Protein Backbone and Side Chain Assignment using Proton Detection under Ultra-Fast Magic Angle Spinning

Venita Decker
Bruker BioSpin GmbH, Rheinstetten, Germany

Introduction

Recently, a basic set of three $^1$H-detected experiments using solid-state NMR has been announced in the Bruker Application Note “Introduction to Proton Detection in Biological Samples under Ultra-Fast Magic Angle Spinning”[1]:

1. hNH: 1D and 2D, pulse program (pulprog): hNH2D.dcp, Fig. 2A
2. hCaNH: 1D-3D, pulprog: hCaNH3D.tcp, Fig. 2A
3. hCONH: 1D-3D, pulprog: hCONH3D.tcp, Fig. 2B.

Based on this Note, we are pleased to present, once again in cooperation with Guido Pintacuda and co-workers, a full set of $^1$H-detected experiments that allows for an accelerated resonance assignment of a protein’s backbone and side chains[2-4]:

4. hCH: 1D and 2D, pulprog: hCH2D.dcp, Figs. 2A and 5A, B
5. hCO(Cα)NH: 1D-3D, pulprog: hCOcaNH3D.tcp, Figs. 2A, C
6. h(CO)Cα(CO)NH: 1D-3D, pulprog: hCOcaCOcaNH3D.tcp, Figs. 2A, B
7. h(Cα)Cβ(Cα)NH: 1D-3D, pulprog: hCbcacoNH3D.tcp, Figs. 2A, E
8. h(Cα)Cβ(Cα)C(NH): 1D-3D, pulprog: hCbcacoNH3D.tcp, Figs. 2B, F.

To guarantee a smooth workflow, we implemented two additional pulse sequences for a straight-forward optimization of specific $^{13}$C pulses:

9. hCOH: 1D, pulprog: hCOH1D.dcp, Figs. 5A, C
10. hCaH: 1D, pulprog: hCaH1D.dcp, Figs. 5A, D.

Figure 1: 1H-detected 2D hCO(Cα)NH (red) and 3D h(Cα)Cβ(Cα)C(NH)NH (blue) spectra of a u-$^{13}$C,$^{15}$N-labeled, deuterated and 100% $^1$H back-exchanged GB1 sample recorded on a Bruker 600 MHz spectrometer at 60 kHz MAS in about 95 minutes (2D) and 1 day (3D), yielding intraresidue (CO-N-H) red and interresidue CB,N-H side chain-backbone blue information, respectively.
The Experiments

All experiments presented here yield either correlations within one amino acid (intraresidue, Fig. 2A) or between consecutive residues (interresidue, Fig. 2B). Whereas the previously introduced hNH\textsuperscript{[1]} (Fig. 2A, cyan) correlates the amide proton to its nitrogen, in the hCH experiment (Fig. 2A, dark green, Fig. 5A) carbons are correlated to protons. Since no discrimination between different carbon nuclei is made here and no nitrogen is involved, there are experiments of higher selectivity: the previously introduced hC\textsubscript{\textalpha}C\textsubscript{\textbeta} experiment (Fig. 2A, purple) and the hC\textsubscript{\textalpha}C\textsubscript{\textbeta} experiment (Figs. 2A, light green and 2C) correlate the amide group to the C\textsubscript{\textalpha} and CO atom of the same amino acid residue, respectively.

Knowing the chemical shifts of the backbone atoms N, C\textsubscript{\textalpha} and CO can reveal important information about the protein's secondary structure. Nevertheless, the amino acid type, which is specified by its side chain, has to be determined for 3D protein structure elucidation as well, since most distance information is encoded in contacts between different side chain nuclei. Because of chemical shift overlap, backbone atoms alone are not suited for the determination of the amino acid type. The hC\textsubscript{\textalpha}C\textsubscript{\textbeta} experiment (Figs. 2A, orange and 2E) closes this gap by providing side chain C\textbeta chemical shifts correlated to the intraresidue amide group. Different from the hC\textsubscript{\textalpha}NH, C\textalpha does not evolve, but is used to selectively transfer the signal to the C\textbeta of the same residue. After evolution of C\textbeta, polarization is transferred back to C\textalpha further to the amide group of the same residue. In the literature, this scalar-based transfer scheme is described as 'out-and-back'\textsuperscript{[2]}, thus, it will be in the following as well. A detailed description is given below.

