Dynamic structural order of a low-complexity domain facilitates assembly of intermediate filaments

Vasiliy O. Sysoev[a], Masato Kato[a,b], Lillian Sutherland[a], Rong Hu[c], Steven L. McKnight[a,1], and Dylan T. Murray[c,1]

[a]Department of Biochemistry, UT Southwestern Medical Center, Dallas, TX 75390; [b]Institute for Quantum Life Science, National Institutes for Quantum and Radiological Science and Technology, 263-8555 Chiba, Japan; and [c]Department of Chemistry, University of California, Davis, CA 95616

Contributed by Steven L. McKnight, August 10, 2020 (sent for review May 22, 2020; reviewed by Leonard J. Mueller and Ueli Schibler)

The coiled-coil domains of intermediate filament (IF) proteins are flanked by regions of low sequence complexity. Whereas IF coiled-coil domains assume dimeric and tetrameric conformations on their own, maturation of eight tetramers into cylindrical IFs is dependent on either "head" or "tail" domains of low sequence complexity. Here we confirm that the tail domain required for assembly of *Drosophila* Tm1-I/C IFs functions by forming labile cross-β interactions. These interactions are seen in polymers made from the tail domain alone, as well as in assembled IFs formed by the intact Tm1-I/C protein. The ability to visualize such interactions in situ within the context of a discrete cellular assembly lends support to the concept that equivalent interactions may be used in organizing other dynamic aspects of cell morphology.

intermediate filaments | cross-beta polymerization | low-complexity proteins | solid-state NMR

Deposition of germ granules at the posterior tip of *Drosophila melanogaster* oocytes specifies formation of cells of the germinal lineage (1). Forward genetic studies have illuminated the importance of a number of genes essential to germ cell formation. Many of these genes encode RNA-binding proteins, of which RNA granules are themselves composed (2–5). Perplexingly, mutations proximal to the locus encoding the fly tropomyosin gene also impede the deposition of germ granules and subsequent formation of germ cells (6). These mutations interfere with the formation of a nonmuscle isoform of tropomyosin, designated Tm1-I/C, that is somehow required for germ cell specification (7, 8).

A program of alternative pre-mRNA splicing allows fly oocytes to produce an isoform of tropomyosin that replaces the domains necessary for interaction with troponin and the resulting sensitivity to regulation of muscle contraction by free calcium with protein segments of low sequence complexity. Previously reported experiments have shown that the oocyte-specific Tm1-I/C isoform assembles into intermediate filaments, and that its C-terminal low-complexity (LC) domain is essential for filament assembly (7). Here we present a combination of biochemical and biophysical studies providing evidence that the C-terminal LC domain of Tm1-I/C facilitates the assembly of intermediate filaments by forming structurally specific cross-β interactions that are unusually dynamic. These studies of the fly Tm1-I/C isoform of tropomyosin may be relevant to the behavior of prototypic intermediate filaments long studied in a variety of more complex organisms. They may further be instructive as to the behavior of LC domains in many other aspects of cell morphology.

Results

An oocyte-specific program of alternative pre-mRNA splicing allows production of an isoform of tropomyosin, Tm1-I/C, that differs markedly from that made in muscle cells. Whereas the two proteins share sequences specifying a central, coiled-coil domain, alternative splicing yields a unique C-terminal tail domain within the oocyte-specific Tm1-I/C isoform. This tail domain is 69 residues long, of low sequence complexity (18 asparagine residues, 15 serine residues, and 6 threonine residues; *SI Appendix*, Fig. S1), evolutionarily limited to the *Sophophora* subgenus of *Drosophila*, and required for the protein to assemble into intermediate filaments (7).

Expression, purification, and incubation of the isolated Tm1-I/C tail domain leads to the assembly of amyloid-like polymers. X-ray diffraction studies of Tm1-I/C tail domain polymers have revealed a prominent diffraction ring at 4.7 Å. These and other studies led to the conclusion that the Tm1-I/C tail domain can, in an isolated state, assemble into amyloid-like polymers that are labile to disassembly (7). In this regard, the Tm1-I/C tail domain displays properties similar to the LC domains of numerous RNA-binding proteins (9, 10), the FG repeats forming the permeability filter of nucleopores (11–14), and six different intermediate filament proteins expressed in mammalian cells and tissues (15).

