Supporting Information

Andreas et al. 10.1073/pnas.1602248113

Structure Calculation Details

The NMR structure of GB1 was calculated with the UNIO software package (version 2.6.0), by taking as inputs the backbone chemical shifts (determined using UNIO–MATCH), the aliphatic side-chain chemical shifts [manually assigned from the (H)CCH total correlation spectroscopy (TOCSY) spectrum], the assignment of the characteristic aromatic ^{1}H and ^{13}C spins (for Trp-43), TALOS+ predictions (65) of 108 backbone dihedral angles based on ^{13}CO, ^{13}Cα, ^{13}Cβ, ^{1}H^{N}, and ^{1}Hα chemical shifts, and three unrefined 3D peak lists that probe proton–proton contacts. The three peak lists were obtained by manual signal identification in the corresponding 3D NMR spectra. In total, 703, 1,335, and 59 peaks were identified in the (H)NHH, aliphatic (H)CHH, and aromatic H(H)CH spectra, respectively. Note that only peaks giving rise to contacts to assigned resonances of Trp-43 were selected in the latter spectrum. Peak amplitudes, not volumes, were used to prevent integration errors due to signal overlap. Table S1 summarizes the identified cross-peaks and conformational restraints used in the structure calculation, and the structure quality in terms of root-mean-square deviation (rmsd) within a bundle of 20 conformers used to represent the results.

The standard unsupervised UNIO protocol was used, which consists of seven cycles of cross-peak assignment, conversion into meaningful distance restraints, and structure calculation using simulated annealing. The chemical-shift–based assignment tolerances were set to the corresponding experimental linewidths, namely 0.15 and 0.4 ppm for proton and heavy-atom dimensions, respectively. At the outset of the spectral analysis, UNIO–CANDID used highly permissive criteria to assign a comprehensive set of cross-peaks. Only the knowledge of the covalent polypeptide structure and the chemical shifts were initially exploited to guide NOE cross-peak identification and NOE assignment. In the second and subsequent UNIO–CANDID cycles, the intermediate protein 3D structures were used as an additional guide for the interpretation of the unrefined input peak lists. In each cycle, 80 initial random structures were generated in UNIO, and the 20 conformers with the lowest CYANA (71) target function values were selected as an additional filter for the subsequent cycle. Because the precision of the calculated protein structures normally improves with each subsequent cycle, the criteria for NOE assignments were successively tightened during the iterations. In each UNIO–CANDID cycle, the output consisted of an updated list of assigned NOE cross-peaks for each input peak list and a final set of meaningful upper-limit distance restraints, which constituted the input for the torsion angle dynamics algorithm of CYANA (71) for 3D structure calculation. Default calibration of peak intensities, assuming a median of deduced distances of 4 Å, was used. r^{-6} scaling of peak intensities was assumed. In addition to the distance restraints, torsion angle restraints for the backbone dihedral angles ϕ and ψ were automatically generated by UNIO from all backbone chemical shifts and added to the input for each cycle of structure calculation. Also, tighter torsion angle restraints from TALOS+ (65) were used during simulated annealing. During the first six UNIO–CANDID cycles, ambiguous distance restraints (72) were used. For the final structure calculation in cycle 7, only those distance restraints were retained by UNIO that are unambiguously valid based on the protein 3D structure from cycle 6.

Table S2 below summarizes the convergence of the GB1 structure calculation. In particular, we report the average backbone rmsd to the mean structure of the bundle, the rmsd to the X-ray reference structure (PDB ID code 2QMT), and the number of long-range restraints per residue. The reliability of the calculation is confirmed by the fact that less than 15% of cross-peaks remained unassigned in cycle 7, and backbone rmsd to mean structure in the first cycle is about 3 Å. UNIO structure calculations that meet these criteria are empirically considered as reliable and to yield the correct protein fold (35).

The NMR structure of AP205CP was calculated with the identical UNIO protocol as used for GB1. In total, 723 and 1,363 peaks were manually selected in the (H)NHH and aliphatic (H)CHH spectra, respectively. The UNIO software package (version 2.6.0) was modified to handle structural analysis of symmetric dimers, a feature that was previously not available. The evolution of UNIO cross-peak assignment and structure calculation is summarized in Table S3, again confirming the reliability of the resulting NMR structure as stated above. To improve convergence, the UNIO–CANDID protocol was supplemented with four ^{1}H-^{1}H distance restraints and 27 hydrogen bonds restraints. The four distance restraints were manually identified based on unique chemical shifts of involved side-chain ^{1}H nuclei. In detail, two spectrally unambiguous intermolecular helix–helix restraints (Val104 Hγ1–Trp97 Hε1 and Asp105 Hα–Trp97 Hε1) were identified as intermolecular, because intramolecular restraints from one end of the helix to the other would distort the helix. In addition, we entered one intermolecular helix–strand restraint (Val104 Hγ1–Ile83 Hδ1), and one intramolecular helix–loop restraint (Asn103 Hα–Ile124 Hδ1) in the calculation. Based on the observed chemical shifts that indicated beta sheet secondary structure, and on the observation of cross-strand Hα-Hα, H^{N}-H^{N}, and Hα-H^{N} RFDR contacts, 6 intermolecular and 21 intramolecular hydrogen bonds were imposed in the calculation between β-strands as detailed in Fig. S1. Because the β-strand arrangement could not be spectroscopically determined, test calculations were performed where intermolecular hydrogen bonds were entered as intramolecular. These calculations failed to converge and resulted in clearly distorted secondary structures, such as shown in Fig. S1*B*.

