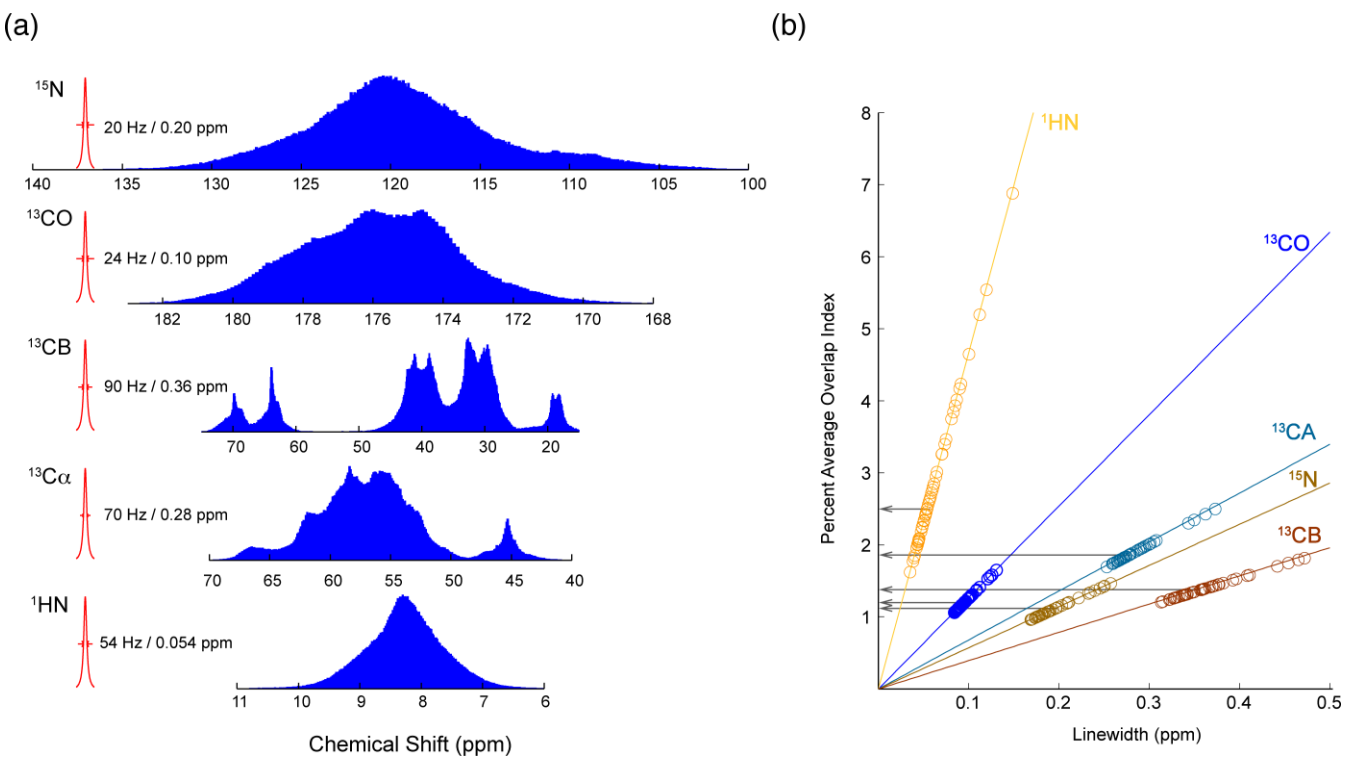

Fig. 4 In (a), histograms of chemical shifts of (from top to bottom) ^{15}N, ^{13}CO, ^{13}CB, ^{13}CA and $^{1}H^{N}$ with amino acid contributions weighted according to their occurrence in vertebrates. The horizontal scale of the histograms was adjusted to be in the same proportion to the respective linewidth in ppm for all nuclei. Indicated in red is the observed median of linewidths for microcrystalline SH3 at 23.5 T and at 60 kHz MAS. In (b) is shown the average percentage overlap index (approximately the probability of ambiguous matching) of ^{13}CA, ^{13}CB, ^{13}CO, ^{15}N and $^{1}H^{N}$ for a given linewidth. Experimental values of observed linewidths are superimposed as open circles onto respective curves to enable direct comparison between nuclei. The arrows indicate the overlap index for the median of linewidth in each case. Note that small increments in overlap index result in a dramatic increase in the overall assignment complexity (see Eq. 4).