Interresidue correlations are analyzed to get the sequential backbone assignment. Accordingly, not only remaining unknown amino acid types can be specified, but the protein's secondary structure can be evaluated. The previously introduced hC\textalpha C\textalpha experiment (Fig. 2B, cyan) experiment correlates the backbone CO of one residue to the amide group of its consecutive neighbor. Sequential correlations between the C\textalpha of a residue 'i-1' and the amide group of residue 'i' can be obtained with the hC\textalpha C\textalpha experiment (Figs. 2B, green and 2D), which relies on the out-and-back transfer scheme between CO and C\textalpha. After evolution of C\textalpha, polarization is transferred back to CO and further to the amide group of the consecutive residue.

In an unique experiment, the hC\textalpha C\textbeta C\textalpha C\textalpha experiment (Figs. 2B, orange and 2F), the side chain C\textbeta of a residue 'i-1' is correlated to the backbone amide group of the consecutive residue 'i'. As
before, polarization is transferred in the out-and-back scheme via Ca to CB and, after evolution of CB, back to Ca. Eventually, polarization is transferred from the CO of residue ‘i-1’ forward sequentially to the amide group of residue ‘i’.

Note, that especially the combined analysis of intra- and inter-residue correlation spectra, reconciled with the protein amino acid sequence, allows for a fast ‘sequential walk’ assignment (compare with Figs. 2A, B).

**Insensitive Nuclei Enhanced by Polarization Transfer (INEPT)**

In the presented experiments, homonuclear $^{13}$C-13C polarization transfer is mediated through-bond by the scalar-based INEPT, a well-known scheme in solution NMR spectroscopy (Fig. 3A).[5] A building block consists of the spin echo sequence ‘delay – π pulses on both spins I and S – delay’ (blue). If the delay is set to $1/(4|\text{J}|_\alpha\beta)$ and the two π pulses are applied in the center of the sequence, scalar $J$-coupling between spins I and S evolves into an anti-phase term on spin I at the end of the spin echo. If a second spin echo is applied consecutively to the π/2 pulses (Fig. 3B, green), $J_{\alpha\beta}$ coupling is refocused at the end of the sequence as well, resulting in in-phase magnetization on the S spin.

**Figure 3:** (A) A non-refocused INEPT starts with x magnetization on spin I (I). During a time of $1/(2|\text{J}|_\alpha\beta)$, scalar $J_{\alpha\beta}$ coupling is evolving into the anti-phase term $2|\text{I}|_\alpha|\text{S}|_\beta$ on spin I, if two π pulses (white) are applied simultaneously on both spins in the center of the sequence. With the π pulse applied on spin I, chemical shift evolution is refocused as well. At the end of this so-called spin echo (blue), polarization is transferred by two π/2 pulses (black) from I to S and becomes $-2|\text{S}|_\beta$. (B) In the refocused version, scalar coupling is refocused after a second INEPT block (green), but due to the polarization transfer from spin I to S, in-phase magnetization is now on the S spin (S). (C) Building block for Bloch-Siegert (BS) shift compensation (yellow) resulting from homonuclear $J_{\text{CO-Ca}}$ decoupling during CO $t_1$ evolution using band-selective π pulses (broad bell shaped) on Ca (light green) and CO (dark green). The symmetry delay Δ is needed for initial $t_1$ compensation. (D) Integration of CO $t_1$ evolution with BS shift and initial $t_1$ compensation into an INEPT step (striped background) to evolve $J_{\text{Ca-Co}}$ coupling while refocusing initial $t_1$ and BS shift evolution. BS shift inducing pulses are labeled with ‘#’, BS shift compensating pulses with ‘##’. Further information is given below.