In the future, to assemble a model of the complete capsid by NMR, interdimer contacts would be required. However, interdimer contacts are particularly challenging to identify in AP205CP, due to a low number of expected contacts for the assigned regions of the protein, compounded by signal degeneracy. Interdimer contacts could not be definitively identified from the 3D RFDR spectra, but might be resolved using more complex sample preparation, such as amino acid-type specific labeling, or with acquisition of 4D spectra. Determination of the number of dimers and their global symmetrical arrangement would be extremely challenging, or possibly impossible by NMR. Alternatively, the full capsid structure could be modeled using both the NMR data and a low- or moderate-resolution cryo-EM map.

Shimming and Magnet Stability

We were able to shim to about 20–40 Hz in ^{1}H using a sample of silicone grease or adamantane. Because the contribution of field inhomogeneity adds geometrically to the total, the imperfect shimming therefore contributes less than 13 Hz to the inhomogeneous linewidth in GB1, and less than 10% of the line overall. Data were acquired without the use of a lock. Drift in the magnetic field was minimal, generally about 10–20 ^{1}H Hz per day, and because it was fairly linear it could be largely corrected

in processing. Acquisition times were kept to a minimum to avoid problems with drift, and split into blocks when longer acquisitions were necessary.

Detection Sensitivity

The high-resolution spectra demonstrated in Figs. 2–4 were acquired with no compromise in sensitivity compared with the previous state of the art at 60-kHz MAS in 1.3-mm rotors. An increase in the spinning rate requires reduced sample dimensions. This provides increased detection sensitivity associated with improved inductive coupling of smaller coils (almost a factor of 2 with respect to 1.3-mm rotors) but invariably entails a reduction in the sample volume (here by a factor of 4–5 with respect to 1.3-mm rotors) (18). Combined, we therefore expect a theoretical 2.2- to 2.7-fold loss in the sensitivity of a single-pulse spectrum.

We therefore compared the sensitivity of the 0.7-mm probe with our 1.3-mm probe at a field of 1 GHz. To remove contributions from solvent in the detected signal, we compared ^{15}N-^{1}H CP-HSQC or ^{13}C-^{1}H CP-HSQC spectra. Multiple factors affect this measurement, including (*i*) how well the sample is packed, (*ii*) the fact that we use rubber spacers to seal the 1.3-mm rotor (reducing the sample volume to about 2 μL), and (*iii*) the efficiency of the RF circuit. Each rotor was filled, and each CP condition was optimized to compare the best achievable signal in each case. Surprisingly, we measured approximately equal sensitivity for ^{15}N-^{1}H and ^{13}C-^{1}H CP-HSQC spectra in a 0.7- and a 1.3-mm probe, as shown by the comparison of the total area in the 1D proton spectra of Figs. S2 and S3. (This also indicates that, despite the expectation of a ∼2.7 reduction in sensitivity, our 0.7-mm probe is approximately as sensitive as our 1.3-mm probe for highly deuterated proteins or disordered samples, where resolution is dominated by inhomogeneous effects, and is not significantly improved at higher spinning rates.) This observation points to an improvement in other factors, such as RF homogeneity, matching of RF field profiles along the rotor axis, and probe electronics.

Coherence Lifetimes and Linewidths

For ^{15}N and ^{1}H, T_2' was measured using a ^{15}N-^{1}H CP-HSQC pulse sequence with a variable-time echo period on the respective channel. For ^{13}C, we used a carbon-detected CP sequence with a selective ^{13}C refocusing pulse.

We observed improvements in T_2' for ^{13}C′, ^{13}Cα, and ^{15}N, when increasing the MAS rates from the previous state of the art at 60 to ∼111 kHz. Specifically for GB1, T_2' for ^{13}C′ increased from 14 to 46 ms, for ^{13}Cα from 42 to 77 ms, and for ^{15}N from 60 to 120 ms. This results in improved *J*-transfer efficiency and therefore higher sensitivity or reduced experiment time for sequences whose efficiency depends heavily on the T_2'. For example, backbone assignment spectra can involve one or more scalar-based carbon–carbon transfers in which C′ or Cα decay according to T_2'. The efficiency of a scalar-based Cα → Cβ half-transfer (in-phase to antiphase) is expected to improve from 57% to 83%, and similarly, a C′ → Cα half-transfer from 81% to 89% (48). Thus, for GB1, there is a sensitivity gain of about 2.1 for an out-and-back Cα → Cβ transfer, and 1.6 for a full (refocused, in-phase to in-phase) C′ → Cα transfer. For AP205CP at 100 kHz, we observed a proton T_2' of about 3.0, 2.2, and 2.0 for amide protons, α protons, and all side-chain protons, respectively. Compared with measurements at 60-kHz MAS, T_2' for ^{13}Cα increased from 8.6 to 22 ms, whereas for ^{15}N and ^{13}C′ the T_2' was relatively unchanged, changing from about 21 to 23 ms, and 25 to 27 ms, respectively.

