Protocols for the Sequential Solid-State NMR Spectroscopic Assignment of a Uniformly Labeled 25 kDa Protein: HET-s(1-227)

ARTICLE *in* CHEMBIOCHEM · JULY 2010

Impact Factor: 3.09 · DOI: 10.1002/cbic.201000124 · Source: PubMed

8 AUTHORS, INCLUDING:

Christian Wasmer
Harvard Medical School
18 PUBLICATIONS **1,158** CITATIONS
SEE PROFILE

Birgit Habenstein
Institut Européen De Chimie Et Biologie
28 PUBLICATIONS **441** CITATIONS
SEE PROFILE

Beat H Meier
ETH Zurich
267 PUBLICATIONS **13,246** CITATIONS
SEE PROFILE

DOI: 10.1002/cbic.201000124

Protocols for the Sequential Solid-State NMR Spectroscopic Assignment of a Uniformly Labeled 25 kDa Protein: HET-s(1-227)

Anne Schuetz,[a] Christian Wasmer,[a] Birgit Habenstein,[b] René Verel,[a] Jason Greenwald,[a] Roland Riek,[a] Anja Böckmann,*[b] and Beat H. Meier*[a]

The sequence-specific resonance assignment of a protein forms the basis for studies of molecular structure and dynamics, as well as to functional assay studies by NMR spectroscopy. Here we present a protocol for the sequential ^{13}C and ^{15}N resonance assignment of uniformly [^{15}N,^{13}C]-labeled proteins, based on a suite of complementary three-dimensional solid-state NMR spectroscopy experiments. It is directed towards the application to proteins with more than about 100 amino acid residues. The assignments rely on a walk along the backbone by using a combination of three experiments that correlate nitrogen and carbon spins, including the well-dispersed C^{β} resonances. Supplementary spectra that correlate further side-chain resonances can be important for identifying the amino acid type, and greatly assist the assignment process. We demonstrate the application of this assignment protocol for a crystalline preparation of the N-terminal globular domain of the HET-s prion, a 227-residue protein.

Introduction

Solid-state NMR spectroscopy has over the last few years developed to a point at which a breakthrough in the field of structure determination for insoluble peptides and proteins is forthcoming. Indeed, the first atomic-resolution 3D structures determined from solid-state NMR spectroscopic data have been presented recently.[1–8] So far, these proteins have all been smaller than 100 amino acid residues, and the progress towards the structure determination of larger membrane and fibrillar proteins will require the adaptation and extension of the current protocols. The sequential resonance assignment, that is, the correlation of the NMR frequencies with the nuclear spins of each amino acid in the primary sequence, is the basis of all further NMR spectroscopy studies, such as site-specific dynamics, interaction measurements and 3D structure determination.

A growing number of sequential resonance assignments of solid-state spectra from uniformly labeled proteins are available today.[9–17] Most of these assignments were achieved from 2D spectroscopy, but the resolution in such spectra severely limits the size of the proteins that can be addressed. Consequently, the few available assignments of larger proteins ($>$200 amino acids) were rendered feasible by employing specific isotope-labeling schemes, for example, reverse labeling of only a subset of spins,[18–21] or checkerboard labeling schemes,[22] in combination with three-dimensional spectroscopy.

Herein we describe an assignment protocol based on three-dimensional NMR spectra that can be applied to larger, uniformly [^{13}C,^{15}N]-labeled proteins. It could also be combined with specific labeling schemes and extended to even larger proteins. The protocol we present is based on a number of previously developed pulse schemes[23–29] that were chosen to provide optimum efficiencies for the three or more polarization-transfer steps involved. In addition, our approach takes advantage of recent advances in spectrometer technology and sample preparation to achieve the necessary resolution and sensitivity in the experiments described. High static magnetic fields, optimized probe heads, and advances in sample preparation, including optimized filling of the protein into the rotor,[30] have brought better resolution and a substantial increase of the signal-to-noise ratio. These advances currently allow for the recording of 3D spectra that include up to four transfer steps between heteronuclei within 3–5 days by using a sample of less than 25 mg of a 25 kDa protein.

The HET-s prion from the fungus *Podospora anserina* is involved in a genetically controlled cell-death reaction (heterokaryon incompatibility) and consists of two domains, a prion domain and a globular domain (1–227).[31] In the following, we study the globular domain, HET-s(1–227), of which the crystal structure has recently been solved.[32] The protein consists of eight α helices and three short 3_{10} helices, accounting for the vast majority of residues. It has very little β-sheet content with only two strands of three residues each. The rather uniform secondary structure content makes it a challenging system for

[a] A. Schuetz, C. Wasmer, Dr. R. Verel, Dr. J. Greenwald, Prof. R. Riek, Prof. B. H. Meier
Physical Chemistry, ETH Zürich
8093 Zürich (Switzerland)
E-mail: beme@ethz.ch

[b] B. Habenstein, Dr. A. Böckmann
Institut de Biologie et Chimie des Protéines,
UMR 5086 CNRS/Université de Lyon 1
7 passage du Vercors, 69367 Lyon (France)
E-mail: anja.bockmann@ibcp.fr

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/cbic.201000124.

solid-state NMR spectroscopy and a good model for studies of α-helical membrane proteins.

In the following, the solid-state NMR spectroscopy protocols and the rationale behind the experiments are discussed, and it is demonstrated how they can be implemented to maximize resolution and sensitivity. We illustrate how the near-complete backbone assignments and 68% of side-chain assignments were obtained by using only a uniformly labeled HET-s(1-227) protein sample and five different 3D spectra that were recorded in 3–5 days each. Spectral ^{13}C linewidths for the protein preparation studied are around 0.5 ppm. One-dimensional and two-dimensional homo- and heteronuclear correlation experiments are shown in Figures 1B and 2, respectively. The INEPT experiment (Figure 1C) recorded on the HET-s(1–227) crystals is almost devoid of resonances; this indicates the absence of highly dynamic stretches in the protein. We compare the chemical shifts in solution to those in crystals and analyze the extent to which solution-state chemical shifts, as well as the X-ray structure, can speed up the assignment process for solid-state spectra.

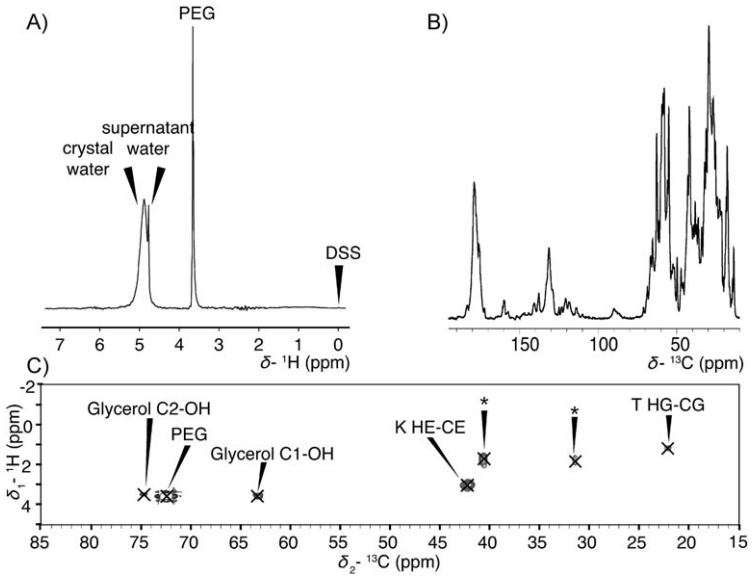

Figure 1. A) One-dimensional ^{1}H NMR spectrum showing the broad resonance from crystal water and a minor, sharp line from supernatant water.[30] The resonance at 3.75 ppm comes from the crystallization agent PEG 4000. B) 1D ^{13}C-CP-MAS spectrum, 64 scans. C) INEPT spectrum revealing highly dynamic residues. No backbone resonances are detected, only flexible side chains. The two correlations marked with an asterisk do not correspond to random-coil chemical shifts and are not assigned. The spectra were recorded at 850 MHz, 18 kHz MAS (A and B) and 600 MHz, 13 kHz (C).

