Published in final edited form as:
Angew Chem Int Ed Engl. 2007 ; 46(44): 8380–8383. doi:10.1002/anie.200702905.

Solid-State Protein Structure Determination with Proton-Detected Triple Resonance 3D Magic-Angle Spinning NMR Spectroscopy[**]

Donghua H. Zhou,
Department of Chemistry, University of Illinois at Urbana-Champaign, Urbana, IL 61801 (USA)

John J. Shea,
Department of Chemistry University of Illinois at Urbana-Champaign Urbana, IL 61801 (USA)

Andrew J. Nieuwkoop,
Department of Chemistry University of Illinois at Urbana-Champaign Urbana, IL 61801 (USA)

W. Trent Franks,
Department of Chemistry University of Illinois at Urbana-Champaign Urbana, IL 61801 (USA)

Benjamin J. Wylie,
Department of Chemistry, University of Illinois at Urbana-Champaign, Urbana, IL 61801 (USA)

Charles Mullen,
Varian, Inc., Fort Collins, CO

Dennis Sandoz, and
Varian, Inc., Palo Alto, CA

Chad M. Rienstra [Prof.][*]
Department of Chemistry, University of Illinois at Urbana-Champaign, Urbana, IL 61801 (USA)

Keywords

Solid-state NMR spectroscopy; protein structures; indirect proton detection; distance restraints; magic-angle spinning

Advances over the last decade in magic-angle spinning solid-state NMR (MAS SSNMR) have enabled the complete structure determination of several small proteins.[1] In principle, SSNMR is not limited by molecular size, which is one major advantage over solution NMR in challenging applications such as membrane protein complexes and high molecular weight protein aggregates. However, solid-state structure determination of larger proteins is typically hindered by the low sensitivity and relatively short measurable distances imposed by the observation of nuclei with low gyromagnetic ratios (γ), such as ^{13}C and ^{15}N. The large γ of ^{1}H, while providing high detection sensitivity and NOE distance restraints for solution NMR, [2] results in large dipolar couplings in the solid state, which may degrade both spectral resolution and sensitivity.[3] Recent studies by Reif and Zilm and their respective coworkers have demonstrated that these challenges in resolution and sensitivity can be overcome by using spin dilution, replacing all non-exchangeable protons with deuterons.[3] In combination with

[**]This research was supported by the National Institutes of Health (R01 GM-75937 to C.M.R). We thank Philippe Nadaud and Prof. Christopher Jaroniec (Ohio State University) for advice regarding expression of deuterated GB1, Mircea Cormos and John A. Stringer (Varian, Inc.) for advice on probe performance

[*]Fax: (+1) 217-244-4655, E-mail: rienstra@scs.uiuc.edu, http://www.scs.uiuc.edu/~rienstra.

Supporting information for this article is available on the WWW under http://www.angewandte.org or from the author.

high magnetic fields and ~20 kHz MAS, spin dilution has led to greatly improved resolution, significant sensitivity enhancement and long-range ^{1}H-^{1}H correlations.[3,4] Reif and coworkers have further obtained resolution rivalling solution NMR of larger proteins by back-exchanging with 10% H_2O and 90% D_2O.[5]

Here we investigate the potential of using ^{1}H-^{1}H distance restraints for solid-state protein structure determination. We prepared a sample of the β1 immunoglobulin binding domain of protein G (GB1) uniformly ^{13}C,^{15}N,^{2}H-labeled and back-exchanged with H_2O. The combination of spin dilution, high field (750 MHz), fast MAS (39 kHz) and triple resonance experiments yielded ^{1}H-detected spectra of very high resolution and sensitivity. Hundreds of ^{15}N- and ^{13}C-resolved ^{1}H-^{1}H distance restraints were obtained to determine a high-resolution structure, assisted only by empirical backbone dihedral angles from TALOS.[6]