For both GB1 and AP205CP, the homogeneous component contributes about equally to the line compared with inhomogeneous contributions (70 Hz of a 100-Hz line for GB1, and 100 Hz of a 150-Hz line for AP205CP).

Considering that the ^{1}H homogeneous linewidth is expected to decrease at higher magnetic fields (38, 45, 46), and the inhomogeneous component will increase linearly, we expect that the linewidth in hertz remains approximately constant or increases less than linearly. We thus expect further resolution improvement with increasing B_0 field. Because the sensitivity also improves (theoretically as $B_0^{3/2}$), there is a strong motivation for the use of high field spectrometers for proton-detected studies of fully protonated proteins.

For spectra with indirect proton acquisition, such as the RFDR spectra, there is also an improvement in sensitivity that depends on the indirect sampling. This improvement stems from the larger integral of the signal envelope at faster spinning, due an improved total linewidth (longer apparent coherence time, T_2^*). By simply integrating the signal for two different values of T_2^*, we find that the sensitivity improvement factor [$S_i(t)$] is as follows:

$$S_i(t) = \frac{T_{2n}^*}{T_{2o}^*}\left(\frac{1-e^{-\frac{t}{T_{2n}^*}}}{1-e^{-\frac{t}{T_{2o}^*}}}\right),$$

where T_{2n}^* is the new T_2^*, and T_{2o}^* is the old T_2^*. As can be seen from the equation, the improvement in sensitivity reaches the ratio of T_2^* values in the limit of long sampling periods. For GB1, the improvement in T_2^* was about a factor of 1.5–1.8 when comparing 60 and 111 kHz. In Fig. S2, $S_i(t)$ is plotted for these two cases, and the *x* axis is shown in units of the new T_2^* value. At $1.5T_{2n}^*$, the sensitivity improves by a factor of 1.3–1.5. Note that this analysis applies to the acquired signal, and a further difference may be observed depending on the apodization function used.

Sample Preparation

Uniformly ^{13}C,^{15}N-labeled GB1 with the T2Q mutation was purchased from Giotto Biotech. The sample was dialyzed extensively against phosphate buffer (52). The concentration was increased to about 25 mg/mL using a 3-kDa Amicon concentrator (EMD Millipore) and microcrystallized by serial addition (an equal volume to the protein solution was added three times) of a mixture of methyl-2-4-pentane-diol and isopropanol according to a previously described protocol (52).

Uniformly ^{13}C,^{15}N-labeled AP205 coat protein (73) was expressed in *Escherichia coli* using a modified pETDuet vector (Novagen). Bacteria were grown in H_2O medium enriched with ^{13}C-glucose (2 g/L) and ^{15}N-labeled ammonium chloride (1 g/L) until they reached OD_{600} = 0.7. Isopropyl β-D-1-thiogalactopyranoside was added to 1 mM final concentration and cells were grown for 4 more hours before being centrifuged and frozen. Lysis buffer (40 mM Tris·HCl, pH 8.0, 300 mM NaCl, 1 mg/mL lysozyme, 10 μg/mL DNase, 10 mM $MgCl_2$) was added (3 mL per g of cells), and cells were further lysed by sonication. The resulting solution was centrifuged, and the supernatant was loaded on a Sepharose CL-4B (GE Healthcare) column. The resulting fractions containing the capsids were further purified on a Fractogel (Merck) ion-exchange column. The eluent was concentrated to 10 mg/mL using a 10-kDa Amicon concentrator and crystallized in hanging drops by addition of an equal volume of precipitant solution [10 mM Hepes, pH 7.5, 0.1 M NaCl, and 10% PEG (wt/vol) 4000].

Microcrystals were harvested and packed by ultracentrifugation at 165,000 × *g* for 15 h at 12 °C directly into the NMR rotor using a device provided by Giotto Biotech, similar to those described in literature (74, 75).

NMR Spectroscopy

All spectra were recorded at $\omega_{0H}/2\pi$ = 1 GHz and a MAS rate $\omega_r/2\pi$ of 100 kHz (AP205CP) or 111.111 kHz (GB1) using a

Bruker 0.7-mm HCN probe. Sample temperature was maintained at about 10 °C using a Bruker cooling unit with regulated N_2 gas directed at the rotor. The temperature of this gas measured just before reaching the sample were 260 K (AP205CP) and 255 K (GB1).

The dipolar-based ^{15}N-^{1}H and ^{13}C-^{1}H CP-HSQC experiments follow, with little modifications, those introduced by Rienstra and coworkers (21, 46). For ^{1}H-^{15}N and ^{1}H-^{13}C CP, we optimized around nutation frequencies of 5/4 ω_r, and 1/4 ω_r, respectively for proton and ^{15}N (or ^{13}C), with a 10% linear ramp applied on the ^{1}H channel. For ^{13}C-^{15}N CP, a 10% tangent ramp was applied on the ^{15}N frequency at 2/5 ω_r and the ^{13}C nutation frequency was about 3/5 ω_r. Low-power WALTZ-16 decoupling of 10 kHz was applied for heteronuclear decoupling. DIPSI-2 of $\gamma B_1/2\pi$ = 20 kHz was used for ^{13}C decoupling during acquisition due to the presence of homonuclear ^{13}C-^{13}C J-couplings in the uniformly labeled sample. Suppression of solvent signals (9) was applied using the MISSISSIPPI scheme (76) without the homospoil gradient for 100–200 ms, and the interscan delay ranged from 0.8 to 1 s for GB1, and 1–1.4 s for AP205CP.