The assignment concept and its implementation

The characteristics of solid-state NMR spectroscopy require assignment protocols that deviate considerably from the ones used in solution-state NMR. Two major differences must be accounted for: 1) The resolution in the proton (^{1}H) dimension of NMR spectra is much reduced in the solid state compared with solution, with the exception of highly deuterated samples,[33] making it more attractive to record spectral dimensions representing heteronuclei (^{15}N,^{13}C). 2) Whereas in isotropic solutions, all polarization transfer steps must be based on scalar couplings, both scalar and dipolar couplings can be used in solids, and the latter are usually advantageous due to their considerably larger magnitude. The dipolar coupling decreases with the third or a higher power of the internuclear distance. Due to the phenomenon of dipolar truncation,[34] the polarization transfer in recoupled experiments in uniformly labeled systems is practically limited to one-bond transfers; this makes it particularly well suited for assignment experiments.

As a consequence a dipolar N–C^{α} polarization transfer is unidirectional, meaning that it takes place virtually only between the amide nitrogen and the C^{α} of the same residue. In contrast, the *J*-based N–C^{α} transfer in solution-phase NMR spectroscopy is bidirectional because it connects the ^{15}N resonance with the $^{13}C^{\alpha}$ chemical shifts of both residues *i* and *i*−1, and the magnitudes of the $^{1}J_{NC\alpha}$ and $^{2}J_{NC\alpha}$ couplings are comparable. In solution, polarization can be transferred in a second step to the C^{β} spins through the large $^{1}J_{C\alpha C\beta}$ couplings, resulting in an HNCACB spectrum. From the information contained in this kind of spectrum it is possible to link each residue to its predecessor. Because most C^{α}–C^{β} and H^{N}–N chemical-shift pairs are unique (i.e., the H^{N}–N and C^{α}–C^{β} correlation spectra are resolved), a single bidirectional HNCACB spectrum is often sufficient to assign the entire backbone of a protein. In practice, several spectra with partially redundant information, like a set of HNCA, HNCACB, and HNCOCA, usually have to be recorded.[35–38]

In solid-state NMR spectroscopy, in which unidirectional experiments are used, a combination of 3D NCOCA, NCACO, and CANCO experiments is, given sufficient spectral resolution, sufficient for a backbone walk. Note that all three experiments involve two polarization transfer steps and each combination of resonances appears in two experiments. For crowded spectra, this scheme has, however, the drawback in that it uses a C′ dimension with a relatively low chemical-shift dispersion (compared to C^{α} and C^{β}). A comparison of the resolving power of different resonance pairs involving C^{β}, C^{α}, N, and C′ is shown in Figure 3.

As a consequence, an alternative set of three 3D experiments that use a C^{β} dimension instead of a C′ dimension is proposed: namely NCACB, N(CO)CACB and CAN(CO)CA (Figure 4A). The set also involves only experiments with one-bond polarization-transfer periods, either 2 or 3, and also leads to connectivities in which each pair of spins is correlated in two different experiments, which is important to resolve spectral overlap. In addition to this set of experiments, supplementary schemes that correlate the N, C^{α}, and C^{β} resonances with further side-chain resonances are useful, not only for the assignment of the side chains (important for 3D structure determination), but for the identification of amino acids, thus providing

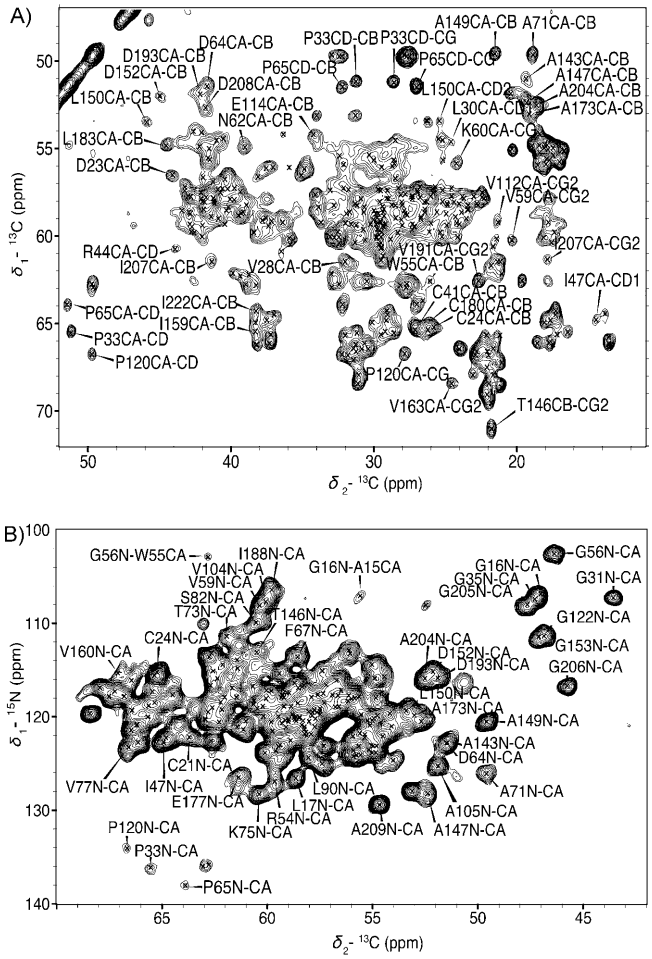

Figure 2. Two-dimensional spectra for initial sample characterization of crystalline HET-s(1–227) recorded at 850 MHz ^{1}H frequency, 18 kHz MAS.
A) Close-up of the aliphatic region of a ^{13}C–^{13}C DARR/MIRROR spectrum with 100 ms mixing time revealing intraresidual correlations of the C$^{\alpha}$ spins.
B) NCA spectrum revealing intraresidue (and very few interresidue) NCα correlations with the C$^{\alpha}$ spins. The peak positions marked by crosses are taken, in both spectra, from the final chemical-shift assignment listed in Table S2. Some isolated peaks with a unique assignment are labeled.

