Combining DNP NMR with segmental and specific labeling to study a yeast prion protein strain that is not parallel in-register

Kendra K. Frederick[a,1,2], Vladimir K. Michaelis[b,c,3], Marc A. Caporini[d,4], Loren B. Andreas[b,c], Galia T. Debelouchina[b,c,5], Robert G. Griffin[b,c], and Susan Lindquist[a,e,f,6]

[a]Whitehead Institute for Biomedical Research, Cambridge, MA 02142; [b]Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA 02139; [c]Francis Bitter Magnet Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139; [d]Bruker BioSpin, Billerica, MA 01821; [e]Howard Hughes Medical Institute, Massachusetts Institute of Technology, Cambridge, MA 02138; and [f]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02138

Edited by Angela M. Gronenborn, University of Pittsburgh School of Medicine, Pittsburgh, PA, and approved March 2, 2017 (received for review November 22, 2016)

The yeast prion protein Sup35NM is a self-propagating amyloid. Despite intense study, there is no consensus on the organization of monomers within Sup35NM fibrils. Some studies point to a β-helical arrangement, whereas others suggest a parallel in-register organization. Intermolecular contacts are often determined by experiments that probe long-range heteronuclear contacts for fibrils templated from a 1:1 mixture of ^{13}C- and ^{15}N-labeled monomers. However, for Sup35NM, like many large proteins, chemical shift degeneracy limits the usefulness of this approach. Segmental and specific isotopic labeling reduce degeneracy, but experiments to measure long-range interactions are often too insensitive. To limit degeneracy and increase experimental sensitivity, we combined specific and segmental isotopic labeling schemes with dynamic nuclear polarization (DNP) NMR. Using this combination, we examined an amyloid form of Sup35NM that does not have a parallel in-register structure. The combination of a small number of specific labels with DNP NMR enables determination of architectural information about polymeric protein systems.

[*PSI*$^+$] prion | solid-state NMR | amyloid | Sup35 | dynamic nuclear polarization

Yeast prions are protein-based epigenetic mechanisms of inheritance (1). They permit biological information to be encoded and inherited solely through a self-propagating conformation of a protein. The best-characterized yeast prion, [*PSI*$^+$] (2), is an amyloid form of the translation termination factor Sup35 (3). The amyloid fold is polymeric fiber of protein monomers that is rich in β-sheets that run perpendicular to the fiber axis. Amyloid formation sequesters Sup35 from the ribosome, changing the rate of stop-codon read-through and creating a variety of new yeast phenotypes (4, 5). Different regions of the Sup35 protein are responsible for prion templating, prion inheritance, and translation termination. The N-terminal domain "N" is Q/N-rich and required for the formation and templating of amyloid. The highly charged K/E-rich middle domain "M" promotes solubility in the nonprion form and is required for chaperone-mediated prion inheritance (6, 7). The C-terminal domain is necessary and sufficient for translation termination. The N and M domains together, NM, contain all of the necessary features to act as an amyloid-based prion.

A wide variety of approaches have been used to investigate Sup35 prion structures, but a consensus structural picture has yet to emerge. Amyloids are notoriously difficult to study, mainly because they are insoluble yet noncrystalline and are therefore not easily amenable to traditional structural biology methods. In all models, at least some part of the N domain is involved in the amyloid fold (8–11), and most of the M domain is thought to be solvent-accessible and intrinsically disordered (11–13). However, there are two conflicting models of how monomers of NM are arranged in amyloid fibers. In one model, called the β-helical model, monomers make intermolecular contacts in discrete regions, with the region in-between making intramolecular contacts. This model is supported by the patterns of interaction and cross-linking determined from a series of site-directed mutants (8). The work has the caveat that mutations may perturb the fibril structure. However, fibers made from the mutated versions of the protein can transform yeast into the corresponding prion state (14, 15), suggesting that the fiber structure is unperturbed. In the other model, called parallel in-register, each residue in one monomer makes intermolecular contact with the identical residue in the neighboring monomers in the fiber. This model is supported by a magic-angle spinning (MAS) NMR experiment that measures dipole–dipole relaxation rates for specifically installed isotopic labels (PITHIRDS-CT) (16). Isotopically labeled sites near other isotopically labeled sites will have faster signal decay rates than isolated sites. If the arrangement of isotopic labels in the fiber is modeled as a linear array, PITHIRDS-CT

Significance

Self-propagating changes in the conformation of amyloidogenic proteins play vital roles in normal biology and disease. Despite intense research, the architecture of amyloid fibers remains poorly understood. In this work, we used both segmental and specific isotopic labeling schemes in combination with dynamic nuclear polarization (DNP) NMR to measure long-range interactions to distinguish between two models for the arrangement of monomers in amyloid fibers. These measurements enabled us to determine that the monomers in one variant of an amyloid form of NM do not adopt a parallel in-register arrangement. The combination of segmental and specific labeling schemes with DNP NMR enables the testing of structural models for systems for which it was previously impossible due to low experimental sensitivity.

Author contributions: K.K.F., R.G.G., and S.L. designed research; K.K.F., V.K.M., M.C., L.A., and G.T.D. performed research; K.K.F. analyzed data; and K.K.F., R.G.G., and S.L. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

[1]To whom correspondence should be addressed. Email: kendra.frederick@utsouthwestern.edu.

[2]Present address: Department of Biophysics, University of Texas Southwestern, Dallas, TX 75390.

[3]Present address: Department of Chemistry, University of Alberta, Edmonton, AB, Canada T6G 2G2.

[4]Present address: Amgen, Inc., Cambridge, MA 02142.

[5]Present address: Department of Chemistry, Princeton University, Princeton, NJ 08544.

[6]Deceased October 27, 2016.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1619051114/-/DCSupplemental.

relaxation rates for many amino acids in NM are consistent with the 4.8-Å spacing for between beta-strands (12, 17, 18). However, the PITHIRDS-CT does not report on the chemical identity of the other isotopically labeled site, and the extracted distances are very sensitive to how the geometry of neighboring spins is modeled (16). Moreover, the highly degenerate nature of NM complicates interpretation of inter- versus intramolecular interactions for specifically labeled molecules, particularly for aromatic residues that may participate in ring stacking (19). Thus, differentiating between these two models for amyloid quaternary structure requires additional experimental approaches.