**Band-Selective Pulses**

Though depicted as three different channels in Fig. 2, CO, Ca and CB belong to the same logical channel in the spectrometer hardware ($^{13}$C channel: F3 in ‘edasp’). Band-selective π (broad bell shaped) and π/2 (narrow bell shaped) pulses have to be applied on CO (dark green), Ca (light green) and Ca-CB (dark red), respectively, whenever band-selectivity is essential. For example, during CO and Ca $t_1$ evolution in the hCOcaNH3D.tcp and hccoCacoNH3D.tcp pulse programs, respectively, $J_{\text{CO-Ca}}$ needs to be decoupled using an off-resonance Ca (light green) or CO (dark green) band-selective π pulse (Figs. 2C, D).

In theory this is true for the CB $t_1$ evolution as well, where decoupling of $J_{\text{Ca-CB}}$ would be advantageous. But due to the fact that the CB region (~70 to 20 ppm) includes the Ca band (~65 to 45 ppm), there is a principle limitation for pure Ca band-selective decoupling. To prevent unwanted signal loss, the hcaCbcaNH3D.tcp and hcaCbcacoNH3D.tcp pulse programs work without a discrimination of Ca and CB. Nevertheless, band-selective π pulses for the whole Ca-CB region (dark red pulses) are used to prevent signal losses from unwanted Ca to CO transfer (Figs. 2E, F).

Because CO is not bonded to CB, in case of pure CO to Ca INEPTs, no band-selective pulses are necessary (Figs. 2D, F). Ca to CO transfer on the other hand is only possible without signal loss to CB using the out-and-back scheme (‘CO-Ca-CO’, Fig. 2D), which intrinsically ensures for selectivity (see below).

That is the reason why there is no hCa(CO)NH sequence, complementary to the hCO(Ca)NH sequence.

**Refocusing Bloch-Siegert shifts during INEPT**

An off-resonance band-selective π pulse causes phase evolution of the on-resonance spin, which is referred to as Bloch-Siegert (BS) shift.[6] This shift is induced during any homonuclear INEPT transfer that needs to use band-selective π pulses to discriminate CO from Ca/Ca-CB to prevent from unwanted transfer losses (Fig. 2C, green and Fig. 2F, blue striped background). Furthermore, it occurs for every $t_1$ evolution, where band-selective homonuclear decoupling is needed, e.g. $J_{\text{CO-Ca}}$ decoupling during CO (Fig. 2C, blue striped background) and Ca $t_1$ evolution (Fig. 2D, yellow), respectively. Here, an additional compensation delay (Δ) is needed, since the initial $t_1$ is greater than zero, which otherwise leads to unwanted chemical shift evolution in the first TD point. A detailed description is given in the previous Application Note.[1] The BS shift is compensated by a combination of one on-resonance band-selective π pulse between two off-resonance ones (Fig. 3C, yellow; here with initial $t_1$ compensation). Such a building block can be combined with an INEPT by introducing the two INEPT delays, $1/(4|\text{J}|_\alpha\beta)$, before the on-resonance π pulse and after the second off-resonance one, respectively, to allow for evolution of scalar couplings while refocusing the BS shift (Fig. 2C, green and Fig. 2F, blue striped background).

In case the BS shift is induced by a decoupling pulse during $t_1$ evolution (Fig. 3D, light green pulse labeled with ‘#’), the first INEPT delay is prolonged by Δ, while the second INEPT delay follows $t_1$. That way, both initial $t_1$ and BS shift compensation are integrated into an INEPT step (Fig. 2C, blue striped background).
Another critical point is the presence of a third spin, K, which is coupled to the first, spin I. In the ‘I to S only’ scheme the selectivity is reduced, since parts of the initial magnetization may get lost to spin K anti-phase coherences (Fig. 4A, gray). In contrast, using the out-and-back scheme prevents from this unwanted signal loss, because no spin I-K coherences are present, which could dephase into 2I-K₂ anti-phase (compare with Fig. 4B) making this transfer highly band-selective.[3]

Which Experimental Conditions to Choose?