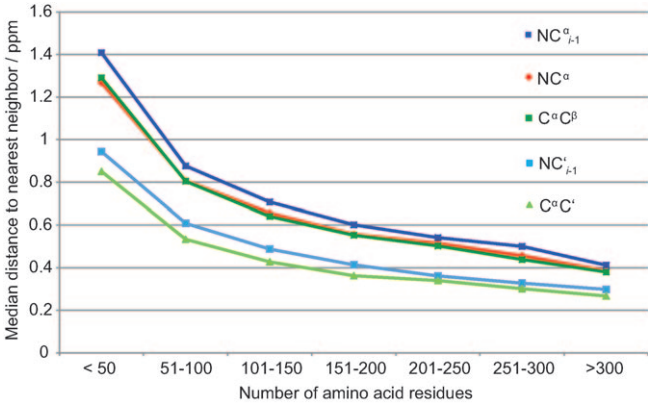

Figure 3. Statistics of the median distance (*d*) in ppm between nearest-neighbor peaks in generic two-dimensional spectra as a function of the size (number of residues) of the protein. Resonance frequencies were taken from assignments deposited in the BMRB.[45] Only nonparamagnetic proteins with extended stretches with a complete assignment (N^{H}, C^{α}, C^{β}, and C′) were taken into account. Amongst the different bonded spin pairs in the backbone, correlations involving C′ clearly have the lowest spectral dispersion. The $N_iC^{\alpha}_{i-1}$ spin pair has the best dispersion followed by NC^{α}, and $C^{\alpha}C^{\beta}$, which are about equally dispersed. We can use these trends to judge the dispersion in the different 3D spectra. The NCACO, NCOCA and CANCO experiments rely on NC^{α}, $C^{\alpha}C'$ and $C'_{i-1}N_i$ for dispersion whereas the set of experiments without C′ dimension, NCACB, N(CO)CACB, and CAN(CO)CA, rely on NC^{α}, $C^{\alpha}C^{\beta}$, and $C^{\alpha}_{i-1}N_i$. The second set of experiments is therefore expected to give spectra with a significantly higher spectral resolution. This comparison assumes that ^{15}N and ^{13}C linewidth are roughly equal in ppm which is the case in our sample.

starting points for the sequential assignment while also resolving assignment ambiguities. 3D experiments for side-chain assignments ideally consist of two well-dispersed dimensions with known chemical shifts (N, C^{α}, C^{β}) and a third one in which new side-chain resonances can be identified. Different transfer schemes are sketched in Figure 4 B, namely ^{15}N-resolved (N(CA)CBCX) and triple ^{13}C correlation (CXCYCZ, in the following referred to as CCC) experiments.[39,41]

Results and Discussion

Polarization-transfer sequences

The pulse sequences used to record the NCACB, N(CO)CACB and CAN(CO)CA spectra are shown in Figure 5 A–C, planes from the corresponding spectra are given in Figure 6. For the C′-resolved set of experiments, pulse sequences and planes are provided in Figure S1 and S2 in the Supporting Information.

The pulse sequences are all combinations of an adiabatic cross-polarization heteronuclear transfer[24] and the following homonuclear transfer schemes: DREAM[27] for C^{α}–C^{β} transfer and MIRROR[29] or DARR[41] optimized for C′–C^{α} transfer. The polarization-transfer steps, which constitute the building blocks of the experiments, have been chosen with regard to their high efficiency and specificity of the required transfer. The relative signal-to-noise achievable in these experiments can be estimated based on the experimental transfer efficiencies for the isolated building blocks in relation to a proton–carbon cross-polarization step.[42] The following estimates for the transfer steps involved were obtained for HET-s(1–227). A high-power proton–carbon cross-polarization step with a ^{1}H field at ~75 kHz, and ^{13}C at ~56 kHz, was used as a reference experiment for all comparisons and set to 100% transfer efficiency. A ^{1}H–^{15}N CP followed by a ^{15}N–^{13}C CP on C^{α} or C′ spins was found to have roughly 30% of the intensity on the desired spin compared to the direct CP optimized for the transfer to C^{α}/C′. A DREAM optimized for transfer between C^{α} and side-chain nuclei has an integral of 30% in side-chain intensity (defined as resonances below 40 ppm) with varying transfer efficiency for different spin systems. For the spectrally isolated alanine region, the C^{α}–C^{β} transfer shows ~75% efficiency, whereas for isoleucine $C^{\delta1}$ only about 10% transfer efficiency was found. A MIRROR transfer optimized for C′–C^{α} correlation has 35% of the intensity on the desired spin compared to a direct proton–carbon CP optimized for intensity on C^{α}/C′. Based on the effieciencies of the individual transfer steps, the N(CO)CACB and CAN(CO)CA experiments are less sensitive by a

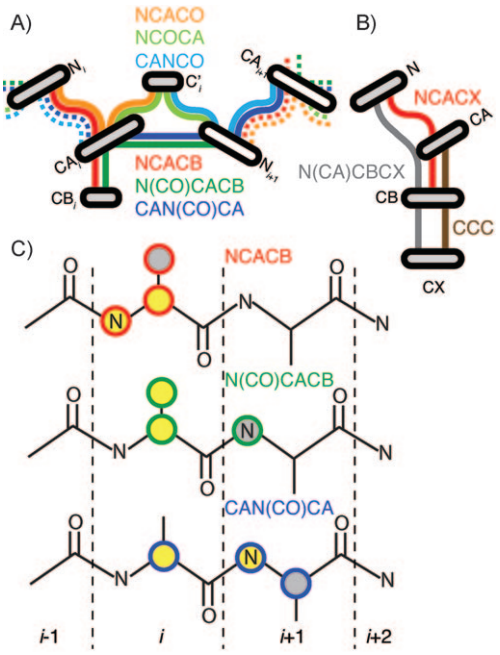

Figure 4. A) Schematic overview of the three-dimensional correlation experiments used for the sequential backbone resonance assignment in this study. Only the desired coherence transfers are shown. Minor pathways into the side chains, which in fact can be beneficial for assignment, are omitted for clarity. The transfer steps needed to walk sequentially from residue (*i*) to residue (*i*+1) are highlighted in bold lines. Indirect or direct acquisition is done on the spins in circles. The experiments can be classified into two categories: those containing a carbonyl-resolved dimension and those with an additional transfer step and no evolution period on the carbonyl spins.
B) Scheme of 3D experiments for side-chain resonances: N(CA)CBCX, NCACX, and CCC. CX is an arbitrary side-chain carbon spin. C) Assignment strategy by using the NCACB, N(CO)CACB, and CAN(CO)CA experiments. For a sequential walk, two known frequencies (highlighted in yellow) are always used to connect to the next spin (highlighted in grey).

factor of ~3 when compared to the NCACB, which is expected to show a transfer efficiency of ~9%. Concerning side-chain correlations, the corresponding pulse sequences are shown in Figure 5D, E. The N(CA)CBCX experiment is based on an NCACB, but with an additional DREAM-transfer appended that has been optimized for transfer from C^{β} further into the aliphatic side chain. Instead of the C^{α} chemical shifts, those of C^{β} are recorded in the second indirect dimension, because they are better dispersed.[43] The third dimension then shows correlations further out the side chain. The triple ^{13}C correlation experiment CCC is implemented by a selective CP on the aliphatic resonances (to reduce the spectral widths to be sampled in the indirect dimensions), followed by a DREAM transfer correlating the aliphatics among each other and then a long spin-diffusion period to mix with further side-chain spins. Whereas the DREAM-based experiments correlate only the aliphatic resonances (in this experimental setup by using 18 kHz MAS at 850 MHz ^{1}H frequency), the DARR or MIRROR mixing step potentially transfers polarization to all close-by ^{13}C nuclei, revealing precious information from the aromatic and carbonyl regions. A NCACX has also been recorded in which a DARR transfer replaces the DREAM transfer in Figure 4A to further distribute the polarization out to the side chains.