The (H)CCH pulse sequence is similar to the HCCH-TOCSY (77) as well as a related sequence (62) for MAS based on total through-bond correlation spectroscopy (78, 79). In the present implementation, composite ^{13}C pulses were applied with a low nutation frequency of one-quarter the MAS frequency, a condition made efficient by the fast spinning regime. Simulations using the software package SIMPSON (70) indicate that the transfer is primarily mediated by the *J* coupling (Fig. S5).

In 3D (H)NHH and (H)CHH spectra, ^{1}H-^{1}H RFDR recoupling (80) was applied after the back-CP at a ^{1}H RF frequency of 100 kHz, yielding ^{1}H-^{1}H contacts resolved using the shift of ^{15}N, aliphatic ^{13}C, or aromatic ^{13}C.

Spectra were apodized in each dimension with 50–90° shifted squared sine-bells ("qsine 3" or "qsine 2" in Bruker Topspin), and zero-filled to at least twice the number of points in the indirect dimensions. Where linewidths are reported, no apodization was applied for the reported frequency. Acquisition and processing parameters specific for each dataset are summarized in Tables S4 and S5.

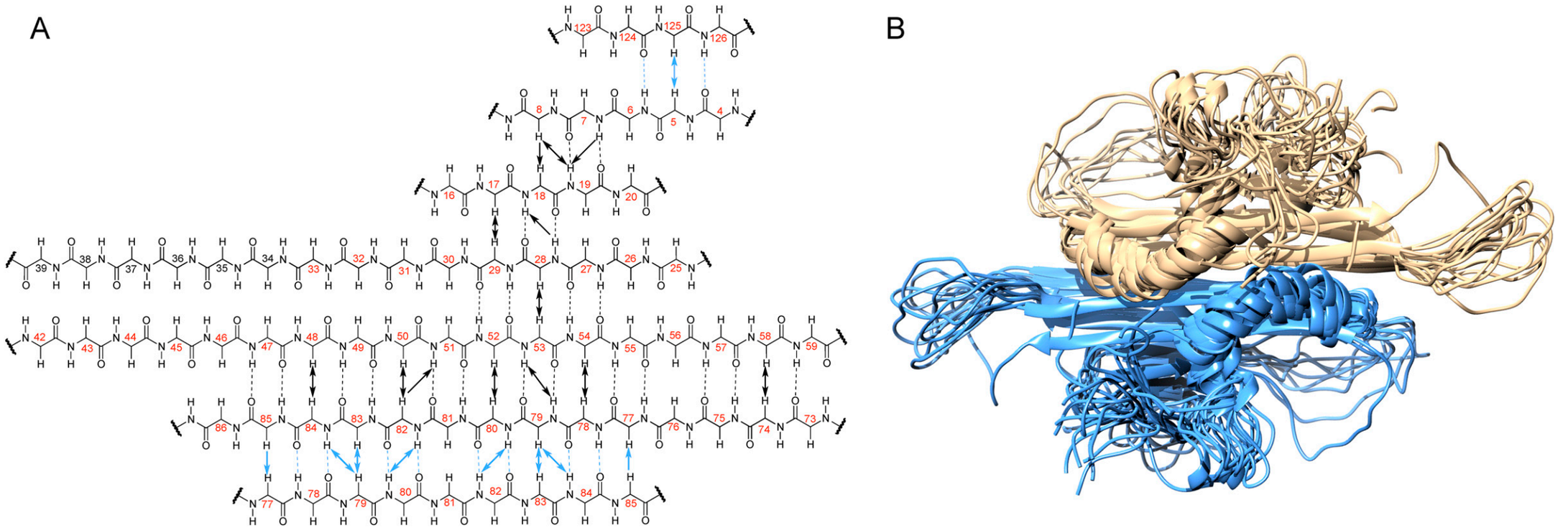

Fig. S1. In *A*, the β-strand arrangement in AP205CP. Residues identified by TALOS+ as β-sheet secondary structure are labeled in red. Black arrows indicate contacts present in the (H)CHH and (H)NHH spectra. When the cross-peak could be identified from both directions (the majority of cases), a double-headed arrow is drawn. Dashed lines indicate the hydrogen bonds entered in the structure calculation. Intermolecular hydrogen bonds are colored in blue. In *B*, the 10 lowest-energy conformers of a test calculation with an incorrect arrangement of one of the β-strands. Two of the hydrogen bonds (residues 6–124 and 4–128) are entered as intramolecular instead of intermolecular. The bundle is not well defined, and the helices are severely bent, indicating disagreement with the experimental data.