In the case of non-crystalline samples, such as sedimented proteins, fibrils, or membrane proteins, bulk ^{13}C coherence lifetimes at 60 kHz MAS are significantly shorter than in microcrystals, resulting in some reduction of transfer efficiency (see Figure S1b in the Electronic Supplementary Material). Coherence lifetimes are improved at 111 kHz spinning as compared to 60 kHz, as exemplified by the 178-residue deuterated single-stranded DNA-binding protein (SSB) from *E. coli*, a sedimented (Mainz et al. 2009;

Bertini et al. 2011) protein known to undergo dynamic exchange processes (Marchetti et al. 2012). In this case, we observed a 23% improvement in ^{13}C coherence lifetimes T2′ to 16.8 ms (^{13}CO) and 18.7 ms (^{13}CA). For ^{1}H and ^{15}N, the improvement was 60% and 40% to 8.8 and 22 ms, respectively.

We measured T2′ values with a standard Hahn echo sequence, and for ^{13}C, with a selective Q3 pulse centered on the ^{13}CO or ^{13}CA of 350 or 600 μs. Low power swept TPPM (Lewandowski et al. 2010) at one quarter ω_R was applied during ^{13}C and ^{15}N echo times, in order to limit RF sample heating. If higher ^{1}H decoupling were applied, we could potentially improve the coherence times at lower spinning frequencies. However, in order to avoid interference with the 40 to 60 kHz sample spinning, high power decoupling would entail the use of above 120 to 180 kHz decoupling fields, which is quite demanding for both the hardware and sample stability. The improvement in coherence lifetimes at 111 kHz does not significantly impact the spectral resolution because the lines are largely inhomogeneously broadened, however the improvement does increase the sensitivity of the amide matching spectra by about 20% compared to 60 kHz due to improved $^{1}J_{CC}$ transfer efficiency (see Figure S1b in the Electronic Supplementary Material), partially compensate for the approximately 2-fold loss in detection sensitivity. We therefore expect that deuterated proteins would be better studied in the 1.3 mm probe at 60 kHz. Surprisingly, despite the reduced sample amount in the 0.7 mm system, initial measurements suggest that the detection sensitivity was no worse than in the 1.3 mm probe. Based on the same considerations of coherence lifetime, sample volume, and coil diameter, a 1.9 mm probe spinning at 40 kHz would be expected to further improve the efficiency of these experiments (although at some loss in resolution). The loss in sensitivity is a factor of about 0.6 from the CC transfer (see Figure S1), but there is an expected gain of about 3-fold for detection sensitivity (Andreas et al. 2015). However, recent measurements determined that the 1.3 mm probe results in higher sensitivity than the 40 kHz probe for the (H)CA(CO)NH experiment, indicating that the heteronuclear cross-polarization transfer efficiencies are also likely influenced by the spinning frequency (Nieuwkoop et al. 2015). A final verdict on the probe with best sensitivity for this experiment therefore awaits further measurements.

At 111 kHz MAS in a 0.7 mm rotor, we recorded an (H)N(CACO)NH spectrum of SSB with 45% efficiency for each N-H and H-N transfer step and above 3% efficiency for the remaining three transfers combined (Figure 5). Due to incomplete H/D exchange, however, we were not able to characterize the completeness of the resulting dataset or to precisely determine the experimental efficiency.

For fully-protonated microcrystalline GB1 microcrystals, we found a significant improvement in the carbon homogeneous coherence lifetimes (T_2') at 111 kHz compared with 60 kHz MAS. The ^{13}CO T_2' was 77 ms, and the ^{13}CA was 46 ms at 111 kHZ, which is about 50% and 300% improvement, respectively, compared with 60 kHz spinning. Although we expect lower total sensitivity compared with deuterated proteins, the 111 kHz probe will enable the approach presented herein for interesting proteins for which partial deuteration is not amenable. Figure 3b shows the (H)N(CACO)NH spectrum recorded on fully protonated microcrystalline GB1. In this 56-residue protein, all of the possible cross-peaks were observed extending from residue 2 to 56.