Impressive signal-to-noise ratios (SNR) of 490±180 (average ± standard deviation) were obtained in a ^{1}H-detected ^{15}N-^{1}H 2D spectrum acquired within 30 min for only 0.9 μmol GB1 (Fig. 1). Line widths (Δ) were 140±30 Hz and 37±5 Hz for ^{1}H and ^{15}N, respectively. The intrinsic proton line width, calculated from the observed 7 ms overall spin-spin relaxation time (T_2'), was only 45 Hz. Magnetic field and sample heterogeneities account for the remainder of the line widths. Compared to a ^{15}N-detected ^{1}H-^{15}N 2D (Supporting Information Fig. S1), proton detection enhanced SNR by a factor of 18±3. To ensure a fair sensitivity comparison, each dimension was truncated to $3T_2^*$ ($T_2^*=1/\pi\Delta$) and processed without apodization. The scroll resonator construction of the probe used in this study is optimized for ^{1}H sensitivity; therefore, the enhancement may be slightly less for conventional solenoid designs.[7] However, even for this probe (*cf.* Experimental Section) there is room for further instrumental innovation. Compared to a ^{15}N-^{1}H 2D acquired with 20 kHz MAS, proton line widths were reduced by a factor of 1.9±0.6 (Fig. S2), indicating that at least up to ~40 kHz MAS the line width is still approximately linear with the inverse spinning rate,[8] and that faster spinning will benefit resolution further.

The ^{15}N-^{1}H 2D spectrum resolved 50 of 55 amide correlations, which could be readily assigned based on published ^{13}C and ^{15}N assignments.[9] Outlying peaks such as T49 and G14 were assigned based on unique ^{15}N chemical shifts. The large majority (46 of 55 backbone amides) of the peaks could be assigned based on the strongest peak in the 3D CON(H)H spectrum (Fig. 2), which correlates N[i] and HN[i] with C'[i-1]. This experiment is analogous to the HN(H)H method of Paulson and Zilm,[4] which was employed with slight modifications (Fig. S3); in both cases, the RFDR[10] recoupling method was used to enhance the rate of polarization transfer among protons. With 2 ms RFDR, sequential correlations (i.e., to HN[i±1]) were observed as peaks of lower intensity. Shorter ^{1}H-^{1}H mixing times in this 3D experiment may be beneficial to confirm assignments by virtue of spectral simplification in larger proteins, but were not necessary for GB1. The 2D CON projection from the 3D CON(H)H (Fig. S4) resolved all peaks except for four partially overlapping pairs (T51/T53, E19/Q2, F30/V29, and D36/Q32), which have distinct ^{1}H chemical shifts (Fig. 1). Furthermore, the HN(H)H spectrum (Fig. 2b) enabled independent confirmation of most assignments. Therefore, two 3D spectra together enabled unique assignments of all backbone amide protons.

Sequential, medium and long-range correlations were next assigned in these 3D spectra (2 ms RFDR) and additional N(H)H 2D spectra (2 and 3 ms RFDR). The strip plot (Fig. 2) illustrates one stretch from D46 to F52. For each C'[i-1]-N[i] frequency pair, in addition to the strong HN[i] peak, several weaker peaks are observed. Among these, many could be unambiguously assigned to long-range correlations (five or more residues away sequentially). For example, at the uniquely resolved F52 ^{15}N frequency (129.6 ppm), peaks are observed at K4, L5 and I6 ^{1}H frequencies, as well as the T51 sequential correlation (Fig. 3). At the unique K10 ^{1}H frequency (9.9 ppm), correlations to E56, L12, G9, and T11 amide ^{15}N frequencies are

observed. In addition to the amide protons, the W43 Hε (10.5 ppm) was correlated to amide ^{15}N frequencies of V39, D40, G41, E42 and W43. In total, ~300 well-resolved cross peaks could be uniquely assigned without making any assumptions about the structure.

These cross peaks were all assumed to arise from ^{1}H-^{1}H distances of less than 8 Å, a conservatively large estimate for initial rounds of calculations by XPLOR-NIH.[11] In addition to the distance lists, TALOS dihedral restraints were utilized during the simulated annealing and refinement calculations.[9] A bundle of ten lowest-energy structures out of 100 showed a consistent but poorly ordered fold with a backbone RMSD of ~2.5 Å. From this fold, additional peaks could be assigned that were ambiguous based on chemical shifts alone, but had only one possible coupling partner within 10 Å among candidates with similar proton chemical shift (Fig. S5). This process of iterative assignment could be automated for larger proteins.[12] In our case, we repeated the process manually for several iterations, enabling ~200 additional peaks to be assigned.

Next, the structure calculations were repeated assuming empirical, semi-quantitative relationships between the peak intensities and distances. The distance ranges were determined as follows. Among helical residues, ~30 HN[i]-HN[i±1] correlations were identified, with their average intensity $\langle I \rangle$ and standard deviation σ(I) calculated; peaks with intensity greater than $\langle I \rangle - \sigma(I)$ were assigned a 3.5-Å upper distance limit. Likewise, among β-sheet residues, ~50 HN[i]-HN[i±1] correlations were used to compute an intensity threshold, above which a 5.5-Å distance limit was assumed. These limits are based upon the known conformations of secondary structure elements, allowing for ~1 Å uncertainty. All other peaks in the spectra, of intensity less than the calibration points above, were assumed to correspond to distances 8.5 Å or less. Complete restraints lists are included in the supporting information (Tables S1–S4). The conservative upper limit on the distance range avoided violations arising from multi-spin transfer events, although the exact upper limit value had no significant bearing on the results of the structure calculations.