The Tm1-I/C Tail Domain Adopts a Similar Structural Conformation in Isolation and in Assembled Intermediate Filaments. Uniform labeling of the Tm1-I/C tail domain with ^{13}C and ^{15}N isotopes allowed analysis of amyloid-like polymers by solid-state NMR (ss-NMR) spectroscopy. That the polymers adopt a partially ordered conformation is confirmed by the spectra shown in Fig. 1*A*. When analyzed by ss-NMR at 16 °C, the Tm1-I/C tail domain-only polymer sample gave rise to strong signals in the spectrum recorded using ^{1}H-^{13}C cross polarization (CP) and high-power ^{1}H decoupling (16, 17). The spectrum likewise gave rise to

Significance

The main point of our manuscript is focused on the structure of the low-complexity (LC) domain of the Tm1-I/C intermediate filament protein in the context of assembled intermediate filaments. We found that the LC tail domain of Tm1-I/C exists in precisely the same cross-β conformation within its proper biologic assembly as it does in labile, amyloid-like polymers made from the tail domain alone. This science represents a conceptually distinct advance that may form the cornerstone understanding of how the thousands of LC domains expressed in eukaryotic cells operate in a mechanistic sense, and stands in conflict with previous research claiming that LC domains function in the absence of molecular structure.

Author contributions: M.K., S.L.M., and D.T.M. designed research; V.O.S. and L.S. cloned, mutagenized, and expressed all recombinant proteins in both unlabeled and isotopically labeled forms; V.O.S. assembled all tail domain-alone polymers and intermediate filaments formed from the Tm1-I/C protein; R.H. and D.T.M. recorded and interpreted all NMR spectra; and V.O.S., M.K., S.L.M., and D.T.M. wrote the paper.

Reviewers: L.J.M., University of California Riverside; and U.S., University of Geneva, Sciences III.

The authors declare no competing interest.

[1]To whom correspondence may be addressed. Email: Steven.McKnight@UTSouthwestern.edu or dtmurray@ucdavis.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2010000117/-/DCSupplemental.

First published September 9, 2020.

intense signals when recorded using ^{1}H-^{13}C insensitive nuclei enhanced by polarization transfer (INEPT) without ^{1}H decoupling (18). Signals appearing in the CP-based spectrum arise from immobilized regions of the protein in the polymers. Signals appearing in the INEPT-based spectrum arise from highly mobile regions of the protein in the polymers. These data indicate that there are both well-ordered and conformationally dynamic regions within Tm1-I/C tail domain polymers.

To study the conformation of the Tm1-I/C tail domain in the context of fully assembled intermediate filaments, we used a split-intein system to produce a segmentally labeled form of the intact Tm1-I/C protein (19). The tail domain, including residues 373 to 441, was expressed separately from the N terminus of the protein (residues 1 to 372). The two segments were chemically ligated via the reaction scheme depicted in *SI Appendix*, Fig. S2, yielding a full-length Tm1-I/C protein containing a single asparagine-to-cysteine change at residue 371. This procedure allowed selective incorporation of ^{13}C and ^{15}N isotopes into the tail domain for the purpose of spectroscopically probing both its structure and dynamics within the full-length Tm1-I/C protein as it exists in assembled intermediate filaments. The native Tm1-I/C protein assembles into intermediate filaments 13 to 16 nm in diameter (7). *SI Appendix*, Fig. S2 shows that segmentally, ^{13}C/^{15}N-labeled Tm1-I/C intermediate filaments are morphologically indistinguishable from those assembled from the native protein as viewed by electron microscopy.