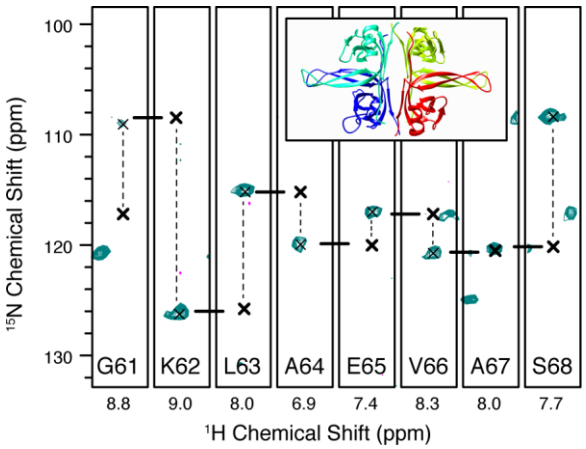

Fig. 5 Selected ^{15}N-^{1}H planes of the (H)N(CACO)NH spectrum, showing connectivity in sedimented SSB. The position of the peak in the second indirect ^{15}N dimension is marked with a bold x. The X-ray structure of tetrameric SSB (PDB 4MZ9 (Shishmarev et al. 2014)) is depicted in the inset. Recorded at 111 kHz MAS and $\omega_{0H}/2\pi$ =1 GHz

Increases in the MAS rate lead to higher centrifugal acceleration of the sample despite the reduction in rotor diameter (Agarwal et al. 2014). For this reason, sealing the rotor with glue or rubber plugs has been applied to 1.3 mm rotors in order to limit solvent loss (Asami et al. 2012). In the case of the 0.7

mm rotor, we found that the Vespel caps are sufficient to provide a watertight seal over a period of several weeks for a variety of protein samples including the sedimented SSB protein.

In conclusion, we have described a method that simplifies protein resonance assignment, providing spectra that can be rapidly and automatically assigned using UNIO-MATCH. We find that amide signals provide particularly unambiguous sequential connectivity and can be applied generally to microcrystalline and sedimented deuterated proteins spinning at 60 and 111 kHz, with as little as 0.5 mg of sample. We expect the method to be of utility for a wide range of insoluble crystalline or non-crystalline samples.

While this manuscript was under review, a very similar approach to residue linking was reported, but using the HORROR condition for the ^{13}C-^{13}C transfer (Xiang et al. 2015).

Electronic Supplementary Material

Sample preparation, comparison of ^{1}H, ^{13}C and ^{15}N resolution and degree of overlap for inter-residue correlation, simulated ^{13}C-^{13}C transfer efficiency curves.

Acknowledgements

We thank the members of the technical staff of the ISA for assistance with the NMR spectrometers. We acknowledge support from CNRS (Fondation pour la Chimie des Substances Naturelles) and from the People Programme of the European Union's FP7 (FP7-PEOPLE-2012-ITN REA grant agreement n°317127 “pNMR” and 316630 “CAS-IDP”). LBA is supported by a MC incoming fellowship (REA grant agreement n°624918 “MEM-MAS”), and JS by an EMBO fellowship (ALTF 1506-2014) and by the Marie Curie Actions of the European Commission (LTFCOFUND2013, GA-2013-609409).

References

Agarwal V, Penzel S, Szekely K, Cadalbert R, Testori E, Oss A, Past J, Samoson A, Ernst M, Bockmann A, Meier BH (2014). De novo 3D structure determination from sub-milligram protein samples by solid-state 100 kHz MAS NMR spectroscopy. Angew Chem Int Ed Engl, 53:12253-12256.

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.

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, 54:155-168.

Barbet-Massin E, Pell AJ, Jaudzems K, Franks WT, Retel JS, Kotelovica S, Akopjana I, Tars K, Emsley L, Oschkinat H, Lesage A, Pintacuda G (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, 56:379-386.

Barbet-Massin E, Pell AJ, Retel JS, Andreas LB, Jaudzems K, Franks WT, Nieuwkoop AJ, Hiller M, Higman V, Guerry P, Ber- tarello A, Knight MJ, Felletti M, Le Marchand T, Kotelovica S, Akopjana I, Tars K, Stoppini M, Vittorio B, Bolognesi M, Ricagno S, Chou JJ, Griffin RG, Oschkinat H, Lesage A, Emsley L, Herrmann T, Pintacuda G (2014). Rapid proton-detected NMR assignment for proteins with fast magic angle spinning. J Am Chem Soc, 136:12489-12497.

Bertini I, Luchinat C, Parigi G, Ravera E, Reif B, Turano P (2011). Solid-state NMR of proteins sedimented by ultracentrifugation. Proc Natl Acad Sci USA, 108:10396-10399.