A detailed summary of parameter recommendations (CP conditions, decoupling, shapes, etc.) has been given in the previous Application Note already[11], which applies to the experiments presented here as well. Additional important parameters are the different band-selective 13C pulses, INEPT delays and specific frequency offsets. A detailed parameter overview is given in Table 1.

In the new experiments the delicate setting of 13C frequency offsets (carrier, frequency switches, etc.) is regulated with the offset constants ‘cnst21,’ ‘cnst22,’ ‘cnst24’ and ‘cnst26’ as the center of the CO, Ca, CO-Ca and Ca-Cβ region, respectively. This does not only prevent from inappropriate experimental settings, but also for the use of smallest spectral widths to save measurement time, since all indirect dimensions are acquired on-resonance with the spectral region of interest.

The Out-and-Back Scheme

In soluble samples, the intense proton dipolar interaction network is averaged out by molecular motion. Because this is not the case for insoluble samples, the coherence decay in solid-state NMR is dominated by coherent contributions from homonuclear proton interactions.[3]

As expected, the effect on CO spins, which do not feature directly bound protons, is least drastic. With a coherence decay time constant (T₂') of approximately 40-50 ms, even in non-perdeuterated samples, the transverse coherence lifetime is sufficiently long to evolve homonuclear scalar couplings of:

\[ \frac{1}{J_{CaCO}} \approx 53 \text{ Hz} \quad \Rightarrow \quad \frac{1}{2 \cdot 1/J_{CaCO}} \approx 9.3 \text{ ms} \quad \Rightarrow \quad < T_{2}' \]

Aliphatic carbons however suffer from a short T₂' (only about 20 ms in perdeuterated/ back-exchanged samples and less than 15 ms in 100% protonated samples). Thus, homonuclear scalar J_{CaCO} and J_{CacB} couplings may not evolve efficiently:

\[ \frac{1}{J_{CaCB}} \approx 35 \text{ Hz} \quad \Rightarrow \quad \frac{1}{2 \cdot 1/J_{CaCB}} \approx 14.3 \text{ ms} \quad \Rightarrow \quad < T_{2}' \]

In contrast to a conventional ‘I to S only’ transfer (Fig. 4A), which depends on the coherence lifetimes of both spins, the out-and-back scheme relies on the T₂' of spin S only (Fig. 4B). Thus, the fast transverse relaxation of Ca spins can be neglected and sensitivity increases dramatically when using a CO to Ca to CO transfer.[3]

Which Experimental Conditions to Choose?

A detailed summary of parameter recommendations (CP conditions, decoupling, shapes, etc.) has been given in the previous Application Note already[11], which applies to the experiments presented here as well. Additional important parameters are the different band-selective 13C pulses, INEPT delays and specific frequency offsets. A detailed parameter overview is given in Table 1.

In the new experiments the delicate setting of 13C frequency offsets (carrier, frequency switches, etc.) is regulated with the offset constants ‘cnst21,’ ‘cnst22,’ ‘cnst24’ and ‘cnst26’ as the center of the CO, Ca, CO-Ca and Ca-Cβ region, respectively. This does not only prevent from inappropriate experimental settings, but also for the use of smallest spectral widths to save measurement time, since all indirect dimensions are acquired on-resonance with the spectral region of interest.

The Out-and-Back Scheme

In soluble samples, the intense proton dipolar interaction network is averaged out by molecular motion. Because this is not the case for insoluble samples, the coherence decay in solid-state NMR is dominated by coherent contributions from homonuclear proton interactions.[3]

As expected, the effect on CO spins, which do not feature directly bound protons, is least drastic. With a coherence decay time constant (T₂') of approximately 40-50 ms, even in non-perdeuterated samples, the transverse coherence lifetime is sufficiently long to evolve homonuclear scalar couplings of:

\[ \frac{1}{J_{CaCO}} \approx 53 \text{ Hz} \quad \Rightarrow \quad \frac{1}{2 \cdot 1/J_{CaCO}} \approx 9.3 \text{ ms} \quad \Rightarrow \quad < T_{2}' \]