MAS NMR is uniquely suited to the study of protein polymers, and provides quantitative structural insights at the atomic level (20). However, experiments on large protein systems, such as Sup35, are often complicated by chemical shift degeneracy. Therefore, a common approach to studying macromolecular assemblies is to use smaller model systems. For NM, one such system is the 7-amino acid-long segment GNNQQNY, which corresponds to residues 7 to 13 of the 253-residue-long protein. This 7-residue peptide segment can assemble into several amyloid-like fibril forms as well as two forms of microcrystals, and has been studied by both MAS NMR (21, 22) and X-ray microcrystallography (23). Although GNNQQNY assemblies have a cross–β-spine that is characteristic of amyloid fibers, all of the studied forms differ from each other, and evidence that any of the resulting structures represents the conformation in the biologically active fibers remains indirect.

Extrapolation of structural studies on model systems to full-length proteins is difficult due to signal overlap. Segmental isotopic labeling has long been proposed as a solution to this problem (24) and has been successfully applied to proteins in solution NMR studies (24–28) and, very recently, been applied in MAS NMR studies (29, 30). However, even with good resolution, measuring long-range interactions remains challenging because of low experimental transfer efficiency, which results in poor sensitivity. Dynamic nuclear polarization (DNP) is a sensitivity-enhancement technique for MAS NMR. DNP MAS NMR has been used to measure intermolecular distances between specifically installed isotopic labels in an 11-amino acid fragment of transthyretin (31) and for a 1:1 mixture of ^{13}C- and ^{15}N-labeled monomers in the 86-amino acid-long protein PI3-SH3 (32). Here we combined segmental and specific isotopic labeling with DNP to obtain unambiguous long-range distance restraints for the amyloid fiber form of NM. The segmental labeling provides the spectral resolution necessary to study large systems, and the DNP provides needed sensitivity enhancement to detect long-range interactions. We report the characterization of the quaternary structure of NM fibers using approaches previously limited by experimental sensitivity.

Results

Segmental Isotopic Labeling of NM Using Split Inteins. We produced segmentally isotopically labeled proteins using split inteins to biosynthetically label only a small segment of the full-length protein (Fig. 1). Contiguous inteins are amino acid sequences that excise themselves from a polypeptide chain, ligating the flanking sequences via a native peptide bond. Split inteins can be cut into (or exist naturally as) two pieces that associate noncovalently upon reconstitution and ligate their flanking sequences (24). We used a natural split-intein system that splices rapidly and with high efficiency, the DnaE intein from *Nostoc punctiform* (Npu) (33). We created full-length NM molecules that are uniformly labeled with NMR-active stable isotopes at the first 14 amino acids (Fig. 1). To do so, we prepared chimeric His^6-NM[$^{(1–14)}$]-Npu[$^{(1–103)}$] proteins expressed in minimal media with [^{13}C]glucose and [^{15}N]ammonium chloride as the sole carbon and nitrogen sources. These were ligated to the matching chimeric construct, Npu[$^{(104–136)}$]-NM[$^{(15–25)}$]-His^6, expressed in natural-abundance media (Fig. 1*C*). Because the molecular recognition of the two parts of the intein is driven by electrostatics and the M domain of NM is highly charged, the kinetics of formation are much slower than those reported previously (33). The $t_{1/2}$ for ligation was about 5 h rather than a minute, and there were a significant number of nonproductive interactions that resulted in truncated versions of NM in the reaction mix (Fig. 1*C*). These truncated versions of NM could be incorporated into seeded fibers, so a secondary purification step based upon Ni-nitrilotriacetic (NTA) resin avidity was necessary to separate full-length product from the truncated version of NM. After the ligation reaction, we separated the product from the reactants, taking advantage of the fact that the correct product has two His^6 tags and therefore increased affinity for Ni-NTA (Fig. 1*D*). After eluting the product with an imidazole gradient, the N-terminal His^6 tags were cleaved off (Fig. 1*E*). The final product was greater than 90% pure by SDS/PAGE and contained a single cysteine mutation (or "scar") at the ligation site. We prepared more than 8 mg of segmentally labeled NM molecules per liter of isotopically labeled growth for a practical yield of about 50% of the His^6-NM[$^{(1–14)}$]-DnaE[$^{(1–103)}$] precursor. To avoid sample heterogeneity derived from multiple fiber forms, the segmentally labeled NM was templated into amyloid fibers using amyloid seeds derived from lysates of yeast carrying the strong [*PSI*$^+$] phenotype (11). Thus, we produced segmentally isotopically labeled NM molecules in sufficient quantities and purity for structural investigations by MAS NMR spectroscopy.