In the first round of calculations with explicit distance ranges, ~30 of the 517 restraints violated the prescribed range by more than 0.5 Å. For those restraints, the ranges were increased to the next longer category (e.g., from 3.5 to 5.5 Å) and the calculation repeated until convergence was achieved. The final set of calculations produced a structure (pdb 2JU6) with 0.82±0.14 Å backbone RMSD (Fig. 4) for the ensemble of 10 structures (of 252); the total atom RMSD was 1.71±0.17 Å. Comparison with a GB1 crystal structure (pdb 2GI9) yielded a backbone RMSD of 1.9 Å, when aligning the entire molecule (Fig. S6); within this alignment, there was slightly better agreement within the residues of the β-sheet (1.5 Å) than the helix (2.4 Å). If the structures were aligned using only the helical residues, the agreement in that region improved to 0.6 Å RMSD. Therefore the majority of variation between the SSNMR and crystal structures is the orientation of the helix relative to the four-stranded beta-sheet, which we attribute to a relative lack of long-range helix-sheet distance restraints, which might be remedied in the future with additional restraints involving side-chain (especially methyl) protons. Examples of correlations in Figure 2 and Figure 3 that exceed 5 Å in the final structure are D46/F52, D46/A48, Y45/D47, F52/V54, F52/T16, F52/A48, W43Hε/V39 and W43Hε/K31, with distances of 5.7, 7.4, 6.7, 7.1, 7.2, 9.6, 7.1 and 7.7 Å, respectively.

Beyond structure determination, MAS SSNMR methods enable reporting of site-specific correlations between the protein and solvent molecules.[13] In Fig. 3, many residues show strong correlations to the H_2O resonance. The observation of strong cross peaks to solvent is restricted to amino acids within ~5 Å of the solvent accessible surface; thus it is likely that polarization transfer between protein molecules is attenuated. Perhaps for this reason, in this study it was not necessary to remove intermolecular cross peaks in any systematic way or to obtain data with samples containing physical mixtures of isotopic labels. This contrasts

with ^{13}C- and ^{15}N-based distance measurements, in which we have observed many intermolecular correlations for GB1 (W. Trent Franks, unpublished data), which may complicate some stages of data analysis. We envision that combinations of distance-measurement techniques involving all available nuclei, including samples with physical mixtures of differently labeled molecules, will enable such intra-and intermolecular correlations to be determined uniquely in general.

In conclusion, we have shown high-resolution proton spectra in the solid state can be obtained for deuterated proteins using high magnetic field and fast MAS. Sensitivity was enhanced by a factor of 18 via ^{1}H *versus* ^{15}N detection. ^{1}H-^{1}H correlations from nuclei separated by up to 9 Å were observed by actively recoupling the dipolar interactions, allowing the fold of GB1 to be determined despite the scarcity of side-chain protons. This contrasts with the NOE-based methods in solution NMR, where 5 Å is the upper limit and methyl protonation is critical for proper global folding.[14] Nevertheless, the use of methyl labeling in the solid state, as demonstrated by Zilm (48th Experimental NMR Conference, Daytona Beach, FL, USA), will surely enhance the overall structure quality. Measurements of methyl-methyl and methyl-HN distances will assist the proper folding of larger proteins. Presuming the feasibility of required isotopic labeling, we anticipate extension to much larger systems.

Experimental Section

Sample Preparation

^{13}C,^{15}N,^{2}H-labeled GB1 (T2Q mutant) was produced in *Escherichia coli* BL21(DE3) from a plasmid kindly provided by A. M. Gronenborn (University of Pittsburgh). Cells grown in 99.9% $^{2}H_2O$-based M9 minimal media containing 1.5 g/l $^{15}NH_4Cl$ and 2.4 g/l of [U-^{2}H, ^{13}C] D-glucose (Cambridge Isotope Laboratories) were harvested 12.5 hours after induction. The protein was purified and then precipitated according to the established procedure,[9] using natural abundance 2-methyl-pentane-2,4-diol and isopropanol (Sigma-Aldrich) as precipitants. About 5 mg (0.9 μmol) GB1 was packed into the 1.6-mm MAS rotor, with rubber discs utilized to maintain hydration.