Sequential assignments

As starting points for the sequential walk, well-isolated cross-peaks are chosen in the NCACB spectrum from residues of which the amino-acid type can be identified from the unique chemical shifts of N, C^{α}, and C^{β} (i.e., Gly, Ala, Ser, and Thr). For other amino acid residues, the fact that more resonances from a spin system are often visible in the NCACB spectrum due to multi-bond transfers during the DREAM helps to identify, for example, Val, Ile, Leu, Arg, Lys, and Glu residues. Also, the CCC or NCACX spectra, revealing side-chain connectivities, are useful to identify spin systems. The N_{i+1} chemical shift can then be identified in the N(CO)CACB spectrum by matching the often well-resolved C^{α}–C^{β} resonance pair. This experiment is the most insensitive in the set of three (see Figure 4A) but excellent in terms of spectral dispersion. Next, the C^{α}_{i+1} is derived from the CAN(CO)CA by matching the $N_iC^{\alpha}_i$ pair. Now the procedure comes full circle and the (*i*+1) amino acid is identified in the NCACB by finding the C^{β}_{i+1} chemical shift. If the present (*i*) and the subsequent (*i*+1) residue can be assigned to a specific amino acid type, and if this combination of amino acid residue types is unique within the protein sequence, the segment can be assigned sequentially.

To give an idea of how this procedure works in practice, a walk along four residues in the protein is given in Figure 7. In each strip in the vertical dimension, the new previously unknown resonance is identified. Wherever there is more than one peak on the vertical line in the center of the strip there is ambiguity. This can be resolved if 1) the expected amino acid residue type is known for the next residue and has characteristic chemical shifts (as expected from BMRB[44] statistics or SPARTA[45] or 2) all potential assignment pathways are further examined until all but one become inconsistent. In the case of HET-s(1–227), the solution-state chemical shifts were also available as a guide. As a concrete example illustrating the encountered degree of ambiguities, in Figure 7 in the fourth strip, the sequential connection from V22 to D23 shows at least two possible C^{β} shifts, with C^{β} of D23 either around 40 ppm or 44 ppm. This ambiguity can be resolved immediately by comparing it to solution-state chemical shifts. If these were unavailable, both possible C^{β} shifts would have to be further investigated. At this point it can be helpful to deduce the spin systems of both N–C^{α}–C^{β} resonance triples from the side-chain assignment experiments (see Figure 8). If the assignment remains ambiguous, further steps in the walk for both possibilities should provide the correct connection. An alternative representation of a strip plot, more of the type often shown in solution-state NMR spectroscopy, is given in Figure S3. The representation of Figure 7 has the advantage though that it gives an immediate impression about the ambiguity encountered in the step.

Representative planes from N(CA)CBCX and CCC spectra for side-chain assignments are shown in Figure 8A, B. By analyzing these 3D spectra, 68% of the carbon side-chain resonances could be assigned. This number is limited by the poor visibility of likely flexible parts of the side chains often encountered in Arg, Lys, and His residues, or the poor efficiency of the longer-

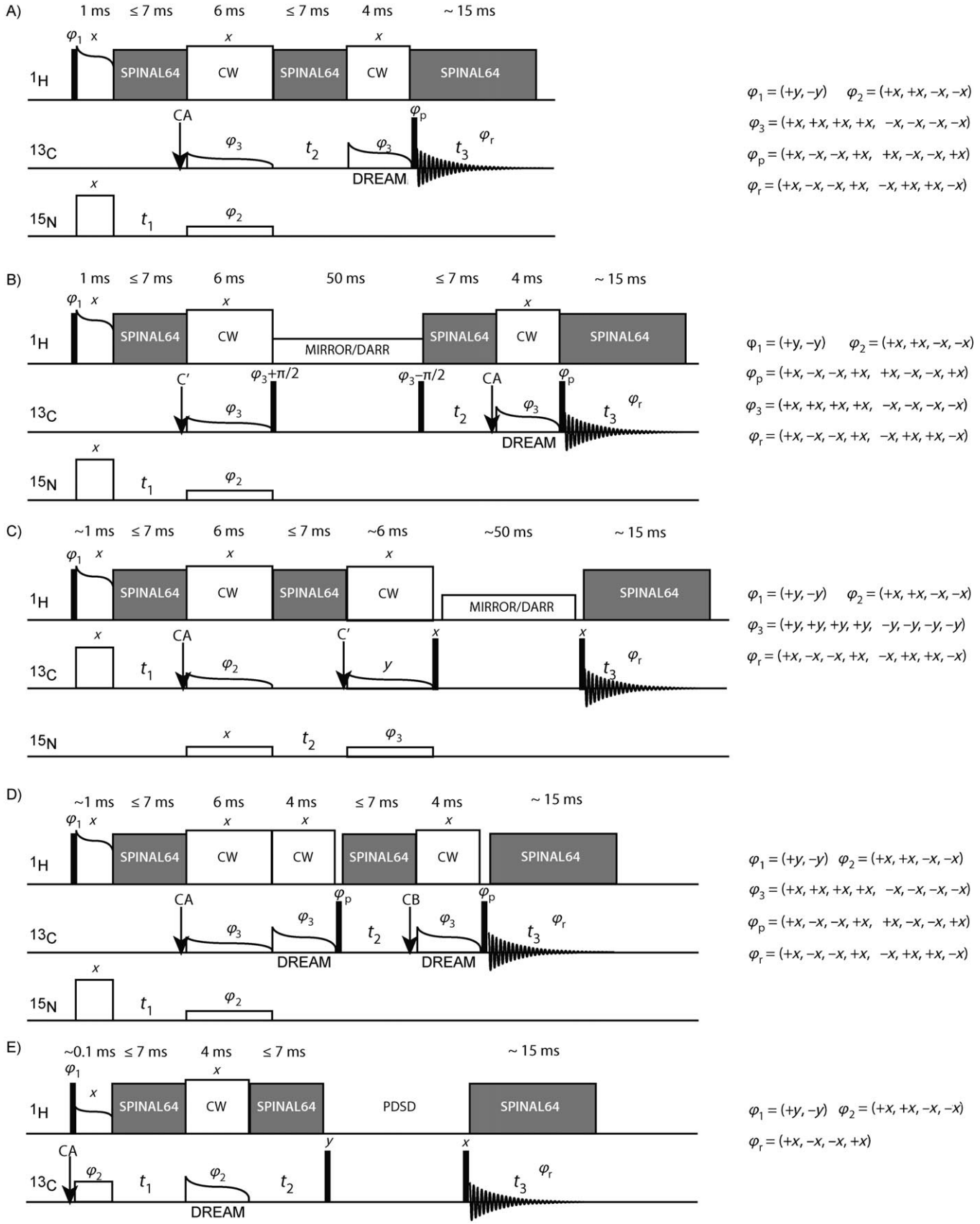

Figure 5. Schematic representation of the pulse sequences NCACB, N(CO)CACB, CAN(CO)CA, N(CA)CBCX and CCC (from A–E). The pulse sequences for the NCACO, NCOCA and CANCO are given in Figure S1. All spectra were recorded with the phase cycle indicated. The arrows with atom names symbolize the position to which the ^{13}C carrier frequency was set: CA corresponds to 56.5 ppm, CB to 42 ppm, C′ to 176.5 ppm. The lengths for transfer periods and the maximum lengths for chemical-shift evolution periods are indicated on top of each pulse sequence. Detailed information including the radiofrequency field strengths, pulse lengths, carrier positions and acquistion times are listed in Table S4.