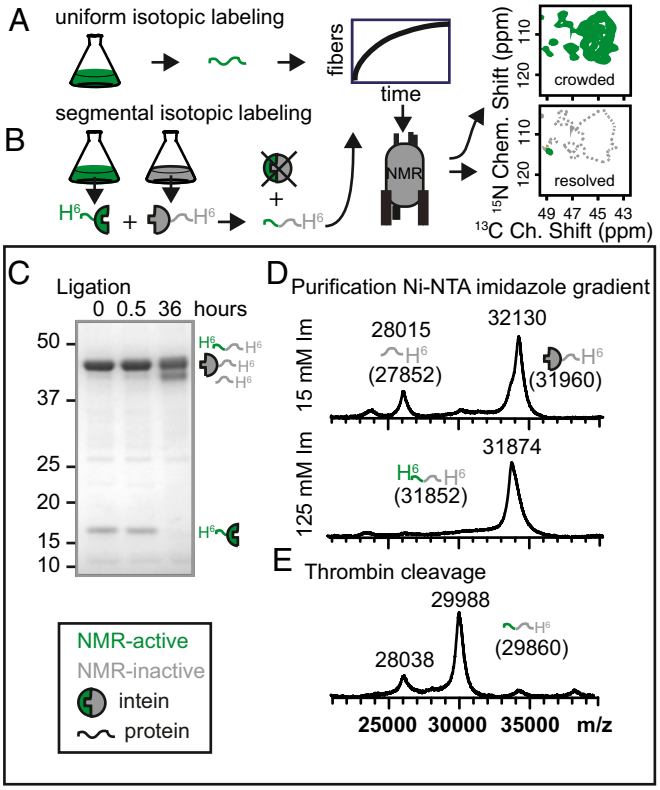

Fig. 1. Segmental isotopic labeling reduces the complexity of solid-state NMR spectra. (*A*) The glycine region of a ^{13}C-^{15}N correlation spectrum for uniformly ^{13}C- and ^{15}N-labeled NM has many peaks, consistent with a protein that contains 23 glycine residues. (*B*) The same region for a full-length NM molecule that is segmentally labeled at the first 14 amino acids has one peak, consistent with a single glycine. NM was segmentally isotopically labeled by intein-mediated ligation of NM–intein chimeras purified from bacteria grown in media containing either NMR-active (^{13}C and ^{15}N; green) or -inactive (^{12}C, ^{14}N; gray) carbon and nitrogen sources. (*C*) Ligation was initiated by mixing purified chimeric proteins and monitored by SDS/PAGE. (*D*) Purification of the ligation mixture was monitored by MALDI-TOF. Expected *m/z* is shown in parentheses. IM, imidazole. (*E*) The N-terminal His tag was removed by thrombin cleavage.

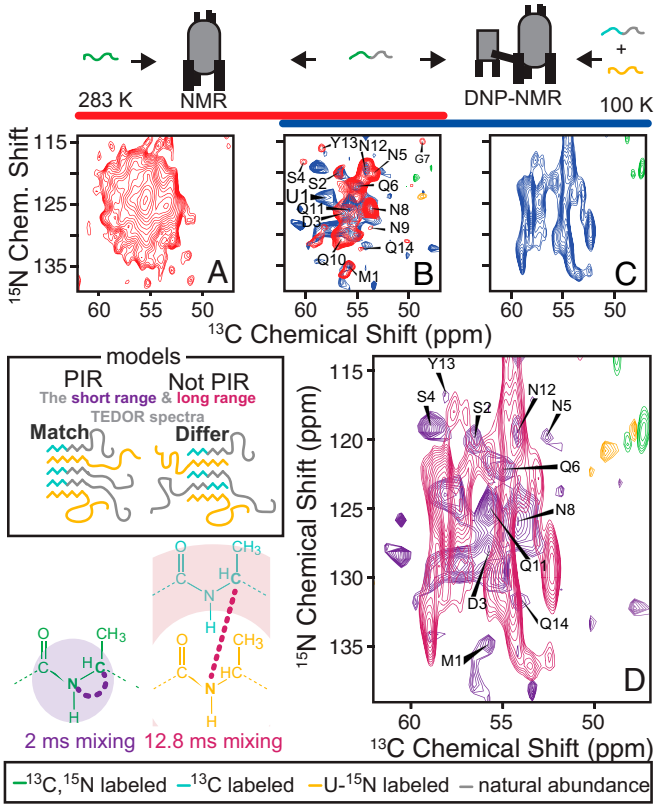

Fig. 2. Monomers of NM are not arranged in parallel in-register (PIR) β-sheets in the N-terminal region when NM is in the amyloid form. The one-bond TEDOR spectrum of a uniformly ^{13}C- and ^{15}N-labeled NM fiber sample at room temperature (*A*) is greatly simplified by using an NM molecule that is isotopically labeled at only the first 14 amino acids (*B*) (red). Despite the large change in sample temperature and the addition of a cryoprotectant (60% d_8-glycerol) and the polarizing agent AMUPol (10 mM), the one-bond TEDOR spectrum of segmentally isotopically labeled NM molecules collected using DNP MAS NMR (blue) is very similar to the room-temperature spectrum (red). However, the long-range TEDOR spectrum of an NM fiber sample made from a 1:1 mixture of segmentally ^{13}C-labeled and uniformly ^{15}N-labeled molecules collected at 100 K with cryoprotection and 10 mM AMUPol is different (*C*), as highlighted by the overlay of the 100 K DNP spectra from *B* and *C* (*D*). Assignments were mapped from the 283 K (red) spectrum onto the 100 K spectrum.

MAS NMR Spectroscopy of Segmentally Labeled NM Fibers. Segmental isotopic labeling allowed examination of a small number of consecutive labeled residues in the context of the full-length unlabeled protein, greatly simplifying the spectroscopy (Figs. 1*B* and 2 *A* and *B*). Spectra of uniformly and segmentally labeled NM fibers were in good agreement (Fig. 1 *A* and *B*). Spin systems in the segmentally labeled sample aligned with the uniformly labeled spectra as expected, but spectral crowding in the segmentally labeled sample was dramatically decreased. For example, the labeled segment contained only one glycine (Fig. 1*B*) and just two serines. Residue-specific assignments were obtained using three 2D spectra collected at room temperature using MAS NMR: a homonuclear ^{13}C-^{13}C dipolar assisted rotational resonance (DARR) (34) spectrum (900 MHz, 18.8 kHz, 60 ms mixing) and two heteronuclear experiments, a *Z*-filtered transferred-echo double-resonance (TEDOR) (35) spectrum (900 MHz, 12.5 kHz, 1.8 ms mixing) and an experiment to correlate side-chain chemical shifts to their main-chain resonances, an NCOCx (36) (800 MHz, 18.8 kHz MAS). Unassigned peaks were not symmetric across the diagonal in homonuclear experiments and/or were only present in one of three spectra, and thus were attributed to limited signal to noise rather than the presence of multiple fiber polymorphs. Every residue in the segment 1 to 14 could be unambiguously assigned (Fig. 3). Whereas chemical shift analysis suggested that some sites, such as Ser2 and Gly7, have unusual conformations as reported previously (13) (Table S1), TALOS+ predicted backbone ϕ–ψ angles that are all consistent with random-coil or β-sheet secondary structure for residues 1 to 14 of NM fibers (37).