NMR Experiments

The SSNMR experiments were performed at −10 °C (cooling gas) on a 750 MHz Varian INOVA spectrometer with a BioFastMAS™ ^{1}H-^{13}C-^{15}N probe (Varian, Inc.) having a scroll resonator for minimal radio-frequency heating and optimal ^{1}H sensitivity.[7] ^{1}H, ^{13}C, and ^{15}N $\pi/2$ pulse widths were 1.75, 3.8, and 4.0 μs with 208, 356, and 984 W input power levels, respectively. The relative channel efficiencies for this probe were $(B_{1W})^{H}/(B_{1W})^{N}=0.50$, where B_{1W} was the B_1 field generated by unit input power.[3] MAS rate was 39 kHz (±100 Hz). Chemical shifts were referenced to DSS using adamantane as a secondary standard.[15] Pulse sequences are in the Supporting Information.

Spectra were processed using NMRPipe,[16] with details in the figure captions. Peak picking was performed in Sparky (T. D. Goddard and D. G. Kneller, University of California, San Francisco).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

1. Castellani F, van Rossum B, Diehl A, Schubert M, Rehbein K, Oschkinat H. Nature 2002;420:98–102. [PubMed: 12422222]Lange A, Becker S, Seidel K, Giller K, Pongs O, Baldus M. Angew. Chem 2005;117:2125–2129.Angew. Chem. Int. Ed 2005;44:2089–2092.
2. Bodenhausen G, Ruben DJ. Chem. Phys. Lett 1980;69:185–189.Jeener J, Meier BH, Bachman P, Ernst RR. J. Chem. Phys 1979;71:4546–4553.
3. Reif B, Griffin RG. J. Magn. Reson 2003;160:78–83. [PubMed: 12565053]Chevelkov V, van Rossum BJ, Castellani F, Rehbein K, Diehl A, Hohwy M, Steuernagel S, Engelke F, Oschkinat H, Reif B. J. Am. Chem. Soc 2003;125:7788–7789. [PubMed: 12822982]Paulson EK, Morcombe CR, Gaponenko V, Dancheck B, Byrd RA, Zilm KW. J. Am. Chem. Soc 2003;125:15831–15836. [PubMed: 14677974]
4. Reif B, Jaroniec CP, Rienstra CM, Hohwy M, Griffin RG. J. Magn. Reson 2001;151:320–327. [PubMed: 11531354]Reif B, van Rossum BJ, Castellani F, Rehbein K, Diehl A, Oschkinat H. J. Am. Chem. Soc 2003;125:1488–1489. [PubMed: 12568603]Paulson EK, Morcombe CR, Gaponenko V, Dancheck B, Byrd RA, Zilm KW. J. Am. Chem. Soc 2003;125:14222–14223. [PubMed: 14624539]
5. Chevelkov V, Rehbein K, Diehl A, Reif B. Angew. Chem 2006;118:3963–3966.Angew. Chem. Int. Ed 2006;45:3878–3881.
6. Cornilescu G, Delaglio F, Bax A. J. Biomol. NMR 1999;13:289–302. [PubMed: 10212987]
7. Stringer JA, Bronnimann CE, Mullen CG, Zhou DH, Stellfox SA, Li Y, Williams EH, Rienstra CM. J. Magn. Reson 2005;173:40–48. [PubMed: 15705511]Zhou DH, Shah G, Cormos M, Sandoz D, Rienstra CM. J. Am. Chem. Soc. 2007in press.
8. Filip C, Hafner S, Schnell I, Demco DE, Spiess HW. J. Chem. Phys 1999;110:423–440.
9. Franks WT, Zhou DH, Wylie BJ, Money BG, Graesser DT, Frericks HL, Sahota G, Rienstra CM. J. Am. Chem. Soc 2005;127:12291–12305. [PubMed: 16131207]
10. Bennett AE, Ok JH, Griffin RG, Vega S. J. Chem. Phys 1992;96:8624–8627.
11. Schwieters CD, Kuszewski JJ, Tjandra N, Clore GM. J. Magn. Reson 2003;160:65–73. [PubMed: 12565051]
12. Fossi M, Castellani F, Nilges M, Oschkinat H, van Rossum BJ. Angew. Chem 2005;117:6307–6310.Angew. Chem. Int. Ed 2005;44:6151–6154.
13. Lesage A, Böckmann A. J. Am. Chem. Soc 2003;125:13336–13367. [PubMed: 14583011]Giraud N, Sein J, Pintacuda G, Böckmann A, Lesage A, Blackledge M, Emsley L. J. Am. Chem. Soc 2006;128:12398–12399. [PubMed: 16984173]
14. Rosen MK, Gardner KH, Willis RC, Parris WE, Pawson T, Kay LE. J. Mol. Biol 1996;263:627–636. [PubMed: 8947563]
15. Morcombe CR, Zilm KW. J. Magn. Reson 2003;162:479–486. [PubMed: 12810033]
16. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. J. Biomol. NMR 1995;6:277–293. [PubMed: 8520220]