Aliphatic carbons however suffer from a short T₂' (only about 20 ms in perdeuterated/ back-exchanged samples and less than 15 ms in 100% protonated samples). Thus, homonuclear scalar J_{CaCO} and J_{CacB} couplings may not evolve efficiently:

\[ \frac{1}{J_{CaCB}} \approx 35 \text{ Hz} \quad \Rightarrow \quad \frac{1}{2 \cdot 1/J_{CaCB}} \approx 14.3 \text{ ms} \quad \Rightarrow \quad < T_{2}' \]

In contrast to a conventional ‘I to S only’ transfer (Fig. 4A), which depends on the coherence lifetimes of both spins, the out-and-back scheme relies on the T₂' of spin S only (Fig. 4B). Thus, the fast transverse relaxation of Ca spins can be neglected and sensitivity increases dramatically when using a CO to Ca to CO transfer.[3]
A special case is the hcoCacoNH3D.tcp sequence, where the \(^{13}\)C carrier frequency ‘o3’ needs to be at the center of the CO-C\(_\alpha\) region. To be on resonance with C\(_\alpha\) in the indirect dimension, the receiver phase is shifted automatically to the frequency defined by ‘cnst22’ in the pulse program. The resulting artificial shift in the processed \(^{13}\)Ca dimension of the spectrum needs to be corrected manually and is given in the status parameter ‘cnst42’ (Hz). When typing ‘spec42’ into the TopSpin\textsuperscript{TM} command line, the value can be copied and added to the ‘edp’ processing parameter ‘sr’ of the ‘F1’ dimension for correct spectral referencing, e.g.:

\[ ST_{F1,initial} = -9 \text{ Hz} \ ; \ \text{cnst}42 = 6 \text{ Hz} \implies ST_{F1,final} = -3 \text{ Hz} . \]

Gaussian pulse cascades\[^{[7]}\] are well-suited as band-selective pulses. The corresponding pulse shapes in the Bruker library are called ‘Q3’ for \(\alpha\) and ‘Q5’ (time-reversed ‘Q5’\(^\text{r}\)) for \(\beta\)/2 pulses. Band-selectivity is regulated with the length of the pulse. While the CO band is well-separated from the aliphatic region, leading to short pulses for the CO- as well as the C\(_\alpha\)-C\(_\beta\) band (200 to 256 us at 600 MHz B\(_0\)), C\(_\alpha\) band-selective pulses need to be rather long to ensure reasonable selectivity from the C\(_\beta\) band (600 to 700 us at 600 MHz B\(_0\)). Note, that because of the better chemical shift dispersion at higher B\(_0\) fields, these band-selective pulse lengths are proportional to 1/B\(_0\).

The corresponding power levels are calculated from the \(^{13}\)C \(\pi/2\) hard pulse using Bruker’s Shape Tool (‘stdisp’). It is recommended to optimize these calculated values. To guarantee a straight-forward workflow, Bruker implemented two additional 1D pulse programs, the hCOH1D.dcp and hCaH1D.dcp, which are based on the hCH2D.dcp sequence (Figs. 2A and 5A). In contrast to the 2D experiment, which is variable in terms of \(^{13}\)C frequency offset, the 1D sequences are written explicitly to fit the respective band-selectivity for proper parameter optimization. Thus, in the hCOH ‘cnst21’, and in the hCaH ‘cnst22’ and ‘cnst24’ have to be set. Then, all relevant frequency offsets (except carrier) will be calculated and set automatically. By setting different flags as ‘zgoptns’, the corresponding band-selective pulses are incorporated into the pulse program and their power levels can be optimized (Figs 5B-D).

In the hcaCbcaNH3D.tcp and hcaCbcacoNH3D.tcp sequences, C\(_\beta\) is involved in the transfer scheme. Here, a z-filter (t\(_z\)) needs to be applied before magnetization can be transferred further to \(^{13}\)N or \(^{13}\)CO, respectively, to ensure for pure C\(_\alpha\) polarization beforehand.\[^{[4]}\] The corresponding parameter ‘db\(_z\)’ takes values in the range of 1 to 5 ms.