The Structures of Microcrystalline GNNQQNY and Full-Length NM Differ. Chemical shift is sensitive to the local conformation, with main-chain and Cβ chemical shifts being strongly correlated to the dihedral ϕ and ψ angles (38, 39). To determine whether the structure of the peptide GNNQQNY—which corresponds to residues 7 to 13 of full-length NM—is preserved in the context of amyloid fibers of full-length NM, we compared chemical shifts for the microcrystalline and fibrillar forms of GNNQQNY (21, 22) with those of segmentally labeled NM fibers. Chemical shifts for the two microcrystalline forms of GNNQQNY differ by an average of 1.40, 0.66, and 0.31 ppm at the backbone nitrogen, carbonyl, and carbon α-sites, respectively (21). Our NM fibrils differed from monoclinic microcrystals by an average of 4.48, 3.53, and 1.32 ppm and from orthorhombic microcrystals by an average of 3.44, 3.65, and 1.29 ppm at the same positions (Fig. S1). Interestingly, the chemical shifts of full-length fibers also differed from all three fibril forms of GNNQQNY (21) by an average of 6.8, 2.3, and 2.0 ppm at these positions (Fig. S1). The chemical shift differences were not simply a result of the chemical differences between the model system and the full-length protein in its amyloid form. Characterization of a variety of mutated forms of the B1 domain of protein G revealed that the average changes in chemical shift from a nonstructurally perturbative mutation are small (~0.1 ppm in ^{13}C and 0.3 ppm in ^{15}N) (40).

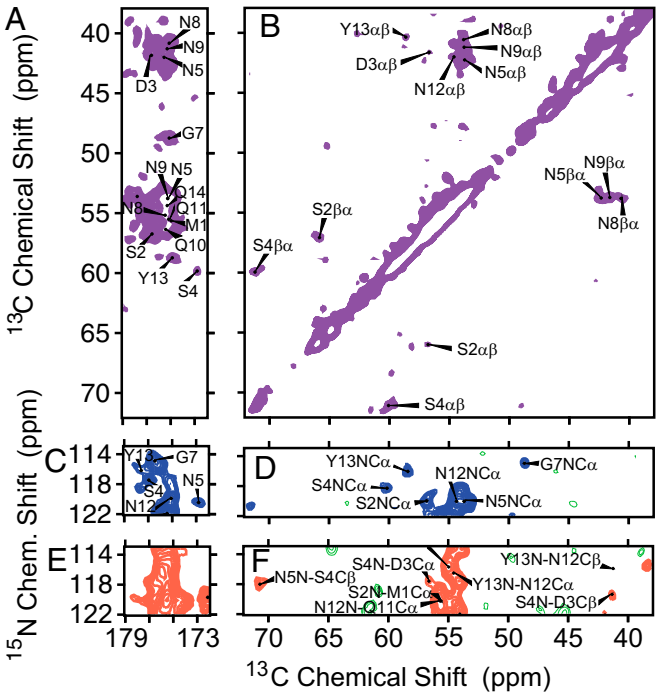

Fig. 3. Solid-state NMR spectra with assignments for segmentally labeled NM (residues 1 to 14) amyloid fibrils. Carbonyl (*A*) and aliphatic (*B*) regions of a ^{13}C-^{13}C correlation spectrum (DARR) recorded at 283 K. Carbonyl (*C*) and aliphatic (*D*) regions of a ^{13}C-^{15}N correlation spectrum (TEDOR). Carbonyl (*E*) and aliphatic (*F*) regions of a ^{13}C-^{15}N NCOCx correlation spectrum, which correlates backbone resonances with the ^{13}C resonances of the side chain of the preceding amino acid.

These differences are smaller than the experimental error in our assignments.

Quaternary Interactions at the N Terminus of NM. Having established that the structures of the full-length protein and those of the model systems are different, we set out to investigate the quaternary arrangement of full-length NM monomers in the amyloid form. ^{13}C-^{15}N MAS NMR correlation spectroscopy (TEDOR) has been successfully used for determination of quaternary interactions for a variety of amyloid fibers, such as those of Het-s (41), β2m (42), PrP (43), PI3-SH3 (32), and Aβ mutants (10). The approach is simple. Even in the absence of assignments, if the TEDOR spectra of fibers made from monomers carrying both ^{13}C and ^{15}N match the TEDOR spectra of fibers made from a mixture of ^{15}N- and ^{13}C-labeled monomers, then the chemical identities of the inter- and intramolecular sites are the same. This will be the case only if the monomers have a parallel in-register arrangement. Because the TEDOR spectra of NM fibers labeled at all 253 amino acids were too crowded for this approach (Fig. 2*A*), we turned to segmental labeling. To determine the arrangement of NM in amyloid fibrils, we compared the short-range TEDOR spectra of fibers made from uniformly segmentally labeled ^{13}C and ^{15}N monomers with the long-range TEDOR spectra of fibers made from a 1:1 mixture of uniformly ^{15}N-labeled monomers and segmentally ^{13}C-labeled monomers.

Collection of a long-range TEDOR spectrum for the mixed sample using traditional MAS NMR was limited by experimental sensitivity. This is largely due to the experimental inefficiency of measurement of long-range interactions, which depend upon the cube of the distance. However, sparse isotopic labeling can eliminate other sources of experimental inefficiencies, such as dipolar truncation (44) and signal modulation from J coupling (35). Indeed, a sparse array of ^{13}C spins can increase the efficiency of long-range TEDOR experiments nearly threefold (35, 42). We used [2-^{13}C]glucose to achieve ^{13}C labeling at the Cα-position and eliminate labeling at the adjacent C′ and Cβ-sites (45). A mixed sample made with this labeling pattern will have about one-eighth of the intermolecular ^{13}C–^{15}N pairs of a uniformly labeled sample due to the random incorporation of monomers into the polymer (∼25%) as well as the inefficiency of ^{13}Cα-labeling by [2-^{13}C]glucose (∼45%) (45). Combined with the order-of-magnitude drop in experimental efficiency of measurement of long-range versus short-range interactions, this experiment sets a high bar for sensitivity. To increase the sensitivity of this experiment, we turned to DNP NMR spectroscopy (Fig. 2 *B* and *C*).

