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

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>As Published</td>
<td><a href="http://dx.doi.org/10.1073/PNAS.1619051114">http://dx.doi.org/10.1073/PNAS.1619051114</a></td>
</tr>
<tr>
<td>Publisher</td>
<td>Proceedings of the National Academy of Sciences</td>
</tr>
<tr>
<td>Version</td>
<td>Final published version</td>
</tr>
<tr>
<td>Accessed</td>
<td>Sun Dec 24 19:17:42 EST 2017</td>
</tr>
<tr>
<td>Citable Link</td>
<td><a href="http://hdl.handle.net/1721.1/112944">http://hdl.handle.net/1721.1/112944</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.</td>
</tr>
<tr>
<td>Detailed Terms</td>
<td></td>
</tr>
</tbody>
</table>
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,b,c,1 Vladimir K. Michaelisb,c,3 Marc A. Caporini,d,4 Loren B. Andreasb,c, Galia T. Debelouchina,b,c,5 Robert G. Griffinb,c,6 and Susan Lindquist,a,b,c,6

aWhitehead Institute for Biomedical Research, Cambridge, MA 02142; bDepartment of Chemistry, Massachusetts Institute of Technology, Cambridge, MA 02139; cFrancis Bitter Magnet Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139; dBruker Biospin, Billerica, MA 01821; eHoward Hughes Medical Institute, Massachusetts Institute of Technology, Cambridge, MA 02138; and fDepartment 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 15C- and 15N-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.

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.


The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

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

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

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

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

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

6Deceased October 27, 2016.

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

10.1073/pnas.1619051114
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 GNNQNY, 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 GNNQNY 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 (DPN) 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 13C- and 15N-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 fibril 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 expected sizes via a native peptide bond. Split inteins can be cut into segments of about 14 amino acids (Fig. 1). To segmentally label and for a 1:1 mixture of 13C- and 15N-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 (13C, 15N; green) or -inactive (13C, 14N; 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. 1. Segmental isotopic labeling reduces the complexity of solid-state NMR spectra. (A) The glycomic region of a 13C,15N correlation spectrum for uniformly 13C- and 15N-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 (13C, 15N; green) or -inactive (13C, 14N; 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.

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. 1C). These truncated versions of NM could be incorporated into seeded fibers, so a secondary purification step based upon Ni-nitritoiracetic (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 His6 tags and therefore increased affinity for Ni-NTA (Fig. 1D). After eluting the product with an imidazole gradient, the N-terminal His6 tags were cleaved off (Fig. 1E). 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 His6-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 “PS1” phenotype (11). Thus, we produced segmentally isotopically labeled NM molecules in sufficient quantities and purity for structural investigations by MAS NMR spectroscopy.