Since the different INEPT delays (‘d22’/ ‘d25’, ‘d24’) derive from scalar coupling constants, their values should not vary greatly. Nevertheless, a fine-tuning on each sample is recommended using solid-state NMR.

Note, that a detailed description of all parameters used in the experiments and corresponding recommended conditions can be found in the header of each pulse program.

### Table 1: Summary of parameter recommendations. Parameters in bold letters should be optimized. For the remaining ones it is sufficient to calculate/set the recommended values.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>hCOcaNH3D.tcp</th>
<th>hcoCacoNH3D.tcp</th>
<th>hcaCbcaNH3D.tcp</th>
<th>hcaCbcacoNH3D.tcp</th>
</tr>
</thead>
<tbody>
<tr>
<td>(^{13})C offset: o3</td>
<td>set to cnst21 (−175 ppm)</td>
<td>set to cnst26 (−115 ppm)</td>
<td>set to cnst24 (−40 ppm)</td>
<td>set to cnst24 (−40 ppm)</td>
</tr>
<tr>
<td>Band-selective pulses(^*)</td>
<td>(^{13})CO (\pi) pulse: p23 @ spnam23, spdb23**</td>
<td>−256 us @ shape ‘Q3.2000’</td>
<td>−256 us @ shape ‘Q3.2000’</td>
<td>−256 us @ shape ‘Q3.2000’</td>
</tr>
<tr>
<td></td>
<td>(^{13})C(<em>\alpha)-(^{13})C(</em>\beta) (\pi) pulse: p20 @ spnam20, spdb20**</td>
<td>−600 to 700 us @ shape ‘Q3.2000’</td>
<td>−600 to 700 us @ shape ‘Q3.2000’</td>
<td>−600 to 700 us @ shape ‘Q3.2000’</td>
</tr>
<tr>
<td>INEPT delays 1/(4J)</td>
<td>CO to C(_\alpha): d22</td>
<td>−4.7 ms (J = −53 Hz)</td>
<td>−4.7 ms (J = −53 Hz)</td>
<td>−4.7 ms (J = −53 Hz)</td>
</tr>
<tr>
<td></td>
<td>C(_\alpha) to CO: d25</td>
<td>−4.7 ms (J = −53 Hz), can be less for shorter T(_z)**</td>
<td>−4.7 ms (J = −53 Hz), can be less for shorter T(_z)**</td>
<td>−4.7 ms (J = −53 Hz), can be less for shorter T(_z)**</td>
</tr>
<tr>
<td>C(<em>\alpha)-C(</em>\beta): d24</td>
<td>-</td>
<td>-</td>
<td>−7.2 ms (J = −35 Hz)</td>
<td>−7.2 ms (J = −35 Hz)</td>
</tr>
<tr>
<td>Offset constants</td>
<td>cnst21: CO (−175 ppm)</td>
<td>☑</td>
<td>☑</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>cnst22: C(_\alpha) (−55 ppm)</td>
<td>☑</td>
<td>☑</td>
<td>☑</td>
</tr>
<tr>
<td></td>
<td>cnst24: C(<em>\alpha)-C(</em>\beta) (−40 ppm)</td>
<td>-</td>
<td>-</td>
<td>☑</td>
</tr>
<tr>
<td></td>
<td>cnst26: CO-C(_\alpha) (−115 ppm)</td>
<td>☑</td>
<td>☑</td>
<td>-</td>
</tr>
</tbody>
</table>

\(^*\) The pulse length changes linearly with 1/B\(_0\). Pulse lengths in the table correspond to 600 MHz B\(_0\) field.

** Open corresponding shape in TopSpin\textsuperscript{TM} Shape Tool: ‘stdisp> Analysis> Integrate Shape’ and enter corresponding values: ‘length of selective pulse, ‘180’ or ‘90’, length of p\(_1\)’ Result tells by how much pldb1 needs to be changed to get the corresponding power level for spdb**.
References