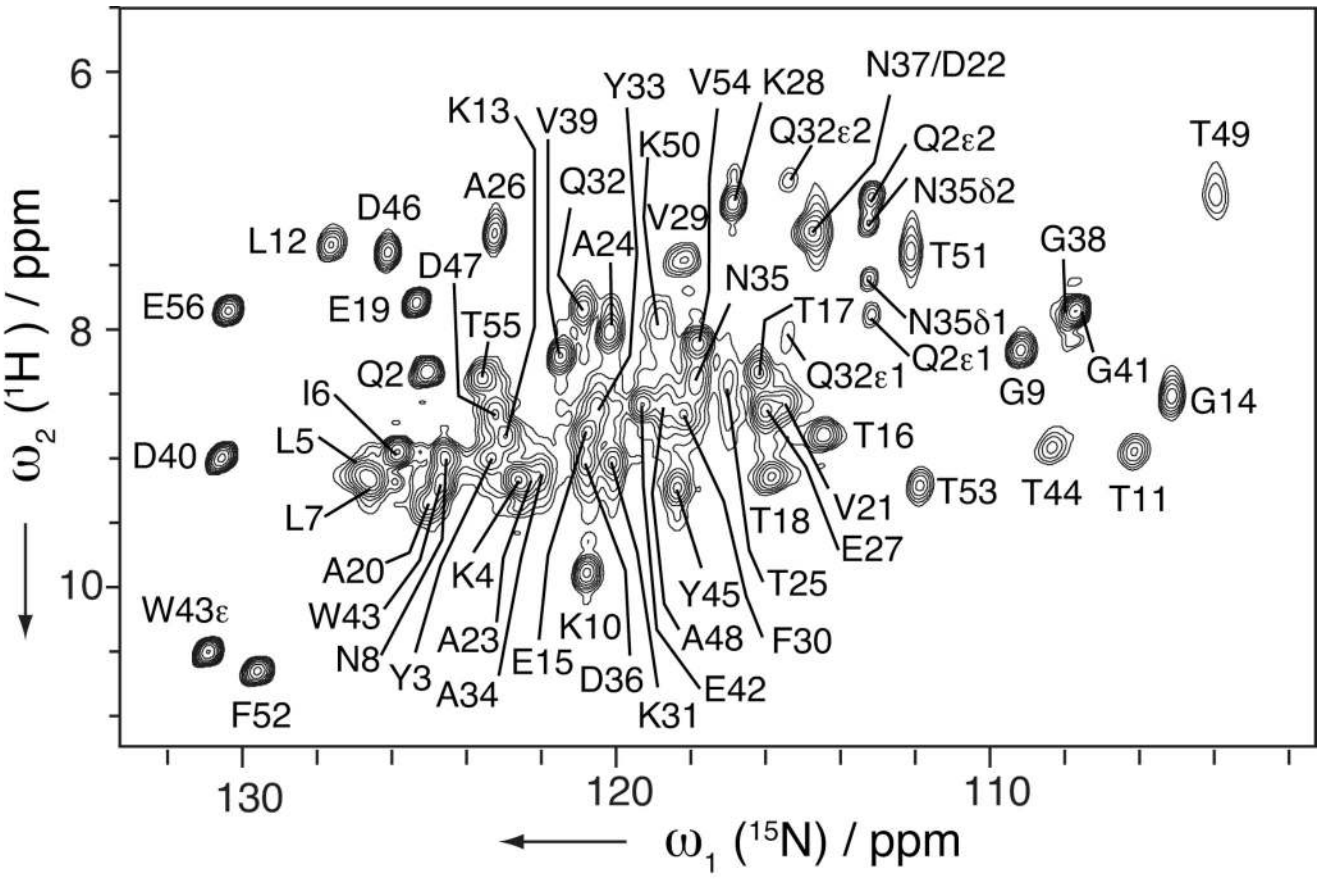

Figure 1.