To determine whether NM fibers are parallel in-register at the N-terminal region, we first collected a short-range TEDOR spectrum on fibrils polymerized from monomers that are segmentally labeled with [^{15}N]glucose and [2-^{13}C]glucose (Fig. 2*B*, blue). This spectrum maintained the features of a segmentally ^{13}C- and ^{15}N-labeled sample collected at room temperature (Fig. 2*B*, red), consistent with the structure not being perturbed by DNP conditions, as was seen for fibrils of the peptide derived from this protein (46). An additional strong peak (labeled U1) is consistent with the average chemical shifts of E, K, and Q. NM fibers have 75 such dynamic sites at room temperature (11), and this signal was attributed to natural abundance. We next collected a long-range TEDOR spectrum on fibrils polymerized from a 1:1 mixture of uniformly ^{15}N-labeled and segmentally [2-^{13}C]glucose–labeled monomers. Because our experiments are so sensitive and the labeled region is so small (∼5% of the full-length protein), we must account for signals derived from natural abundance of ^{13}C (∼1.1% of total carbons) and ^{15}N (0.37% of total nitrogens) that are also present in the sample. Whereas the contribution from neighboring ^{13}C and ^{15}N sites in the unlabeled section of the segmentally labeled NM molecules is negligible, the intramolecular ^{15}N–^{13}Cα sites from natural-abundance carbon in the uniformly ^{15}N-labeled protein at the ^{15}N–^{13}Cα (1.5 Å) and ^{15}N–^{13}Cα$_{(i+1)}$ (2.5 Å) sites each accounts for a little under a third of all ^{15}N–^{13}Cα pairs in the fibril sample. However, the signal intensities of such sites peak at much shorter mixing times (1.8 and ∼6 ms, respectively) (35), and have largely decayed by the longer mixing times (12.8 ms), where the intensity for long-range interactions approaches its maximum. The short-range TEDOR and the long-range TEDOR for these fibrils differ (Fig. 2*D*). Moreover, peaks that correspond to the intramolecular interactions observed in the short-range TEDOR (putatively assigned to M1, D3, S4, N5, N8, and Q11) are absent in the long-range TEDOR spectra (Fig. 2*E*). Although not all of the first 14 amino acids of NM are definitively involved in the amyloid core, most studies agree that at least half of these residues are involved (9, 10). Using site-specific reporters for quaternary structure, these experiments combined segmental isotopic labeling for resolution with DNP for sensitivity. Because the TEDOR spectra of fibers made from monomers carrying both ^{13}C and ^{15}N did not match the TEDOR spectra of fibers made from a mixture of ^{15}N- and ^{13}C-labeled monomers, NM cannot have a parallel in-register quaternary structure at the N-terminal 14 amino acids (Fig. 2*D*).

Quaternary Structure of NM Fibers. Having established that NM is not parallel in-register at the N terminus, we set out to determine whether this was also the case in other regions of the protein. To do so, we used alanine as a probe. There are 15 alanine residues in the NM sequence spread throughout the N and M domains. Five of these alanines are near residues interpreted to be in a parallel in-register quaternary structure based upon relaxation rates determined from PITHIRDS-CT experiments (Fig. 4*B*) (18). We first prepared an NM fiber sample where alanine residues were isotopically enriched at both the Cα and N positions. Short-range TEDOR spectra of these fibers revealed that whereas the majority of alanine sites had chemical shifts consistent with β-sheet or random-coil secondary structure (Fig. 4*A*), a quarter of the alanine sites had chemical shifts consistent with α-helical conformations, in line with previous observations of the M domain upon cryoprotection (47). This establishes that alanine takes on a variety of secondary structures in NM fibrils, consistent with its representation across the primary sequence of NM.

To determine the quaternary arrangement of NM monomers in the fiber, we prepared a mixed-fibril sample. We prepared NM monomers labeled with ^{13}C at only the alanine Cα-position and with ^{15}N at only the alanine backbone position. We templated a 1:1 mixture of these NM monomers into amyloid fibers. If the amyloid core of the protein is parallel in-register, then at long TEDOR mixing times (18 ms) we would expect to observe long-range (∼5-Å) interactions as a peak in the ^{15}N-filtered ^{13}C spectra at ∼52 ppm. We do not observe such a peak (Fig. 4*C*). To ensure that our experiment was sensitive enough to detect such interactions if they were present, we measured short-range interactions with natural-abundance sites in the same sample. As expected, we detected peaks at 52 and 175 ppm (Fig. 4*D*). These signals arose from natural-abundance ^{13}C and ^{15}N sites. Such sites were ∼70-fold less abundant than any hypothetical intermolecular sites. However, experiments that measure short distances (∼1.8 Å) are an order of magnitude more sensitive than those probing longer distances (>5 Å). Thus, the intramolecular natural-abundance sites adjacent to isotopically enriched alanine residues were only an order of magnitude more difficult to detect than any hypothetical intermolecular long-range interactions for a parallel in-register arrangement. Although not all alanines are predicted to be in the amyloid core, if only one alanine residue in the NM fiber was parallel in-register, the peak for an intermolecular interaction would have an intensity comparable to that of the sensitivity control. To ensure our experiment had the sensitivity to detect a single long-range interaction, we signal averaged the intermolecular experiment longer than the sensitivity control. We did not detect any intermolecular interactions between alanine residues (Fig. 4*C*). Therefore, the monomers in

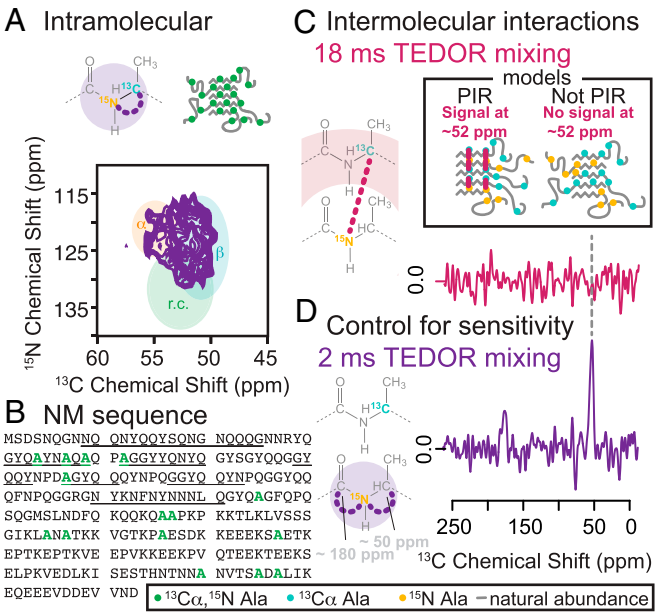

Fig. 4. Monomers of NM are not arranged in parallel in-register β-sheets in the amyloid form throughout the molecule. (*A*) The one-bond TEDOR spectrum of NM fibers prepared from proteins specifically labeled only at the Cα and N positions of alanines reveals that alanines have a variety of secondary structure conformations. (*B*) Alanine is represented through the primary sequence of NM, and a third of the alanine positions fall into a region suggested to assume a parallel in-register arrangement (underlined regions). (*C*) The long-range TEDOR spectrum of NM fibers prepared from a 1:1 mixture of protein specifically labeled at only the Cα site of alanines with protein specifically labeled at only the N site of alanines does not show any long-range interactions. (*D*) The inability to observe any signal is not due to a lack of experimental sensitivity, as evidenced by the short-range TEDOR spectrum on the same sample that reports on natural-abundance interactions, which are 100 times less common.