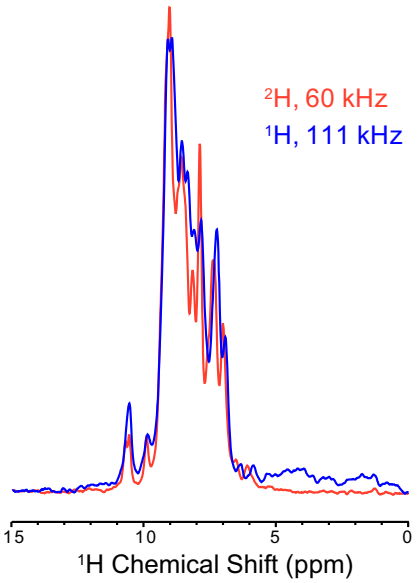

Fig. S2. Comparison of optimized ^{15}N-^{1}H CP-HSQC spectra acquired either in a 1.3-mm probe at 60 kHz on perdeuterated GB1 (red) or in the 0.7-mm probe at 111 kHz on protonated GB1 (blue). Both spectra were acquired on the 1-GHz spectrometer, and 100 Hz of exponential line broadening was applied.

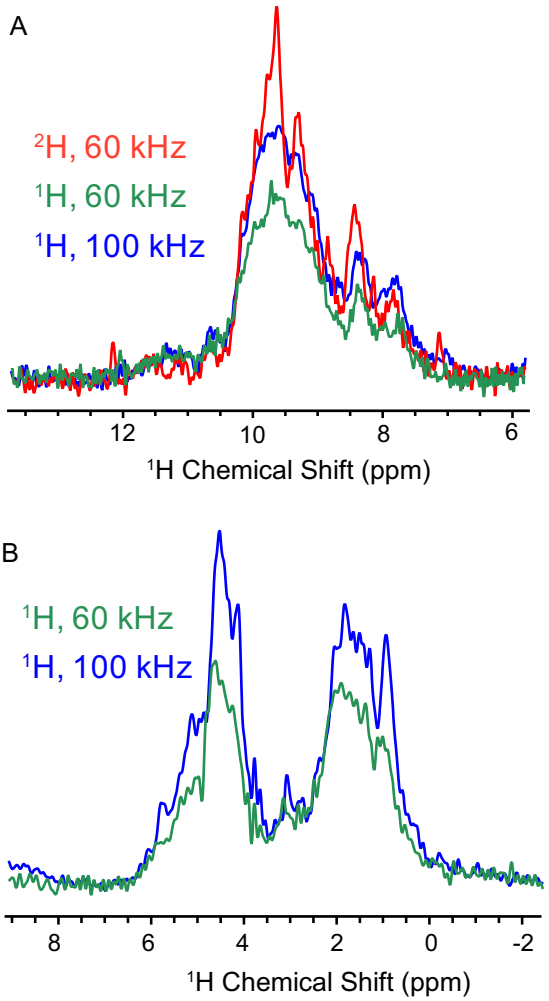

Fig. S3. Comparison of sensitivity of AP205CP in the 1.3- and 0.7-mm probes. In *A*, 1D ^{15}N-^{1}H CP-HSQC spectra acquired either in a 1.3-mm probe at 60-kHz MAS with perdeuterated AP205CP (red), a 1.3-mm probe at 60 kHz with protonated AP205CP (green), or in the 0.7-mm probe at 100 kHz and protonated AP205CP (blue). Spectra were acquired on the 1-GHz spectrometer, and no line broadening was applied. In *B*, 1D ^{13}C-^{1}H CP-HSQC spectra of protonated AP205CP acquired either in a 1.3-mm probe at 60-kHz MAS (green) or in a 0.7-mm probe (blue).