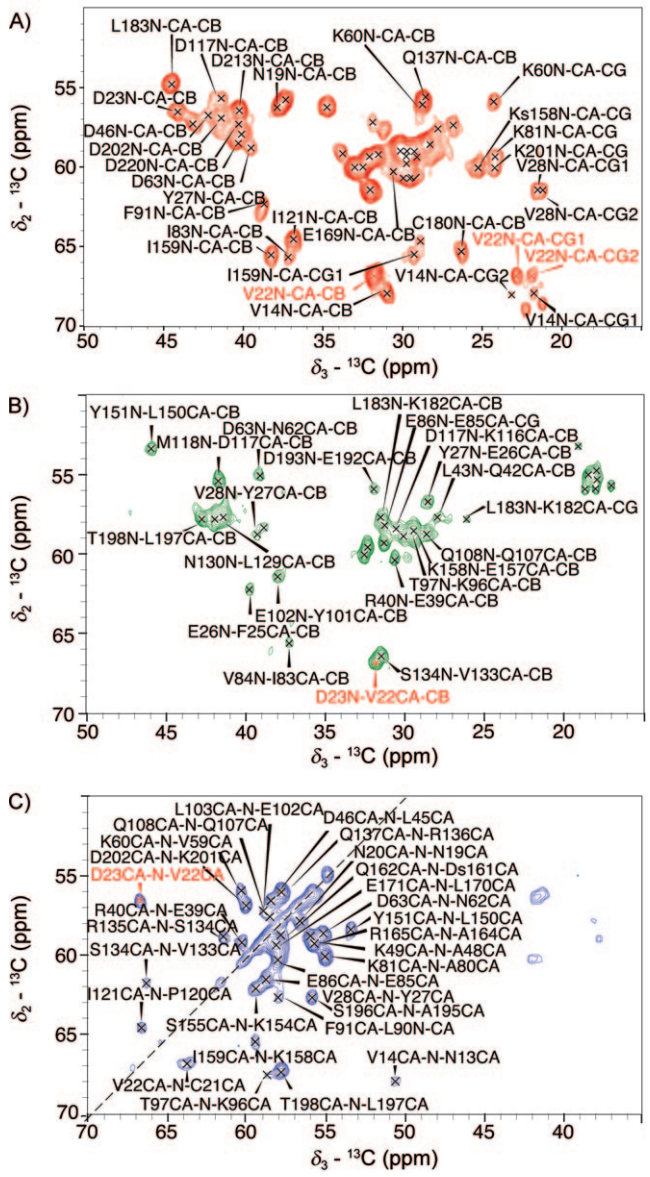

Figure 6. Representative planes from the three 3D spectra used for the backbone assignment. All peaks correspond to sequential correlations picked in the spectra; the assignment is given where space permits. As two representative amino acid neighbors, the correlations used to establish sequential contact from V22 to D23 are highlighted in red. A) δ_2-δ_3-Plane of the NCACB spectrum at $\delta_1 = 117.8$ ppm. The spin system of V22 is highlighted. B) δ_2–δ_3 plane of the N(CO)CACB spectrum at $\delta_1 = 117.1$ ppm. The connection of the V22 $C^{\alpha}C^{\beta}$ resonance pair to the N-resonance of D23 is highlighted. C) δ_1–δ_3 Plane of the CAN(CO)CA spectrum at $\delta_2 = 117.1$ ppm. The connection V22 CA to the NC^{α}-resonance pair of D23 is highlighted.

distance transfer, for example between methionine C^{γ} and C^{ε}. Figure 8C shows a more detailed representation of the completeness of the assignments for the various amino acid types and the statistics itemized by atom types are given in Table S1.

In the solid-state NMR spectra, 13 residues at the N terminus and four residues at the C terminus are invisible. Of the remaining 210 residues, a sequence-specific backbone assignment (N, C^{α}, C^{β}) was obtained for 204 residues (90%). The resonances for residues P4 to A11 are sensitive to the proline *cis–trans* isomerization in solution and are not visible in the solid-state NMR spectra. A gap in the assignment concerns residues K138 to S141, which are invisible also in the INEPT spectra that probe the highly flexible residues,[46] and the adjacent residues on both sides of the gap are broadened. This stretch of residues is part of a flexible loop, as also seen in the X-ray structure and in solution (vide infra). The other gaps are only single missing residues, which are located in stretches with generally low signal intensity. The extent of the assignment is shown on the protein primary and secondary structure in Figure 9 (and compared with the extent of assignment of our solution NMR data in Figure S4). Additionally, 148 carbonyl resonances (65%) were assigned from the C′-resolved set of spectra (shown in Figure S2). A comprehensive list of all assigned resonances is given in Table S2. Concerning the assignments of HET-s(1–227) in solution, 203 out of 227 residues (89%) could be assigned sequentially by using standard procedures. Conformational exchange is evidenced for residues P4 to A11 in the form of a proline *cis–trans* isomerization, and from residues T140 to L150 located in an extended loop. A side-chain assignment was not attempted in solution. Full assignments are given in Table S3. Assignments have been deposited in the BMRB (accession numbers 16964 and 16965 for solid- and solution-state, respectively) and the differences between the solution and solid-state chemical shifts are presented in Figure S5 and S6.

Possible support by solution-state NMR spectroscopy and chemical-shift predictions

The two almost complete backbone-resonance assignments allow a comparison of solution and solid-state chemical shifts. Solution-state chemical shifts can present a valuable help in assigning solid-state NMR spectra. To assess how many resonances could have been guessed directly from the solution data, we attempted to automatically assign the solid-state NCACB spectrum from the solution chemical shift values. Fifty-five peaks unambiguously match the N-C^{α}-C^{β} peaks in all three dimensions within a chemical-shift tolerance given by twice the RMSD between the solid and solution chemical shifts (N, 0.9; C^{α}, 0.5; C^{β}, 1.2 ppm). Out of these, however, twelve (22%) were incorrect suggestions; this makes a direct use of the solution-state shift less attractive. In addition to peak assignments, knowledge of the solution-state chemical shifts can be used to resolve ambiguities in the backbone walk in 3D spectra, and some assignment possibilities can be tentatively excluded at an earlier stage.

For proteins with known X-ray structure but no solution-state NMR spectroscopic assignments, it might help to use chemical-shift predictions, for example, by SPARTA,[45] to assist assignments. For HET-s(1–227), the accuracy of the SPARTA predictions is characterized by an rmsd of 2.5 ppm for nitrogen and 1.0 ppm for carbon resonances, close to the values reported before.[47] With a chemical-shift tolerance set to twice the RMSD, only three out of 13 cross peak assignments were suggested correctly for the NCACB assignments. Table S3 gives an overview of the residues guessed correctly from solution-state shifts, along with the SPARTA prediction for these residues.