amyloid fibrils of lysate-templated NM did not adopt a parallel in-register arrangement.

Discussion

Yeast prions represent a paradigm-shifting, protein-based mechanism for the inheritance of biological phenotypes. A wide variety of approaches, including using small model systems, have been used to investigate Sup35 prion structures, but a consensus structural picture has yet to emerge. By combining specific and segmental isotopic labeling with sensitivity-enhanced DNP NMR, we can observe a small region of the full-length protein. We find that the conformation of the full-length protein differs from those determined for small model systems, so detailed insights derived from small model systems are not generalizable to full-length proteins. Moreover, the intermolecular interactions detected for NM are not consistent with the parallel in-register model for amyloid fibers of NM.

Structural analysis of peptide microcrystals provides atomic-level insights into amyloid structures, yet the relationship of these structures to biologically active amyloids formed from full-length proteins remains unclear. For NM, as with many other amyloid-forming proteins such as TTR (48), Tau (49), and p53 (50), much of the detailed structural work has been focused on small peptides derived from aggregation-prone regions of a full-length protein. The peptide GNNQQNY corresponds to residues 7 to 13 of the 253-residue-long NM. This 7-residue peptide segment forms amyloid-like fibrils as well as two forms of microcrystals, and has been studied by both X-ray microcrystallography (23, 49) and MAS NMR (21, 22). Prior work suggested that the structure of peptide microcrystals differs from the structure of the full-length protein because chemical shifts for residues 7 to 13 are very different (13). However, that work used de novo formed NM fibers with an N-terminal tag adjacent to the prion-forming domain. This modification is known to perturb fiber assembly (51). Using segmentally labeled untagged NM polymerized into the fiber form using seeds derived from cellular lysates from strong [*PSI*$^+$] cultures, we also found that the chemical shifts for these regions differed from those of the microcrystalline forms. Thus, the atomic details of microcrystalline GNNQQNY are not representative of those in the lysate-templated fibers.

This work presents evidence that the monomers in lysate-templated NM fibers do not adopt a parallel in-register arrangement. This is in contrast to a series of MAS NMR studies that report that the amyloid of the N-terminal domain of de novo assembled fibers is parallel in-register (12, 17, 18). All of those studies relied upon an experiment, PITHIRDS-CT, which measures dipole–dipole interactions (16). This coupling is highly distance-dependent, so closer proximity of the probed site to other isotopically labeled sites increases the rate of signal decay. Thus, after correcting for signal decay due to the presence of natural-abundance isotopes in systems of known geometry, the PITHIRDS-CT curves are readily interpreted as distances. Indeed, PITHIRDS-CT curves are so sensitive to the geometry of the coupled spin system that deviations from a linear geometry produce more rapid signal decays and lower plateau values for the same nearest-neighbor distances (16). Difference in our results could be due to sample preparation; however, for systems where the geometry is uncertain, PITHIRDS-CT dipolar recoupling rates should be interpreted with caution.

The TEDOR experiment used in this study also relies upon dipolar couplings, but measures a signal buildup that is dependent on the heteronuclear dipole coupling between spins, rather than a signal decay. Because TEDOR reports on the chemical shifts of both of the coupled spins, it provides rich information about their chemical identity. However, TEDOR is a much less sensitive experiment than PITHIRDS-CT. Moreover, although it is possible to observe a dense array of spins, too many labeled sites can hinder interpretation of the experiment due to chemical shift degeneracy. The strategy reported here combines segmental and specific isotopic labeling with DNP NMR to increase sensitivity and TEDOR to examine the intermolecular contacts in NM amyloid fibers.

Combining segmental and specific isotopic labeling with DNP NMR provides both the resolution and sensitivity to determine whether or not the NM monomers in our lysate-templated fibers are parallel in-register. The parallel in-register arrangement can be readily identified when both short-range (one-bond) and long-range (~4.8-Å interstrand) TEDOR spectra give rise to the same cross-peaks. Intermolecular interactions can be specifically interrogated using samples prepared from a mixture of ^{15}N- and ^{13}C-labeled monomers. We examined 10 residues that were predicted to be parallel in-register by PITHIRDS-CT (5 in the segmentally labeled region and 5 alanine residues present in the first 80 residues of NM). None showed any intermolecular interactions required for the parallel in-register architecture. Our experiments examined two sites, residues 13 and 35, that were also examined by the PITHIRDS-CT method (18). We examined residue 13 using a segmentally labeled molecule and implicitly examined alanine 34 using specific labeling. We did not detect intermolecular interactions in either experiment. Our data are incompatible with a parallel in-register arrangement of NM monomers in the amyloid core of our yeast lysate-templated NM fibrils.

Despite the fact that most proteins have the capacity to form amyloids under the right conditions (52) and the amyloid fold has enormous importance in many realms of biology, there are few detailed structural models for amyloid forms of any protein. Of the handful of amyloid structures, beneficial amyloids of microbial origin have a β-helical fold (41) whereas the amyloids associated with human disease have a parallel in-register arrangement (10, 42, 43). This observation sets up a paradigm of a

disease versus a beneficial amyloid, where one represents an incurable pathological trap and the other is a dynamically regulated biological switch. In this work, we used both segmental and specific isotopic labeling schemes in combination with DNP NMR to measure long-range interactions that define quaternary interactions in a large protein system. These measurements enabled us to determine that the monomers in at least one of the amyloid forms of NM do not adopt a parallel in-register arrangement. The combination of segmental and specific labeling to test structural models is easily adaptable to other large protein systems and is an exciting avenue for future biomolecular studies.