Solid-state ^{1}H-detected ^{15}N-^{1}H 2D spectrum of uniformly-^{13}C,^{15}N,^{2}H-labeled GB1, back-exchanged with H_2O (750 MHz, 39 kHz MAS, 2 s recycle delay, 2 scans per row, $t_1^{max}(^{15}N)$ = 50 ms, $t_2^{max}(^{1}H)$ = 30 ms, total 30 min). No apodization or post-acquisition solvent suppression was applied. Assignments were derived from 3D experiments.

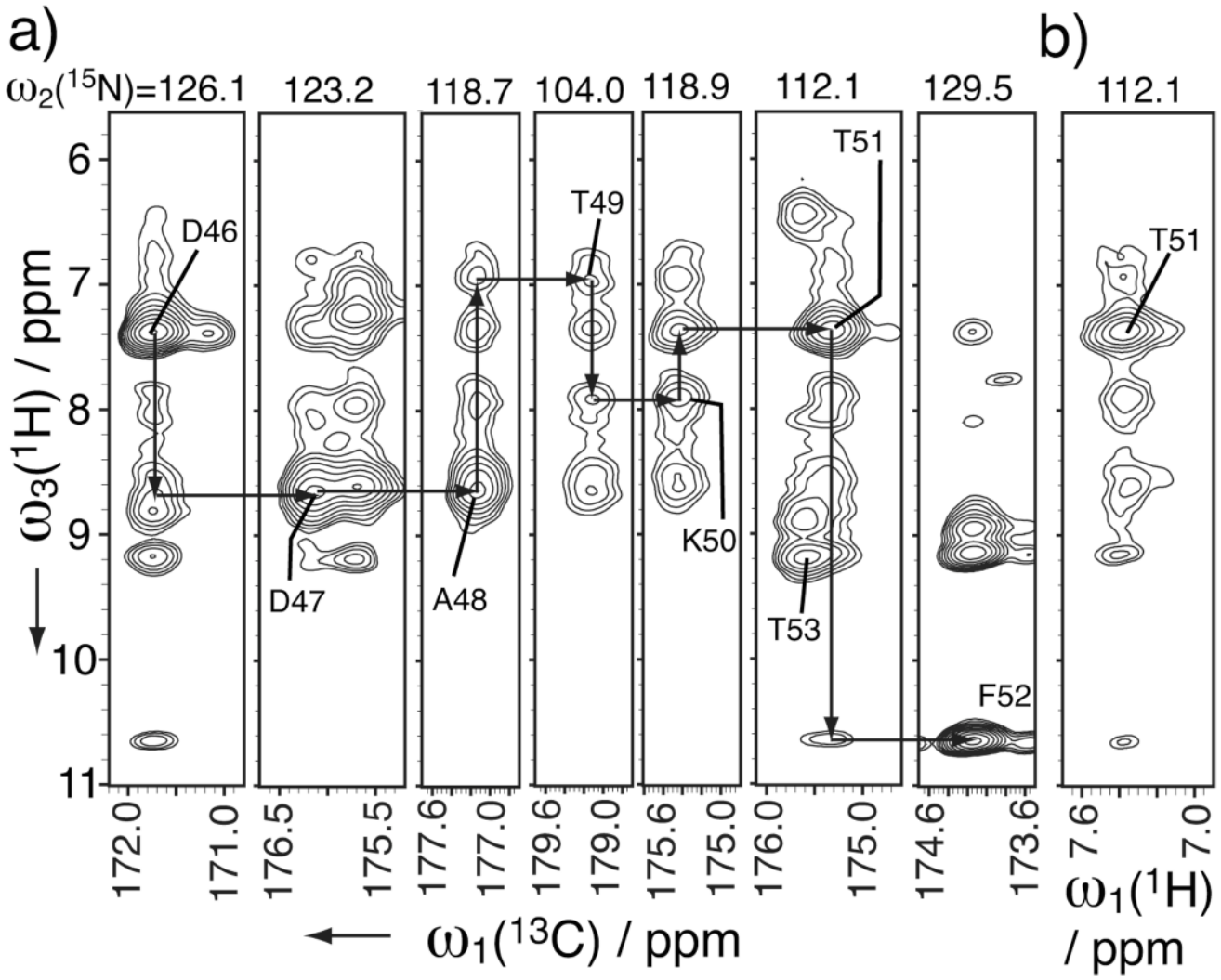

Figure 2.