Böckmann A, Lange A, Galinier A, Luca S, Giraud N, Juy M, Heise H, Montserret R, Penin F, Baldus M (2003). Solid state NMR sequential resonance assignments and conformational analysis of the 2 x 10.4 kDa dimeric form of the Bacillus subtilis protein Crh. J Biomol NMR, 27:323-339.

Bracken C, Palmer AG, Cavanagh J (1997). (H)N(COCA)NH and HN(COCA)NH experiments for ^{1}H-^{15}N backbone assignments in ^{13}C/^{15}N-labeled proteins. J Biomol NMR, 9:94-100.

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, 45:3878-3881.

Dutta SK, Serrano P, Proudfoot A, Geralt M, Pedrini B, Herrmann T, Wüthrich K (2015). APSY-NMR for protein backbone assignment in high-throughput structural biology. J Biomol NMR, 61:47-53.

Emsley L, Bodenhausen G (1992). Optimization of shaped selective pulses for NMR using a quaternion description of their overall propagators. J Magn Reson, 97:135-148.

Franks WT, Zhou DH, Wylie BJ, Money BG, Graesser DT, Frericks HL, Sahota G, Rienstra CM (2005). Magic-angle spinning solid-state NMR spectroscopy of the beta 1 immunoglobulin binding domain of protein G (GB1): ^{15}N and ^{13}C chemical shift assignments and conformational analysis. J. Am. Chem. Soc., 127:12291-12305.

Frueh DP, Sun ZY, Vosburg DA, Walsh CT, Hoch JC, Wagner G (2006). Non-uniformly sampled double-TROSY hNcaNH experiments for NMR sequential assignments of large proteins. J Am Chem Soc, 128:5757-5763.

Hong M (1999). Resonance assignment of ^{13}C/^{15}N labeled solid proteins by two- and three-dimensional magic-angle-spinning NMR. J. Biomol. NMR, 15:1-14.

Igumenova TI, Wand AJ, McDermott AE (2004). Assignment of the backbone resonances for microcrystalline ubiquitin. J Am Chem Soc, 126:5323-5331.

Knight MJ, Pell AJ, Bertini I, Felli IC, Gonnelli L, Pierattelli R, Herrmann T, Emsley L, Pintacuda G (2012). Structure and backbone dynamics of a microcrystalline metalloprotein by solid-state NMR. Proc Natl Acad Sci USA, 109:11095-11100.

Knight MJ, Webber AL, Pell AJ, Guerry P, Barbet-Massin E, Bertini I, Felli IC, Gonnelli L, Pierattelli R, Emsley L, Lesage A, Herrmann T, Pintacuda G (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, 50:11697-11701.

Kobayashi T, Mao K, Paluch P, Nowak-Krol A, Sniechowska J, Nishiyama Y, Gryko DT, Potrzebowski MJ, Pruski M (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, 52:14108-14111.

Lamley JM, Iuga D, Öster C, Sass H-J, Rogowski M, Oss A, Past J, Reinhold A, Grzesiek S, Samoson A, Lewandowski JR (2014). Solid-State NMR of a Protein in a Precipitated Complex with a Full-Length Antibody. J Am Chem Soc, 136:16800-16806.

Lewandowski JR, Sein J, Sass HJ, Grzesiek S, Blackledge M, Emsley L (2010). Measurement of site-specific ^{13}C spin-lattice relaxation in a crystalline protein. J Am Chem Soc, 132:8252-8253.

Linser R, Dasari M, Hiller M, Higman V, Fink U, del Amo JML, Markovic S, Handel L, Kessler B, Schmieder P, Oesterhelt D, Oschkinat H, Reif B (2011). Proton-detected solid-state NMR spectroscopy of fibrillar and membrane proteins. Angew Chem Int Ed Engl, 50:4508-4512.

Liu A, Riek R, Wider G, von Schroetter C, Zahn R, Wuthrich K (2000). NMR experiments for resonance assignments of ^{13}C, ^{15}N doubly-labeled flexible polypeptides: application to the human prion protein hPrP(23-230). J Biomol NMR, 16:127-138.

Mainz A, Jehle S, van Rossum BJ, Oschkinat H, Reif B (2009). Large protein complexes with extreme rotational correlation times investigated in solution by magic-angle-spinning NMR spectroscopy. J Am Chem Soc, 131:15968-15969.