Materials and Methods

Experimental details of sample preparation and NMR spectroscopy are given in *SI Materials and Methods*.

ACKNOWLEDGMENTS. We thank members of the S.L. and R.G.G. laboratories for valuable discussions during the course of this research. S.L. was a Howard Hughes Medical Institute Investigator. K.K.F. was an HHMI Fellow of the Life Science Research Foundation, and V.K.M. was a Banting Fellow of the Natural Science and Engineering Research Council of Canada, Government of Canada. This work was supported by NIH Grants GM-025874 (to S.L.) and EB-003151, EB-002804, and EB-002026 (to R.G.G.).

1. Shorter J, Lindquist S (2005) Prions as adaptive conduits of memory and inheritance. *Nat Rev Genet* 6:435–450.
2. Cox BS (1965) Ψ, a cytoplasmic suppressor of super-suppressor in yeast. *Heredity* 20: 505–521.
3. Glover JR, et al. (1997) Self-seeded fibers formed by Sup35, the protein determinant of [*PSI*$^+$], a heritable prion-like factor of *S. cerevisiae*. *Cell* 89:811–819.
4. True HL, Berlin I, Lindquist SL (2004) Epigenetic regulation of translation reveals hidden genetic variation to produce complex traits. *Nature* 431:184–187.
5. True HL, Lindquist SL (2000) A yeast prion provides a mechanism for genetic variation and phenotypic diversity. *Nature* 407:477–483.
6. Helsen CW, Glover JR (2012) Insight into molecular basis of curing of [*PSI*$^+$] prion by overexpression of 104-kDa heat shock protein (Hsp104). *J Biol Chem* 287:542–556.
7. Liu J-J, Sondheimer N, Lindquist SL (2002) Changes in the middle region of Sup35 profoundly alter the nature of epigenetic inheritance for the yeast prion [*PSI*$^+$]. *Proc Natl Acad Sci USA* 99(Suppl 4):16446–16453.
8. Krishnan R, Lindquist SL (2005) Structural insights into a yeast prion illuminate nucleation and strain diversity. *Nature* 435:765–772.
9. Toyama BH, Kelly MJS, Gross JD, Weissman JS (2007) The structural basis of yeast prion strain variants. *Nature* 449:233–237.
10. Schütz AK, et al. (2015) Atomic-resolution three-dimensional structure of amyloid β fibrils bearing the Osaka mutation. *Angew Chem Int Ed Engl* 54:331–335.
11. Frederick KK, et al. (2014) Distinct prion strains are defined by amyloid core structure and chaperone binding site dynamics. *Chem Biol* 21:295–305.
12. Shewmaker F, Kryndushkin D, Chen B, Tycko R, Wickner RB (2009) Two prion variants of Sup35p have in-register parallel beta-sheet structures, independent of hydration. *Biochemistry* 48:5074–5082.
13. Luckgei N, et al. (2013) The conformation of the prion domain of Sup35p in isolation and in the full-length protein. *Angew Chem Int Ed Engl* 52:12741–12744.
14. King C-Y, Diaz-Avalos R (2004) Protein-only transmission of three yeast prion strains. *Nature* 428:319–323.
15. Tanaka M, Chien P, Naber N, Cooke R, Weissman JS (2004) Conformational variations in an infectious protein determine prion strain differences. *Nature* 428:323–328.
16. Tycko R (2007) Symmetry-based constant-time homonuclear dipolar recoupling in solid state NMR. *J Chem Phys* 126:064506.
17. Shewmaker F, Wickner RB, Tycko R (2006) Amyloid of the prion domain of Sup35p has an in-register parallel beta-sheet structure. *Proc Natl Acad Sci USA* 103:19754–19759.
18. Gorkovskiy A, Thurber KR, Tycko R, Wickner RB (2014) Locating folds of the in-register parallel β-sheet of the Sup35p prion domain infectious amyloid. *Proc Natl Acad Sci USA* 111:E4615–E4622.
19. Reymer A, et al. (2014) Orientation of aromatic residues in amyloid cores: Structural insights into prion fiber diversity. *Proc Natl Acad Sci USA* 111:17158–17163.
20. Comellas G, Rienstra CM (2013) Protein structure determination by magic-angle spinning solid-state NMR, and insights into the formation, structure, and stability of amyloid fibrils. *Annu Rev Biophys* 42:515–536.
21. van der Wel PCA, Lewandowski JR, Griffin RG (2007) Solid-state NMR study of amyloid nanocrystals and fibrils formed by the peptide GNNQQNY from yeast prion protein Sup35p. *J Am Chem Soc* 129:5117–5130.
22. van der Wel PCA, Lewandowski JR, Griffin RG (2010) Structural characterization of GNNQQNY amyloid fibrils by magic angle spinning NMR. *Biochemistry* 49:9457–9469.
23. Nelson R, et al. (2005) Structure of the cross-beta spine of amyloid-like fibrils. *Nature* 435:773–778.
24. Volkmann G, Iwaï H (2010) Protein *trans*-splicing and its use in structural biology: Opportunities and limitations. *Mol Biosyst* 6:2110–2121.
25. Freiburger L, et al. (2015) Efficient segmental isotope labeling of multi-domain proteins using Sortase A. *J Biomol NMR* 63:1–8.
26. Nabeshima Y, Mizuguchi M, Kajiyama A, Okazawa H (2014) Segmental isotope-labeling of the intrinsically disordered protein PQBP1. *FEBS Lett* 588:4583–4589.
27. Busche AEL, et al. (2009) Segmental isotopic labeling of a central domain in a multidomain protein by protein *trans*-splicing using only one robust DnaE intein. *Angew Chem Int Ed Engl* 48:6128–6131.
28. Otomo T, Ito N, Kyogoku Y, Yamazaki T (1999) NMR observation of selected segments in a larger protein: Central-segment isotope labeling through intein-mediated ligation. *Biochemistry* 38:16040–16044.
29. Schubeis T, Lührs T, Ritter C (2015) Unambiguous assignment of short- and long-range structural restraints by solid-state NMR spectroscopy with segmental isotope labeling. *ChemBioChem* 16:51–54.