2D planes of (a) CON(H)H 3D (39 kHz MAS, 750 MHz, 2 s recycle delay, 2 scans per row, $t_1^{max}(^{13}C)$ = 18.8 ms, $t_2^{max}(^{15}N)$ = 30 ms, $t_3^{max}(^{1}H)$ = 30 ms, total 36 h) and (b) HN(H)H 3D (2 scans per row, $t_1^{max}(^{1}H)$ = 12 ms, $t_2^{max}(^{15}N)$ = 30 ms, $t_3^{max}(^{1}H)$ = 30 ms, total 36 h), with ω_2 frequency indicated at the top. For both 2 ms RFDR[10] was used. Line broadening of 10, 10, 40 Hz was applied to ^{15}N, ^{13}C, and ^{1}H dimensions, respectively. The backbone walk is traced by arrows.

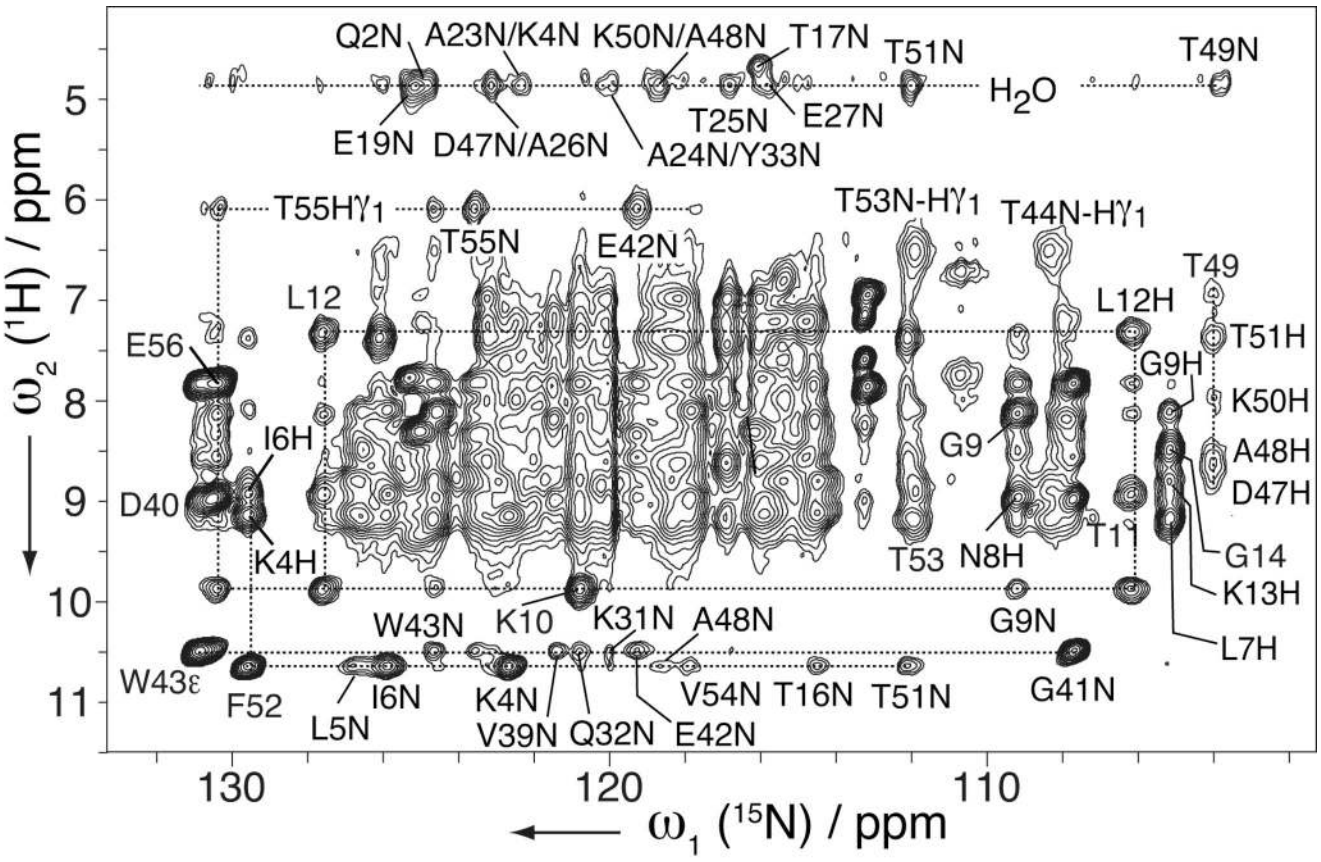

Figure 3.