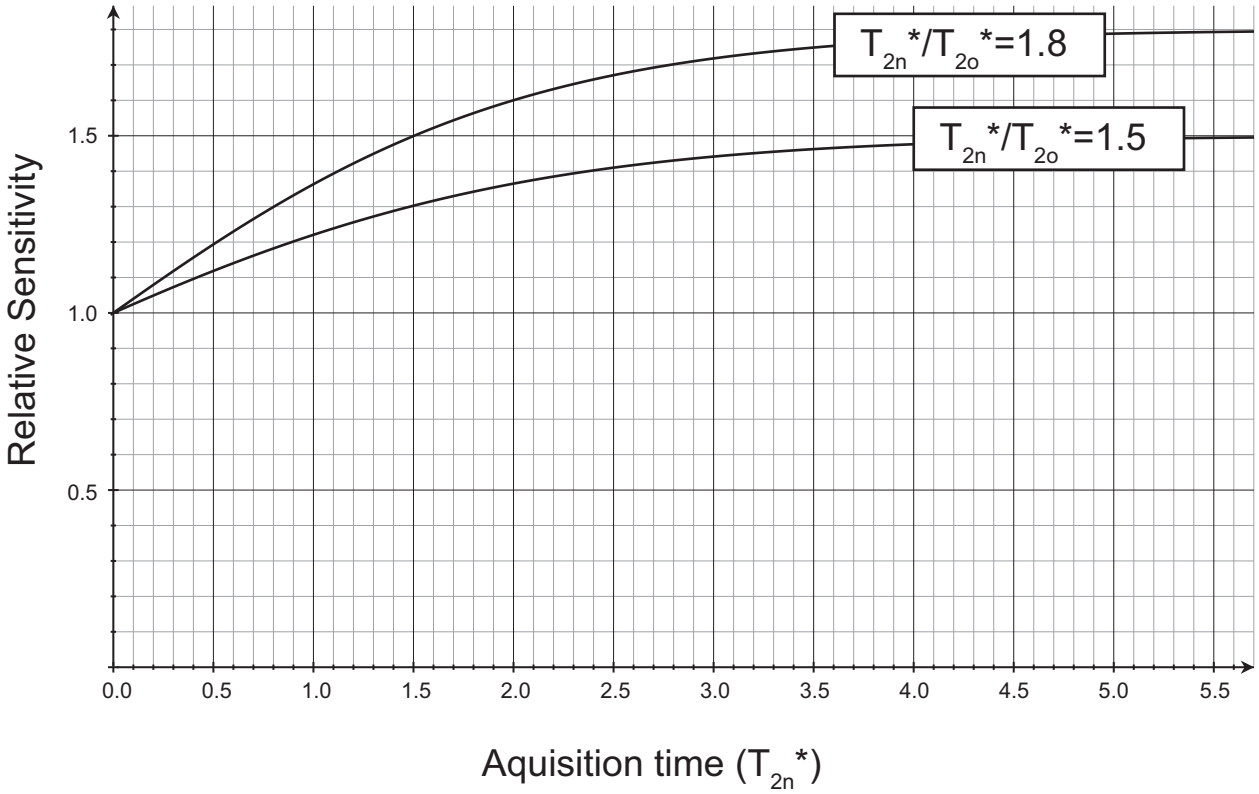

Fig. S4. Improvement in sensitivity due to improved linewidth, as a function of the acquisition time in an indirect dimension. The *x* axis is shown in units of the new T_2* (T_{2n}*).

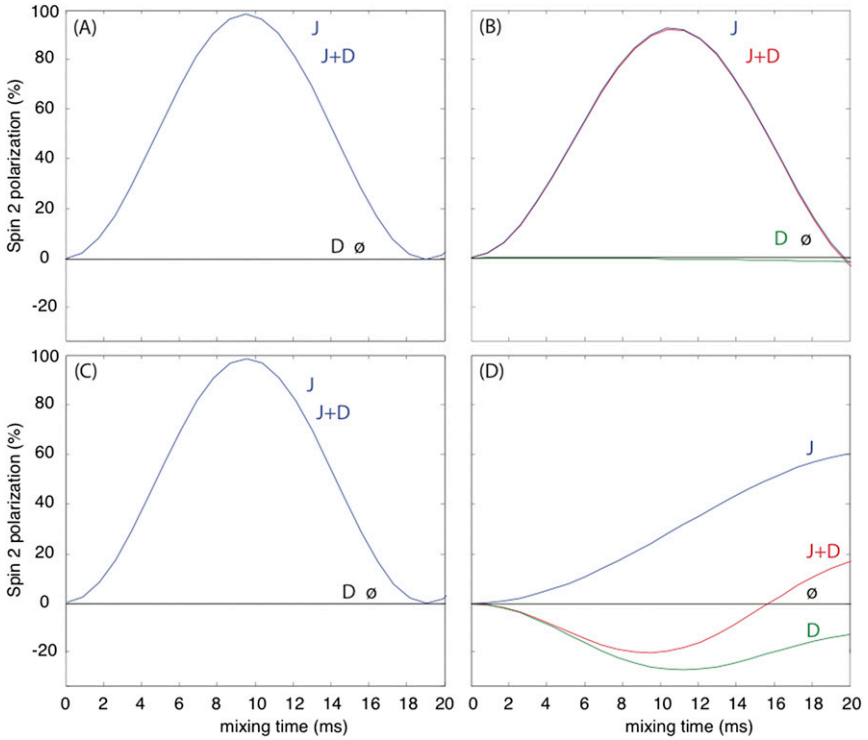

Fig. S5. Simulation of the transfer efficiency of WALTZ-16 sequence at 111-kHz MAS and an external magnetic field of 1 GHz. The applied nutation frequency is one-quarter of the rotor frequency. The spin system corresponds to two ^{13}C spins. The scalar coupling constant was set to 50 Hz, and the dipolar coupling constant was chosen to correspond to a spin–spin distance of 1.5 Å. In *A*, the isotropic chemical shifts were 10 and −10 ppm and the CSA was neglected (aliphatic–aliphatic transfer with little offset). In *B*, the isotropic chemical shifts were larger, 20 and −20 ppm, and the CSA was again neglected (aliphatic–aliphatic transfer with larger offset). In *C*, the isotropic chemical shifts were 10 and −10 ppm, and the chemical shift anisotropy and asymmetry were set to 80 ppm and 0.9, respectively, for both spins (aromatic–aromatic transfer). Finally, in *D*, the isotropic chemical shifts were 80 and −10 ppm, and the chemical shift anisotropy and asymmetry were set to 80 ppm and 0.9, respectively, for the first spin only (aromatic–aliphatic transfer). In all cases, only scalar or dipolar coupling was switched on for "J" and "D" curves, respectively; both were switched on for "J + D" curves. The "ø" curves show no transfer (both J and D off). The simulations were performed by using SIMPSON software (70).