30. Schubeis T, et al. (2015) Untangling a repetitive amyloid sequence: Correlating biofilm-derived and segmentally labeled curli fimbriae by solid-state NMR spectroscopy. *Angew Chem Int Ed Engl* 54:14669–14672.
31. Debelouchina GT, et al. (2013) Higher order amyloid fibril structure by MAS NMR and DNP spectroscopy. *J Am Chem Soc* 135:19237–19247.
32. Bayro MJ, et al. (2011) Intermolecular structure determination of amyloid fibrils with magic-angle spinning and dynamic nuclear polarization NMR. *J Am Chem Soc* 133:13967–13974.
33. Zettler J, Schütz V, Mootz HD (2009) The naturally split Npu DnaE intein exhibits an extraordinarily high rate in the protein *trans*-splicing reaction. *FEBS Lett* 583:909–914.
34. Takegoshi K, Nakamura S, Terao T (2001) ^{13}C-^{1}H dipolar-assisted rotational resonance in magic-angle spinning NMR. *Chem Phys Lett* 344:631–637.
35. Jaroniec CP, Filip C, Griffin RG (2002) 3D TEDOR NMR experiments for the simultaneous measurement of multiple carbon-nitrogen distances in uniformly (13)C,(15)N-labeled solids. *J Am Chem Soc* 124:10728–10742.
36. Egorova-Zachernyuk TA, et al. (2001) Heteronuclear 2D-correlations in a uniformly [^{13}C, ^{15}N] labeled membrane-protein complex at ultra-high magnetic fields. *J Biomol NMR* 19:243–253.
37. Shen Y, Delaglio F, Cornilescu G, Bax A (2009) TALOS+: A hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J Biomol NMR* 44:213–223.
38. Wishart DS, Sykes BD (1994) Chemical shifts as a tool for structure determination. *Methods Enzymol* 239:363–392.
39. Wang Y, Jardetzky O (2002) Probability-based protein secondary structure identification using combined NMR chemical-shift data. *Protein Sci* 11:852–861.
40. Goehlert VA, Krupinska E, Regan L, Stone MJ (2004) Analysis of side chain mobility among protein G B1 domain mutants with widely varying stabilities. *Protein Sci* 13:3322–3330.
41. Wasmer C, et al. (2008) Amyloid fibrils of the HET-s(218–289) prion form a beta solenoid with a triangular hydrophobic core. *Science* 319:1523–1526.
42. Debelouchina GT, Platt GW, Bayro MJ, Radford SE, Griffin RG (2010) Intermolecular alignment in β2-microglobulin amyloid fibrils. *J Am Chem Soc* 132:17077–17079.
43. Helmus JJ, Surewicz K, Apostol MI, Surewicz WK, Jaroniec CP (2011) Intermolecular alignment in Y145Stop human prion protein amyloid fibrils probed by solid-state NMR spectroscopy. *J Am Chem Soc* 133:13934–13937.
44. Bayro MJ, et al. (2009) Dipolar truncation in magic-angle spinning NMR recoupling experiments. *J Chem Phys* 130:114506.
45. Lundström P, et al. (2007) Fractional ^{13}C enrichment of isolated carbons using [1-^{13}C]- or [2-^{13}C]-glucose facilitates the accurate measurement of dynamics at backbone Calpha and side-chain methyl positions in proteins. *J Biomol NMR* 38:199–212.
46. Debelouchina GT, et al. (2010) Dynamic nuclear polarization-enhanced solid-state NMR spectroscopy of GNNQQNY nanocrystals and amyloid fibrils. *Phys Chem Chem Phys* 12:5911–5919.
47. Frederick KK, et al. (2015) Sensitivity-enhanced NMR reveals alterations in protein structure by cellular milieus. *Cell* 163:620–628.
48. Fitzpatrick AWP, et al. (2013) Atomic structure and hierarchical assembly of a cross-β amyloid fibril. *Proc Natl Acad Sci USA* 110:5468–5473.
49. Sawaya MR, et al. (2007) Atomic structures of amyloid cross-beta spines reveal varied steric zippers. *Nature* 447:453–457.
50. Soragni A, et al. (2016) A designed inhibitor of p53 aggregation rescues p53 tumor suppression in ovarian carcinomas. *Cancer Cell* 29:90–103.
51. Serio TR, et al. (2000) Nucleated conformational conversion and the replication of conformational information by a prion determinant. *Science* 289:1317–1321.
52. Dobson CM (2001) The structural basis of protein folding and its links with human disease. *Philos Trans R Soc Lond B Biol Sci* 356:133–145.
53. Peterson FC, Gordon NC, Gettins PGW (2001) High-level bacterial expression and ^{15}N-alanine-labeling of bovine trypsin. Application to the study of trypsin-inhibitor complexes and trypsinogen activation by NMR spectroscopy. *Biochemistry* 40:6275–6283.
54. Serio TR, Cashikar AG, Moslehi JJ, Kowal AS, Lindquist SL (1999) Yeast prion [*PSI*$^+$] and its determinant, Sup35p. *Methods Enzymol* 309:649–673.
55. Bennett AE, Rienstra CM, Auger M, Lakshmi KV, Griffin RG (1995) Heteronuclear decoupling in rotating solids. *J Chem Phys* 103:6951–6958.
56. Morcombe CR, Zilm KW (2003) Chemical shift referencing in MAS solid state NMR. *J Magn Reson* 162:479–486.
57. Pines A, Gibby MG, Waugh JS (1973) Proton-enhanced NMR of dilute spins in solids. *J Chem Phys* 59:569–590.
58. Rosay M, Blank M, Engelke F (2016) Instrumentation for solid-state dynamic nuclear polarization with magic angle spinning NMR. *J Magn Reson* 264:88–98.
59. Barnes AB, et al. (2012) Dynamic nuclear polarization at 700 MHz/460 GHz. *J Magn Reson* 224:1–7.
60. Michaelis VK, et al. (2014) Topical developments in high-field dynamic nuclear polarization. *Isr J Chem* 54:207–221.
61. Delaglio F, et al. (1995) NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* 6:277–293.
62. Goddard TD, Kneller DG (2008) SPARKY 3 (University of California, San Francisco).