N(H)H 2D spectrum of GB1 with 3 ms RFDR[10] (39 kHz MAS, 750 MHz, 2 s recycle delay, 8 scans per row, $t_1^{max}(^{15}N)$ = 50 ms, $t_2^{max}(^{1}H)$ = 30 ms, total 2 h). The ^{15}N dimension was apodized by sine bell (shift=45°), and ^{1}H by squared sine bell (shift=72°), −40 Hz Lorentzian, +80 Hz Gaussian functions. The inter-residue peaks are labeled with a suffix to indicate either ^{15}N or ^{1}H resonance frequency.

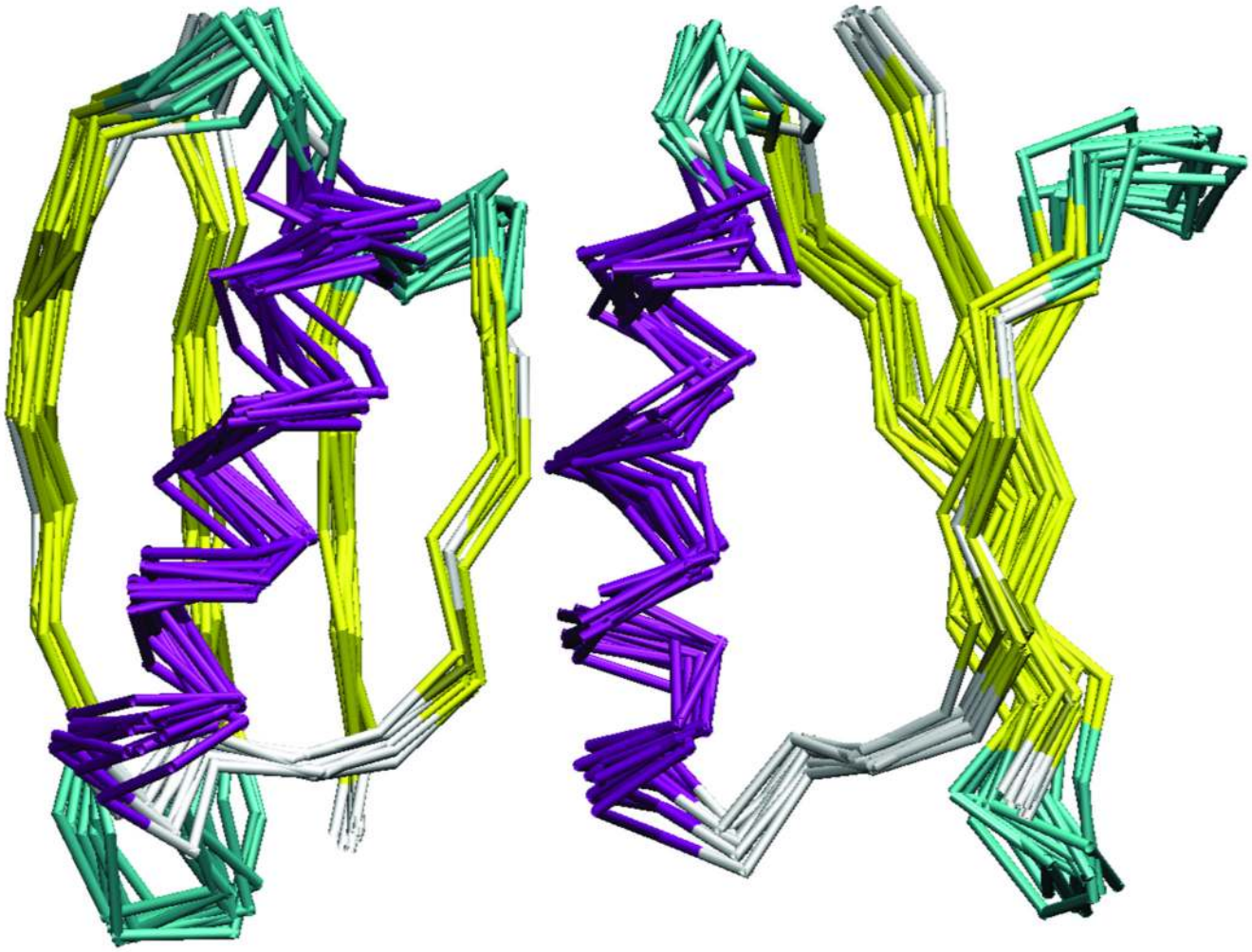

Figure 4.
Ensemble of GB1 backbone structures calculated from SSNMR proton-proton distances and dihedral constraints, shown with two views for the ten lowest energy structures from 252 calculations. The backbone RMSD is 0.82±0.14 Å.