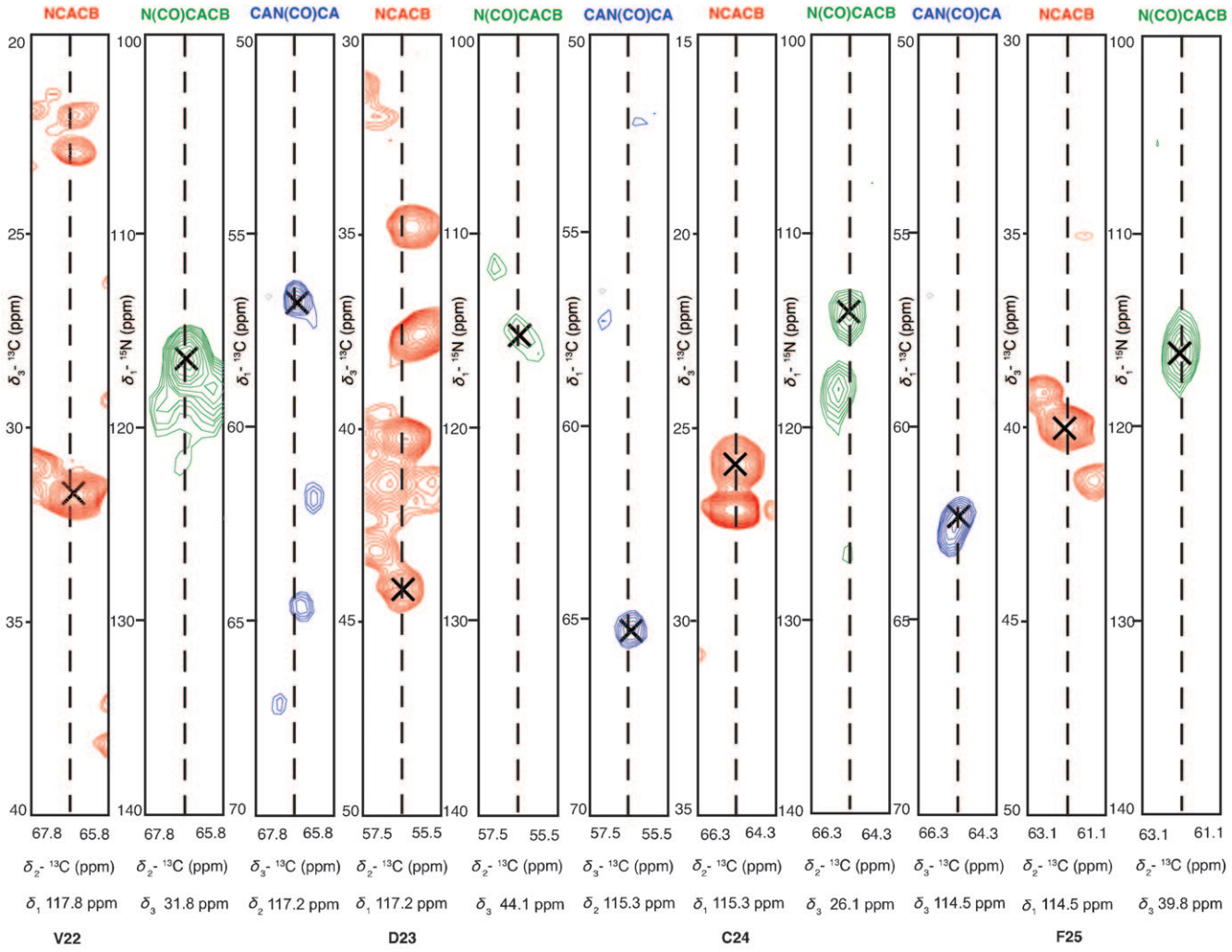

Figure 7. Sequential walk along the protein backbone by using NCACB, N(CO)CACB, and CAN(CO)CA experiments in the direction from the N to C terminus by following the scheme displayed in Figure 4A. For each strip, the previously unknown resonance (highlighted in gray in Figure 4C) is displayed in the vertical dimension. The potential ambiguity of an assignment can be judged from the number of peaks on the dashed line. An alternative representation of the same data is given in Figure S3.

Conclusions

We have devised protocols for the sequential resonance assignment in proteins that have crowded 2D solid-state NMR spectra; our method is applicable to proteins with sizes larger than 100 amino acid residues. We have validated the approach for the 25 kDa crystalline HET-s(1–227) protein. The sequence-specific assignment was obtained from a uniformly labeled sample by using a set of three 3D experiments relying on N, C^{α} and C^{β} spins for dispersion (NCACB, N(CO)CACB, CAN(CO)CA). Side-chain resonances have been identified by using two additional experiments; N(CA)CBCX and CCC. Each of these spectra was acquired in less than five days of measurement time on an 850 MHz NMR spectrometer. The sensitivity is generous for the NCACB and just sufficient for the least sensitive experiment, the N(CO)CACB. Assignments are 90% complete for the backbone resonances and are limited by the presence of dynamics mainly in the N- and C-terminal region. Knowledge of solution-state chemical shifts facilitates the assignment; they can be used when the solid-state spectra are ambiguous and provide an independent control. However, assignments without this information should be feasible and we are currently testing the protocol on proteins with unknown solution-state chemical shifts.

We estimate that the current protocol works for proteins of up to 300–400 amino acids. Larger proteins will demand higher sensitivity and resolution, which can be provided by higher magnetic fields, selective labeling schemes, improved polarization transfer schemes and hyperpolarization.

Experimental Section

Sample preparation for solution-state NMR: Perdeuterared [^{2}H,^{15}N,^{13}C] HET-s(1–227) was recombinantly expressed and purified as described previously.[31] The purified protein was dialysed against a buffer containing H_2O/ D_2O (9:1), NaCl (50 mM), sodium phosphate (20 mM), DTT (2 mM) at pH 7 and concentrated to 1 mM for solution NMR spectroscopic measurements.