Table S1. Summary of identified cross-peaks and conformational restraints used in the structure calculation, and structure quality in terms of rmsd within a bundle of 20 conformers used to represent the results

	GB1			AP205CP	
Assigned cross-peaks	NHH	CHH	Ar. CHH*	NHH	CHH
Intraresidue ($i = j$)	203	598	16	160	453
Sequential ($\|i - j\| = 1$)	213	170	3	202	222
Medium range ($1 \leq \|i - j\| < 5$)	67	65	0	30	48
Long range ($\|i - j\| \geq 5$)	60	123	22	42	89
Distance restraints[†]					
Total		766		1,376	
Intraresidue		128		154	
Sequential ($\|i - j\| = 1$)		258		562	
Medium range ($1 \leq \|i - j\| < 5$)		144		250	
Long range ($\|i - j\| \geq 5$)		236		410 (104[‡])	
No. of restraints per residue		13.7		5.3 (6.2[§])	
Dihedral angle restraints		108		352	
Backbone rmsd		0.48 Å[¶]		1.23 Å[#]	
Heavy-atom rmsd		1.04 Å[¶]		1.84 Å[#]	
Backbone rmsd to PDB 2QMT		1.45 Å[¶]		—	

*Sequence-specific assignment of the aromatic side-chain resonances was made for one of six aromatic residues (Trp-43).

[†]Only meaningful, nonredundant distance restraints are reported, including multiple restraints generated by the UNIO protocol to resolve the multiplicity of assignment of each ambiguous peak, as described in *Supporting Information*.

[‡]Number of intermolecular restraints.

[§]Number of restraints per assigned residue.

[¶]Calculated over the entire polypeptide chain.

[#]Calculated over the regions of the protein in regular secondary structure, namely, residue ranges 3–5, 7–12, 15–20, 23–38, 43–58, 74–86, 90–108, 113–115, and 125–129 of both chains in the dimer.

Table S2. Summary of the GB1 structure calculation statistics

Structure calculation statistics	Cycle 1	2	3	4	5	6	7
Peaks							
Selected	2,097	2,097	2,097	2,097	2,097	2,097	2,097
Assigned, %	96.0	95.9	94.0	93.5	91.7	90.0	86.6
Unassigned, %	4.0	4.1	6.0	6.5	8.3	10.0	13.4
Off-diagonal assignment, %	86.3	86.3	86.0	85.9	85.6	85.4	84.8
Cross-peaks							
With intraresidual assignment, $\|i - j\| = 0$	835	840	840	837	836	832	817
With short-range assignment, $\|i - j\| = 1$	530	532	519	497	471	446	386
With medium-range assignment, $1 < \|i - j\| < 5$	238	239	193	171	148	139	132
With long-range assignment, $\|i - j\| \geq 5$	136	123	143	179	192	195	205
Average assignment per constraint							
In hCHH spectrum	4.5	3.96	2.1	1.9	1.69	1.51	1.24*
In hNHH spectrum	5.16	3.88	2.25	2.13	1.87	1.65	1.27*
In HhCH(aro) spectrum	4.37	3.58	1.98	1.75	1.64	1.58	1.17*
Upper distance limits							
All	1,256	1,150	901	847	790	748	766
Intraresidual, $\|i - j\| = 0$	313	180	177	149	153	155	128
Sequential, $\|i - j\| = 1$	483	443	409	377	340	307	258
Medium range, $1 < \|i - j\| < 5$	318	370	180	160	131	122	144
Long range, $\|i - j\| \geq 5$	142	157	135	161	166	164	236
Long range per residue	2.5	2.8	2.4	2.9	3.0	2.9	4.2
Average target function value, Å^2	183	23	26	6.3	1.9	0.34	0.41
Rmsd (all residues 1–56), Å							
Average backbone to mean	3.02	1.35	0.73	0.62	0.68	0.62	0.48
Average backbone to X-ray 2QMT	2.15	1.4	1.39	1.37	1.16	1.25	1.45
Average heavy atom to mean							1.04

*In previous implementations of UNIO–CANDID, only cross-peaks with highly unambiguous assignments were kept in cycle 7, and therefore the average assignment per constraint (peak) was forced to be 1 at the end of the calculation. In the current UNIO 2.6.0 version, information from unambiguous peaks is kept by splitting peak volumes after the assignment step during the conversion into distance restraints in the final cycle. This results in an average assignment per peak greater than 1. Retained peak ambiguities are ultimately converted into unambiguous distance restraints during the generation of distance restraints.