Marchetti A, Jehle S, Felletti M, Knight MJ, Wang Y, Xu ZQ, Park AY, Otting G, Lesage A, Emsley L, Dixon NE, Pintacuda G (2012). Backbone assignment of fully protonated solid proteins by ^{1}H detection and ultrafast Magic-Angle-Spinning NMR spectroscopy. Angew Chem Int Ed Engl, 51:10756-10759.

Matsuo H, Kupce E, Li H, Wagner G (1996). Use of selective C alpha pulses for improvement of HN(CA)CO-D and HN(COCA)NH-D experiments. J Magn Reson B, 111:194-198.

Nieuwkoop AJ, Franks WT, Rehbein K, Diehl A, Akbey U, Engelke F, Emsley L, Pintacuda G, Oschkinat H (2015). Sensitivity and resolution of proton detected spectra of a deuterated protein at 40 and 60 kHz magic-angle-spinning. J Biomol NMR, 61:161-171.

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.

Panchal SC, Bhavesh NS, Hosur RV (2001). Improved 3D triple resonance experiments, HNN and HN(C)N, for HN and ^{15}N sequential correlations in (^{13}C, ^{15}N) labeled proteins: application to unfolded proteins. J Biomol NMR, 20:135-147.

Reif B, Griffin RG (2003). ^{1}H detected $^{1}H,^{15}N$ correlation spectroscopy in rotating solids. J Magn Reson, 160:78-83.

Rienstra CM, Hohwy M, Hong M, Griffin RG (2000). 2D and 3D $^{15}N^{13}C^{13}C$ NMR Chemical Shift Correlation spectroscopy of Solids: Assignment of MAS Spectra of Peptides. J Am Chem Soc, 122:10979 - 10990.

Schanda P, Meier BH, Ernst M (2010). Quantitative analysis of protein backbone dynamics in microcrystalline ubiquitin by solid-state NMR spectroscopy. J Am Chem Soc, 132:15957-15967.

Shishmarev D, Wang Y, Mason CE, Su XC, Oakley AJ, Graham B, Huber T, Dixon NE, Otting G (2014). Intramolecular binding mode of the C-terminus of *Escherichia coli* single-stranded DNA binding protein determined by nuclear magnetic resonance spectroscopy. Nucleic Acids Res, 42:2750-2757.

van Rossum BJ, Castellani F, Pauli J, Rehbein K, Hollander J, de Groot HJ, Oschkinat H (2003). Assignment of amide proton signals by combined evaluation of HN, NN and HNCA MAS-NMR correlation spectra. J Biomol NMR, 25:217-223.

Volk J, Herrmann T, Wuthrich K (2008). Automated sequence-specific protein NMR assignment using the memetic algorithm MATCH. J Biomol NMR, 41:127-138.

Ward ME, Shi L, Lake E, Krishnamurthy S, Hutchins H, Brown LS, Ladizhansky V (2011). Proton-detected solid-state NMR reveals intramembrane polar networks in a seven-helical transmembrane protein proteorhodopsin. J Am Chem Soc, 133:17434-17443.

Weisemann R, Ruterjans H, Bermel W (1993). 3D triple-resonance NMR techniques for the sequential assignment of NH and ^{15}N resonances in ^{15}N- and ^{13}C-labelled proteins. J Biomol NMR, 3:113-120.

Xiang S, Chevelkov V, Becker S, Lange A (2014). Towards automatic protein backbone assignment using proton-detected 4D solid-state NMR data. J. Biomol. NMR.

Xiang S, Grohe K, Rovo P, Vasa SK, Giller K, Becker S, Linser R (2015). Sequential backbone assignment based on dipolar amide-to-amide correlation experiments. J. Biomol. NMR.

Yoshimura Y, Kulminskaya NV, Mulder FA (2015). Easy and unambiguous sequential assignments of intrinsically disordered proteins by correlating the backbone ^{15}N or ^{13}C' chemical shifts of multiple contiguous residues in highly resolved 3D spectra. J Biomol NMR, 61:109-121.

Zhou DH, Nieuwkoop AJ, Berthold DA, Comellas G, Sperling LJ, Tang M, Shah GJ, Brea EJ, Lemkau LR, Rienstra CM (2012). Solid-state NMR analysis of membrane proteins and protein aggregates by proton detected spectroscopy. J Biomol NMR, 54:291-305.

Zhou DH, Rienstra CM (2008). High-performance solvent suppression for proton detected solid-state NMR. J Magn Reson, 192:167-172.