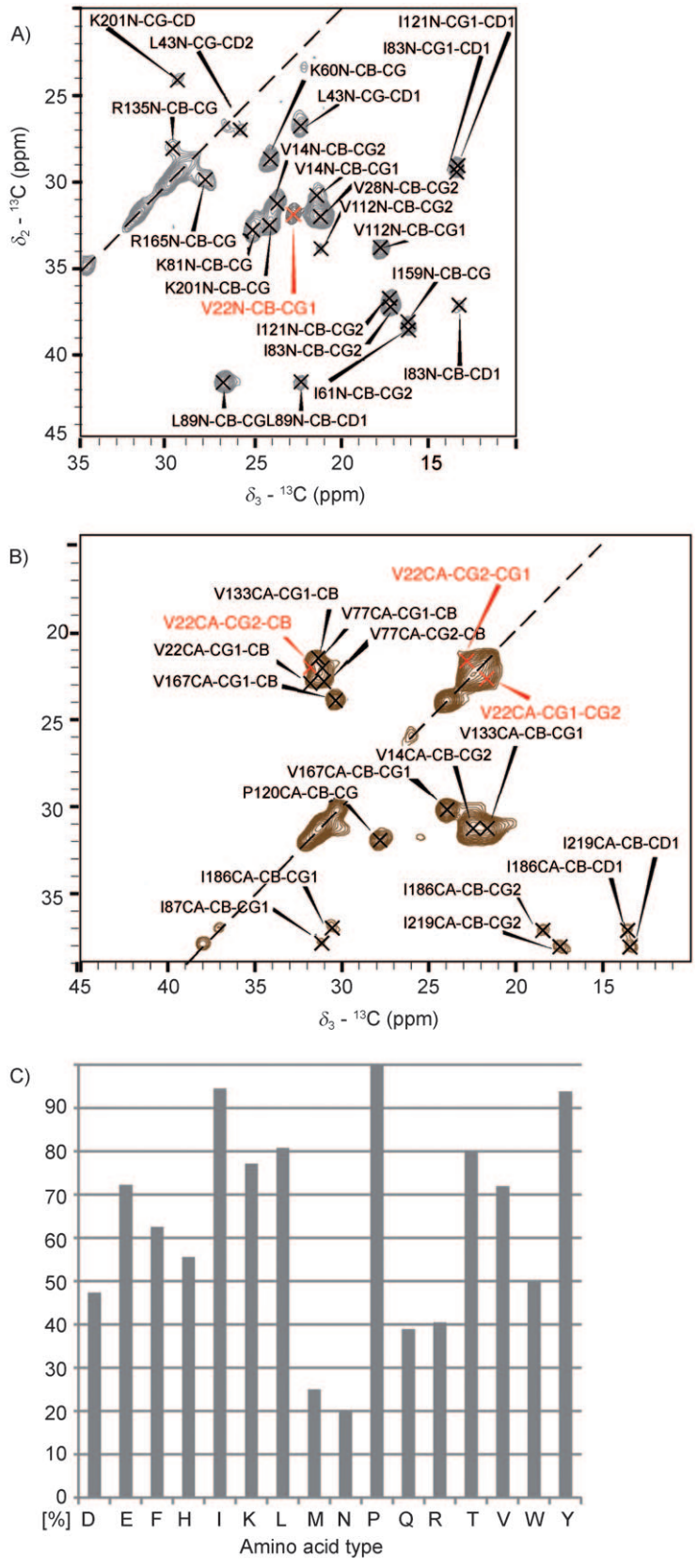

Figure 8. A) Representative $\delta_2-\delta_3$ plane from the N(CA)CBCX spectrum at $\delta_1 = 117.8$ ppm. The peak corresponding to V22 is highlighted in red. B) $\delta_2-\delta_3$ plane from the CCC spectrum at $\delta_1 = 66.8$ ppm. The spin system of V22 is highlighted in red. C) Statistics about side-chain assignments according to amino acid type. For each amino acid, the percentage of assigned side-chain carbon atoms (all except C^{α}, C^{β}, C′) for all assigned residues is given. More detailed statistics itemized by amino acid and atom type are given in Table S2.

Solution-state NMR spectroscopy: All solution-state NMR spectra were recorded on a Bruker Avance 700 spectrometer operating at a static magnetic field of 16.5 T (700 MHz ^{1}H frequency) equipped with a triple resonance cryoprobe and pulsed-field gradients. Measurements were performed on a perdeuterated sample at 297 K. Assignments of HET-s(1–227) were done by using standard TROSY-based[48] solution-state NMR spectroscopy experiments, including HNCACB,[36] HNCA,[35] and HNCOCA.[37] All spectra were processed in TopSpin 2.0 (Bruker Biospin) by zero-filling to no more than double the number of acquired points, multiplied with a squared cosine window function, and Fourier transformed. The spectra were analyzed by using Sparky version 3.115 (T. D. Goddard and D. G. Kneller, University of California, San Francisco).

Sample preparation for solid-state NMR spectroscopy: Uniformly [^{15}N,^{13}C]-labeled HET-s(1–227) was recombinantly expressed as described previously.[31] Crystals thereof were prepared by starting from 34 mg of protein in a 17 mg mL^{-1} protein solution in Tris–HCl (20 mM, pH 8.5) DTT (0.5 mM) and NaN_3 (0.02 %). Crystals were grown in sitting drops of 150 μL by adding precipitation solution (150 μL) containing PEG 4000 (30 %) in Tris-–HCl (20 mM, pH 8.5). The reservoir contained a NaCl solution (2 M, 20 mL). Abundant precipitates appeared within one week and were harvested and centrifuged into a 3.2 mm Bruker rotor at 5000×*g*, followed by a final centrifugation at 84200×*g* for 30 min in an ultracentrifuge.[30] The proton spectrum (Figure 1 A) shows only a very weak supernatant water line, and a strong broad crystal water resonance. The 1D ^{13}C NMR spectrum is shown in Figure 1 B.

Solid-state NMR: The auxiliary INEPT and CANCO experiments were recorded on a Bruker Avance 600 spectrometer operating at a static magnetic field of 14.1 T by using a 2.5 mm Chemagnetics triple resonance probe at a MAS spinning frequency of 13 kHz, at 20 °C (INEPT) and 10 °C (CANCO). All other solid-state NMR spectra were recorded on a Bruker Avance II+ 850 MHz spectrometer operating at a static magnetic field of 20.0 T. A 3.2 mm Bruker triple-resonance probe equipped with an LLC coil was used to reduce radio frequency (RF) heating during the experiments. The spectra were recorded at 17.5–19 kHz MAS at sample temperatures of about 10 °C as determined by using the water resonance.[49] Details of experimental parameters are given in Table S4. The NCACB spectrum was internally referenced to DSS and all other spectra were subsequently referenced to an isolated resonance in this spectrum. All spectra were processed in TopSpin 2.0 (Bruker Biospin) by zero filling to a power of two but not larger than twice the number of points measured, multiplied with a shifted squared cosine window function, and Fourier transformed. Automatic polynomial baseline correction was applied in the direct dimension. The spectra were analyzed by using Sparky version 3.115 (T. D. Goddard and D. G. Kneller, University of California, San Francisco).

Figure 9. Extent of solid-state NMR backbone assignments displayed on the protein primary sequence. The secondary structure (residue 11–224) elements[32] are given on top. Sequentially assigned residues are given in black. The GS prior to the initial M residue is not part of the protein primary sequence but remainders of the cleaved His-tag from protein purification.

Acknowledgements

This work has been supported by the ANR (ANR-PCV08_321323), the Swiss National Science Foundation, the ETH Zürich and the CNRS.

Keywords: prions · proteins · sequence assignment · solid-state NMR spectroscopy · structure determination