Table S3. Summary of the AP205CP structure calculation statistics

Structure calculation statistics	Cycle 1	2	3	4	5	6	7
Peaks							
Selected	2,086	2,086	2,086	2,086	2,086	2,086	2,086
Assigned, %	58.4	81.5	78.7	78.5	77.5	77.1	74.0
Unassigned, %	41.6	18.5	21.3	21.5	22.5	22.9	26.0
Off-diagonal assignment, %	81.7	82.5	81.9	81.9	81.6	81.5	80.8
Cross-peaks							
With intraresidual assignment, $\|i - j\| = 0$	430	654	652	649	645	641	613
With short-range assignment, $\|i - j\| = 1$	343	470	463	457	447	438	424
With medium-range assignment, $1 < \|i - j\| < 5$	129	185	130	127	107	102	78
With long-range assignment, $\|i - j\| \geq 5$	93	94	99	107	120	130	131
Average assignment per constraint							
In hCHH spectrum	4.69	5.41	2.61	2.37	2.11	1.82	1.39*
In hNHH spectrum	4.24	4.40	2.27	2.29	1.97	1.7	1.36*
Upper distance limits							
All	1,199	1,630	1,326	1,256	1,204	1,152	1,376
Intraresidual, $\|i - j\| = 0$	154	138	224	184	182	164	154
Sequential, $\|i - j\| = 1$	500	682	678	636	612	584	562
Medium range, $1 < \|i - j\| < 5$	329	552	250	246	208	192	250
Long range, $\|i - j\| \geq 5$	216	258	174	190	202	212	410
Intermolecular	0	9	46	46	46	48	104
Long range per residue	0.8	0.8	0.7	0.7	0.8	0.8	1.6
Average target function value, Å^2	299	38.8	32.6	13.7	8.7	10.6	8.6
Rmsd (residues 3–5, 7–9, 15–20, 23–34, 43–58, 74–86, 90–108, 113–115, 125–127), Å							
Average backbone to mean	2.16	2.17	1.54	1.43	1.43	1.4	1.23
Average backbone to X-ray 5FS4	3.83	2.63	2.15	1.97	2.19	2.34	2.35
Average heavy atom to mean							1.84

*In previous implementations of UNIO–CANDID, only cross-peaks with highly unambiguous assignments were kept in cycle 7, and therefore the average assignment per constraint (peak) was forced to be 1 at the end of the calculation. In the current UNIO 2.6.0 version, information from unambiguous peaks is kept by splitting peak volumes after the assignment step during the conversion into distance restraints in the final cycle. This results in an average assignment per peak greater than 1. Retained peak ambiguities are ultimately converted into unambiguous distance restraints during the generation of distance restraints.

Table S4. Acquisition and processing parameters of used NMR spectra for GB1

Spectrum	Max indirect evolution	Scans per point	Experimental time	Signal/noise (first FID)	Processing
(H)NH	50 ms (N)	4	24 min	46	60° shifted sine-bell squared
(H)CH	46 ms	8	5 h	79	60° shifted sine-bell squared
(H)NCAH	7.9 ms (CA), 19.7 ms (N)	4	25 h	3.2	90° shifted sine-bell squared
(H)CANH	7.9 ms (CA), 19.7 ms (N)	4	25 h	8.9	90° shifted sine-bell squared
(H)(CO)CA(CO)NH	3.3 ms (CA), 9.2 ms (N)	4	3.5 h	3.5	90° shifted sine-bell squared
(H)(CA)CB(CA)NH	3.3 ms (CA), 8.0 ms (N)	4	6 h	3.6	90° shifted sine-bell squared
(H)N(CA)(CO)NH	8.3 ms (CA, N)	8	7.5 h	3.3	60° shifted sine-bell squared
(H)N(CO)(CA)NH	8.3 ms (CA, N)	8	7.5 h	3.6	60° shifted sine-bell squared
(H)CCH TOCSY (15 ms)	6.8 ms	2	32.5 h	3.0	90° shifted sine-bell squared
(H)NHH (RFDR 1.8 ms)	6.9 ms (H), 13.6 ms (N)	4	15 h	14.3	60° shifted sine-bell squared
(H)CHH (RFDR 0.5 ms)	2.6 ms (H), 5.3 ms (C)	4	14.5 h	67	90° shifted sine-bell squared
H(H)CH aromatic ^{13}C filtered (RFDR 0.5 ms)	5.7 ms (H), 10.2 ms (C)	4	38 h	4.5	90° shifted sine-bell squared

The direct proton dimension was sampled to 10 ms. Spinning rate was 111.111 kHz.

Table S5. Acquisition and processing parameters of used NMR spectra for AP205CP

Spectrum	Max indirect evolution	Scans per point	Experimental time	Signal/noise (first FID)	Processing (indirect/direct)
(H)NH	15.8 ms (N)	32	1 h 17 min	83	60° shifted sine-bell squared
(H)CH	17 ms (C)	8	2 h 46 min	85	60° shifted sine-bell squared
(H)NCAH	6.36 ms (CA), 9.86 ms (N)	2	6 h	7.9	60° shifted sine-bell squared
(H)CANH	6.4 ms (CA), 9.9 ms (N)	2	6 h	5.3	60° shifted sine-bell squared
(H)(CO)CA(CO)NH	6.4 ms (CA), 9.9 ms (N)	8	25 h	6.7	60° shifted sine-bell squared
(H)(CA)CB(CA)NH	—	—	—	—	—
(H)N(CA)(CO)NH	—	—	—	—	—
(H)N(CO)(CA)NH	—	—	—	—	—
(H)CCH TOCSY (14 ms)	7.0 ms	2	55 h	10	60/50° shifted sine-bell
H(H)NH (RFDR 1 ms)	5 ms (H), 10.3 ms (N)	4	60 h	35	72/60° shifted sine-bell squared
H(H)CH (RFDR 1 ms)	5 ms (H), 4.8 ms (C)	4	43 h	35	72/60° shifted sine-bell squared
(H)CO(N)CAH	6.4 ms (CO), 5.3 ms (CA)	24	37 h	10	90° shifted sine-bell squared

The direct proton dimension was sampled to 10 ms. Spinning rate was 100 kHz.