Frederick et al.
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 (4) is greatly simplified by using an NM molecule that is isotopically labeled at only the first 14 amino acids (8). Despite the large change in sample temperature and the addition of a cryoprotectant (60% D$_2$-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. 1B 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. 1B) 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 NCCOx (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+(p) predicted backbone $\phi$–$\psi$ 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 $\phi$–$\psi$ 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 monomeric 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).
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 its model systems are different, we set out to investigate the quaternary arrangement of full-length NM monomers in the amyloid form. \(^{13}\text{C}^{15}\text{N}\) 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), \(\beta2m\) (42), PrP (43), PI3-SH3 (32), and A\(\phi\) mutants (10). The approach is simple. Even in the absence of assignments, if the TEDOR spectra of fibers made from monomers carrying both \(^{13}\text{C}\) and \(^{15}\text{N}\) match the TEDOR spectra of fibers made from a mixture of \(^{15}\text{N}\)- and \(^{13}\text{C}\)-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. 2A), 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 \(^{13}\text{C}\) and \(^{15}\text{N}\) monomers with the long-range TEDOR spectra of fibers made from a 1:1 mixture of uniformly \(^{15}\text{N}\)-labeled monomers and segmentally \(^{13}\text{C}\)-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 \(^{13}\text{C}\) spins can increase the efficiency of long-range TEDOR experiments nearly threefold (35, 42). We used \(^{2}\text{H}^{13}\text{C}\)glucose to achieve \(^{13}\text{C}\) labeling at the C\text{-position} and eliminate labeling at the adjacent C\text{-sites} (45). A mixed sample made with this labeling pattern will have about one-eighth of the intermolecular \(^{13}\text{C}^{15}\text{N}\) pairs of a uniformly labeled sample due to the random incorporation of monomers into the polymer (\(\sim25\%\)) as well as the inefficiency of \(^{13}\text{C}\)-labeling by \(^{2}\text{H}^{13}\text{C}\)glucose (\(\sim45\%\)) (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. 2B 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 \(^{15}\text{N}\)glucose and \(^{2}\text{H}^{13}\text{C}\)glucose (Fig. 2B, blue). This spectrum maintained the features of a segmentally \(^{13}\text{C}\) and \(^{15}\text{N}\)labeled sample collected at room temperature (Fig. 2B, 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 O. 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 \(^{15}\text{N}\)-labeled and segmentally \(^{2}\text{H}^{13}\text{C}\)glucose–labeled monomers. Because our experiments are so sensitive and the labeled region is so small (\(\sim5\%\) of the full-length protein), we must account for signals derived from natural abundance of \(^{13}\text{C}\) (\(\sim1.1\%\) of total carbons) and \(^{15}\text{N}\) (0.3\% of total nitrogens) that are also present in the sample. Whereas the contribution from neighboring \(^{13}\text{C}\) and \(^{15}\text{N}\) sites in the unlabeled section of the segmentally labeled NM molecules is negligible, the intramolecular \(^{13}\text{N}^{15}\text{C}\) sites from natural-abundance carbon in the uniformly \(^{15}\text{N}\)-labeled protein at the \(^{13}\text{N}^{15}\text{C}\) (1.5 Å) and \(^{13}\text{N}^{15}\text{C}\) (2.5 Å) sites each accounts for a little under a third of all \(^{13}\text{N}^{12}\text{C}\) pairs in the fibril sample. However, the signal intensities of such sites peak at much shorter mixing times (1.8 and \(\sim6\) 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. 2D). 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. 2E). 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 \(^{13}\text{C}\) and \(^{15}\text{N}\) did not match the TEDOR spectra of fibers made from a mixture of \(^{13}\text{N}\)- and \(^{13}\text{C}\)-labeled monomers, NM cannot have a parallel in-register quaternary structure at the N-terminal 14 amino acids (Fig. 2D).

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 is 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 PITHRDS-CT experiments (Fig. 4B) (18). We first prepared an NM fiber sample where alanine residues were isotopically enriched at both the C\text{r} and N\text{p} positions. Short-range TEDOR spectra of these fibers revealed that whereas the majority of alanine sites had chemical shifts consistent with \(\beta\)-sheet or random-coil secondary structure (Fig. 4A), a quarter of the alanine sites had chemical shifts consistent with \(\alpha\)-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 \(^{13}\text{C}\) at only the alanine \(\text{C}r\)-position and with \(^{15}\text{N}\) 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 (\(\sim5\) Å) interactions as a peak in the \(^{15}\text{N}\)-filtered \(^{13}\text{C}\) spectra at \(\sim52\) ppm. We do not observe such a peak (Fig. 4C). 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. 4D). These signals arose from natural-abundance \(^{13}\text{C}\) and \(^{15}\text{N}\) sites. Such sites were \(\sim70\%\) less abundant than any hypothetical intermolecular sites. However, experiments that measure short distances (\(\sim1.8\) 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. 4C). Therefore, the monomers in
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 GNNQNY corresponds to residues 7 to 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
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-000206 (to R.G.G.).