[1] F. Castellani, B. van Rossum, A. Diehl, M. Schubert, K. Rehbein, H. Oschkinat, *Nature* **2002**, *420*, 98–102.
[2] S. G. Zech, A. J. Wand, A. E. McDermott, *J. Am. Chem. Soc.* **2005**, *127*, 8618–8626.
[3] A. Lange, K. Giller, S. Hornig, M. F. Martin-Eauclaire, O. Pongs, S. Becker, M. Baldus, *Nature* **2006**, *440*, 959–962.
[4] W. T. Franks, B. J. Wylie, H. L. Frericks Schmidt, A. J. Nieuwkoop, R. Mayrhofer, G. J. Shah, D. T. Graesser, C. M. Rienstra, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 4621–4624.
[5] C. Gardiennet, A. Loquet, M. Etzkorn, H. Heise, M. Baldus, A. Böckmann, *J. Biomol. NMR* **2008**, *40*, 239–250.
[6] A. Loquet, B. Bardiaux, C. Gardiennet, C. Blanchet, M. Baldus, M. Nilges, T. Malliavin, A. Böckmann, *J. Am. Chem. Soc.* **2008**, *130*, 3579–3589.
[7] T. Manolikas, T. Herrmann, B. H. Meier, *J. Am. Chem. Soc.* **2008**, *130*, 3959–3966.
[8] C. Wasmer, A. Lange, H. van Melckebeke, A. B. Siemer, R. Riek, B. H. Meier, *Science* **2008**, *319*, 1523–1526.
[9] J. Pauli, M. Baldus, B. van Rossum, H. de Groot, H. Oschkinat, *ChemBioChem* **2001**, *2*, 272–281.
[10] A. Böckmann, A. Lange, A. Galinier, S. Luca, N. Giraud, M. Juy, H. Heise, R. Montserret, F. Penin, M. Baldus, *J. Biomol. NMR* **2003**, *27*, 323–339.
[11] T. I. Igumenova, A. E. McDermott, K. W. Zilm, R. W. Martin, E. K. Paulson, A. J. Wand, *J. Am. Chem. Soc.* **2004**, *126*, 6720–6727.
[12] T. I. Igumenova, A. J. Wand, A. E. McDermott, *J. Am. Chem. Soc.* **2004**, *126*, 5323–5331.
[13] D. Marulanda, M. L. Tasayco, M. Cataldi, V. Arriaran, T. Polenova, *J. Phys. Chem. B* **2005**, *109*, 18135–18145.
[14] O. C. Andronesi, S. Becker, K. Seidel, H. Heise, H. S. Young, M. Baldus, *J. Am. Chem. Soc.* **2005**, *127*, 12965–12974.
[15] W. T. Franks, D. H. Zhou, B. J. Wylie, B. G. Money, D. T. Graesser, h. L. Frericks, G. Sahota, C. M. Rienstra, *J. Am. Chem. Soc.* **2005**, *127*, 12291–12305.
[16] A. B. Siemer, C. Ritter, M. O. Steinmetz, M. Ernst, R. Riek, B. H. Meier, *J. Biomol. NMR* **2006**, *34*, 75–87.
[17] A. Goldbourt, B. J. Gross, L. A. Day, A. E. McDermott, *J. Am. Chem. Soc.* **2007**, *129*, 2338–2344.
[18] H. Heise, W. Hoyer, S. Becker, O. C. Andronesi, D. Riedel, M. Baldus, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 15871–15876.
[19] M. Etzkorn, S. Martell, O. C. Andronesi, K. Seidel, M. Engelhard, M. Baldus, *Angew. Chem.* **2007**, *119*, 463–466; *Angew. Chem. Int. Ed.* **2007**, *46*, 459–462.
[20] R. Schneider, C. Ader, A. Lange, K. Giller, S. Hornig, O. Pongs, S. Becker, M. Baldus, *J. Am. Chem. Soc.* **2008**, *130*, 7427–7435.
[21] L. Shi, M. Ahmed, W. Zhang, G. Whited, L. Brown, V. Ladizhansky, *J. Mol. Biol.* **2009**, *386*, 1078–1093.
[22] V. Higman, J. Flinders, M. Hiller, S. Jehle, S. Markovic, S. Fiedler, B. Rossum, H. Oschkinat, *J. Biomol. NMR* **2009**, *44*, 245–260.
[23] M. Baldus, A. Petkova, J. Herzfeld, R. Griffin, *Mol. Phys.* **1998**, *95*, 1197–1207.
[24] M. Baldus, D. Geurts, S. Hediger, B. H. Meier, *J. Magn. Reson. Ser. A* **1996**, *118*, 140–144.
[25] R. Verel, M. Baldus, M. Ernst, B. H. Meier, *Chem. Phys. Lett.* **1998**, *287*, 421–428.
[26] S. Hediger, B. H. Meier, R. Ernst, *Chem. Phys. Lett.* **1995**, *240*, 449.
[27] R. Verel, M. Ernst, B. H. Meier, *J. Magn. Reson.* **2001**, *150*, 81–99.
[28] K. Takegoshi, S. Nakamura, T. Terao, *Chem. Phys. Lett.* **2001**, *344*, 631–637.
[29] I. Scholz, M. Huber, T. Manolikas, B. H. Meier, M. Ernst, *Chem. Phys. Lett.* **2008**, *460*, 278–283.
[30] A. Böckmann, C. Gardiennet, R. Verel, A. Hunkeler, A. Loquet, G. Pintacuda, L. Emsley, B. H. Meier, A. Lesage, *J. Biomol. NMR* **2009**, *45*, 319–327.
[31] A. Balguerie, S. Dos Reis, C. Ritter, S. Chaignepain, B. Coulary-Salin, V. Forge, K. Bathany, I. Lascu, J. Schmitter, R. Riek, S. J. Saupe, *EMBO J.* **2003**, *22*, 2071–2081.
[32] J. Greenwald, C. Buhtz, C. Ritter, W. Kwiatkowski, S. Choe, M.-L. Maddelein, F. Ness, S. Cescau, A. Soragni, D. Leitz, S. J. Saupe, R. Riek, **2010**; *submitted*.
[33] V. Chevelkov, K. Rehbein, A. Diehl, B. Reif, *Angew. Chem.* **2006**, *118*, 3963–3966; *Angew. Chem. Int. Ed.* **2006**, *45*, 3878–3881.
[34] M. Bayro, M. Huber, R. Ramachandran, T. Davenport, B. H. Meier, M. Ernst, R. G. Griffin, *J. Chem. Phys.* **2009**, *130*, 114506.
[35] L. E. Kay, M. Ikura, R. Tschudin, A. Bax, *J. Magn. Reson.* **1990**, *89*, 496–514.
[36] M. Wittekind, L. Mueller, *J. Magn. Reson. Ser. B* **1993**, *101*, 201–205.
[37] S. Grzesiek, A. Bax, *J. Magn. Reson.* **1992**, *96*, 432–440.
[38] M. Sattler, J. Schleucher, C. Griesinger, *Prog. Nucl. Magn. Reson. Spectrosc.* **1999**, *34*, 93–158.
[39] S. Li, Y. Zhang, M. Hong, *J. Magn. Reson.* **2009**, *196*, 39–47.
[40] Y. Tian, L. Chen, D. Niks, J. M. Kaiser, J. Lai, C. M. Rienstra, M. F. Dunn, L. J. Mueller, *Phys. Chem. Chem. Phys.* **2009**, *11*, 7078–7086.
[41] K. Takegoshi, S. Nakamura, T. Terao, *Chem. Phys. Lett.* **2001**, *344*, 631–637.
[42] A. Detken, E. H. Hardy, M. Ernst, M. Kainosho, T. Kawakami, S. Aimoto, B. H. Meier, *J. Biomol. NMR* **2001**, *20*, 203–221.
[43] Y. Wang, O. Jardetzky, *Protein Sci.* **2002**, *11*, 852–861.
[44] E. L. Ulrich, H. Akutsu, J. F. Doreleijers, Y. Harano, Y. E. Ioannidis, J. Lin, M. Livny, S. Mading, D. Maziuk, Z. Miller, E. Nakatani, C. F. Schulte, D. E. Tolmie, R. Kent Wenger, H. Yao, J. L. Markley, *Nucleic Acids Res.* **2008**, *36*, D402–408.
[45] Y. Shen, A. Bax, *J. Biomol. NMR* **2007**, *38*, 289–302.
[46] A. B. Siemer, A. A. Arnold, C. Ritter, T. Westfeld, M. Ernst, R. Riek, B. H. Meier, *J. Am. Chem. Soc.* **2006**, *128*, 13224–13228.
[47] K. Seidel, M. Etzkorn, R. Schneider, C. Ader, M. Baldus, *Solid State Nucl. Magn. Reson.* **2009**, *35*, 235–242.
[48] K. Pervushin, R. Riek, G. Wider, K. Wüthrich, *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 12366–12371.
[49] H. Gottlieb, V. Kotlyar, A. Nudelman, *J. Org. Chem.* **1997**, *62*, 7512–7515.

Received: February 25, 2010
Published online on June 22, 2010