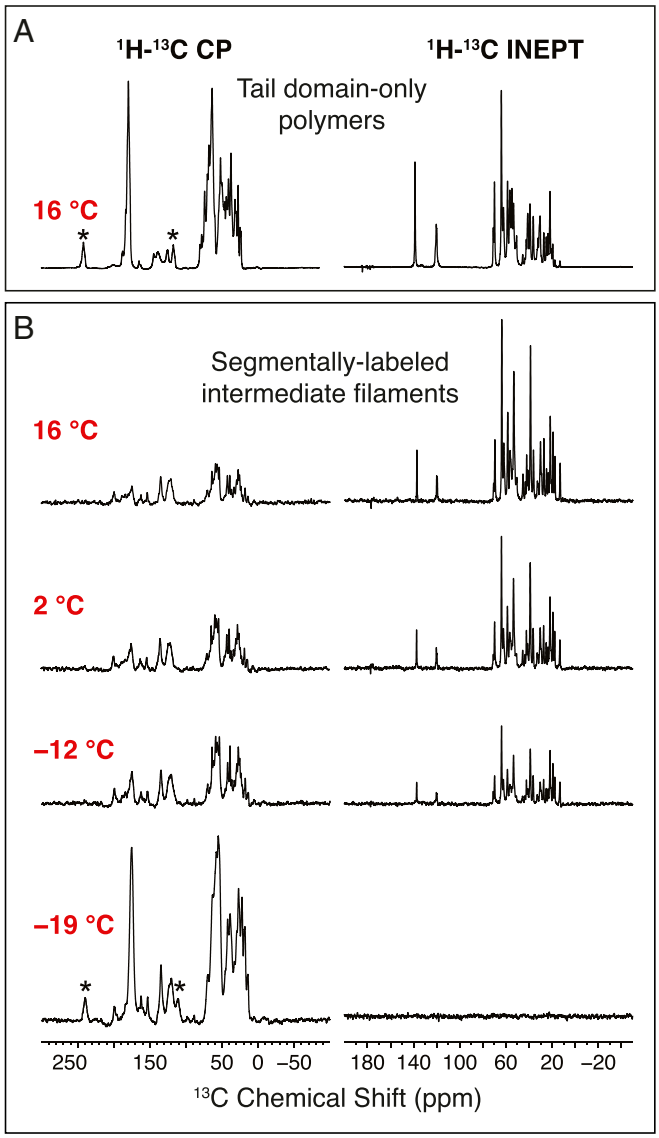

Fig. 1. One-dimensional CP- and INEPT-based ss-NMR spectra of the Tm1-I/C tail domain. (*A*) CP (*Left*) and INEPT (*Right*) spectra of ^{13}C/^{15}N-labeled Tm1-I/C tail domain-only polymers recorded at 16 °C. (*B*) Four sets of spectra representing CP (*Left*) and INEPT (*Right*) spectra of segmentally ^{13}C/^{15}N-labeled Tm1-I/C protein assembled into mature intermediate filaments. Isotopic labeling was restricted to the Tm1-I/C tail domain by the use of intein chemistry as diagrammed in *SI Appendix*, Fig. S2*A*. The CP and INEPT spectra of fully assembled, segmentally labeled Tm1-I/C intermediate filaments were recorded at 16, 2, −12, and −19 °C as displayed from top to bottom. Asterisks in the top and bottom CP spectra indicate magic-angle spinning side bands.

Assembled, segmentally ^{13}C/^{15}N-labeled Tm1-I/C intermediate filaments were packed by centrifugation into an NMR rotor and analyzed at temperatures ranging from 16 to −19 °C (Fig. 1*B*). The initial temperature of 16 °C was chosen based on previous ss-NMR studies on α-synuclein amyloid fibrils that show enhanced signal intensities for CP-based spectra at reduced temperatures due to decreased small-amplitude motions of protein backbone and sidechains (20, 21). Weak CP signals for the sample analyzed at 16 °C indicate that few regions within the Tm1-I/C tail domain are sufficiently immobilized to allow significant CP magnetization transfers. The strong signals in the INEPT-based spectrum at 16 °C are indicative of highly dynamic regions within the Tm1-I/C tail domain as it exists in assembled intermediate filaments. As a function of temperature reduction to 2, −12, and −19 °C, the strength of CP signals increased and that of INEPT signals decreased. These temperatures were chosen because they illustrate when significant changes were observed in the CP- or INEPT-based spectra. The fact that these temperatures do not correspond to "regular" increments reflects a nonlinear response of the actual sample temperature to settings on the NMR spectrometer. These observations are consistent with reduced molecular motion of the Tm1-I/C tail domain in intermediate filament assemblies as the sample temperature was reduced. The overall profile of the CP-based spectra of the Tm1-I/C tail domain as viewed in isolated, tail domain-only polymers is similar to its conformation in ^{13}C/^{15}N-labeled intermediate filaments as assayed at low temperatures (Fig. 1). These one-dimensional spectra are, however, inadequate for reporting on the actual molecular structure of the tail domain in the two different assemblies.

To more carefully probe both the uniformly ^{13}C/^{15}N-labeled Tm1-I/C tail domain-only polymers and assembled, segmentally labeled Tm1-I/C intermediate filaments, we performed two-dimensional (2D) ss-NMR to measure ^{13}C chemical shifts from rigid sites within the assemblies. Fig. 2*A* shows a 2D ^{13}C-^{13}C CP-based dipolar assisted rotational resonance (CP-DARR) (22, 23) spectrum with 50-ms mixing time for the Tm1-I/C tail domain-only polymers. Off-diagonal cross-peaks in CP-DARR spectra recorded with this mixing time arise from rigid and structured carbon atoms that are one to three bonds apart in the protein. The spectrum of the Tm1-I/C tail domain-only polymer sample exhibited sharp NMR resonances typical of a well-ordered protein structure, with full-width-half-maximum (FWHM) line widths of 175 to 225 Hz (0.8 to 1.2 ppm). Signals attributable to threonine, serine, isoleucine, cysteine, asparagine, aspartic acid, lysine, glutamic acid, and alanine are clearly present in the spectrum. The ^{13}C chemical shift values of these signals are consistent with β-strand conformations for all sites that could be unambiguously assigned to a single amino acid type.

That the chemical shift peaks observed in Tm1-I/C tail domain polymers may be diagnostic of bona fide biologic structure is supported by ss-NMR spectra obtained from an ether-precipitated sample of the same protein (Fig. 2*B*). Whereas the ether-precipitated sample revealed chemical shifts for numerous amino acid side chains within the Tm1-I/C tail domain, few

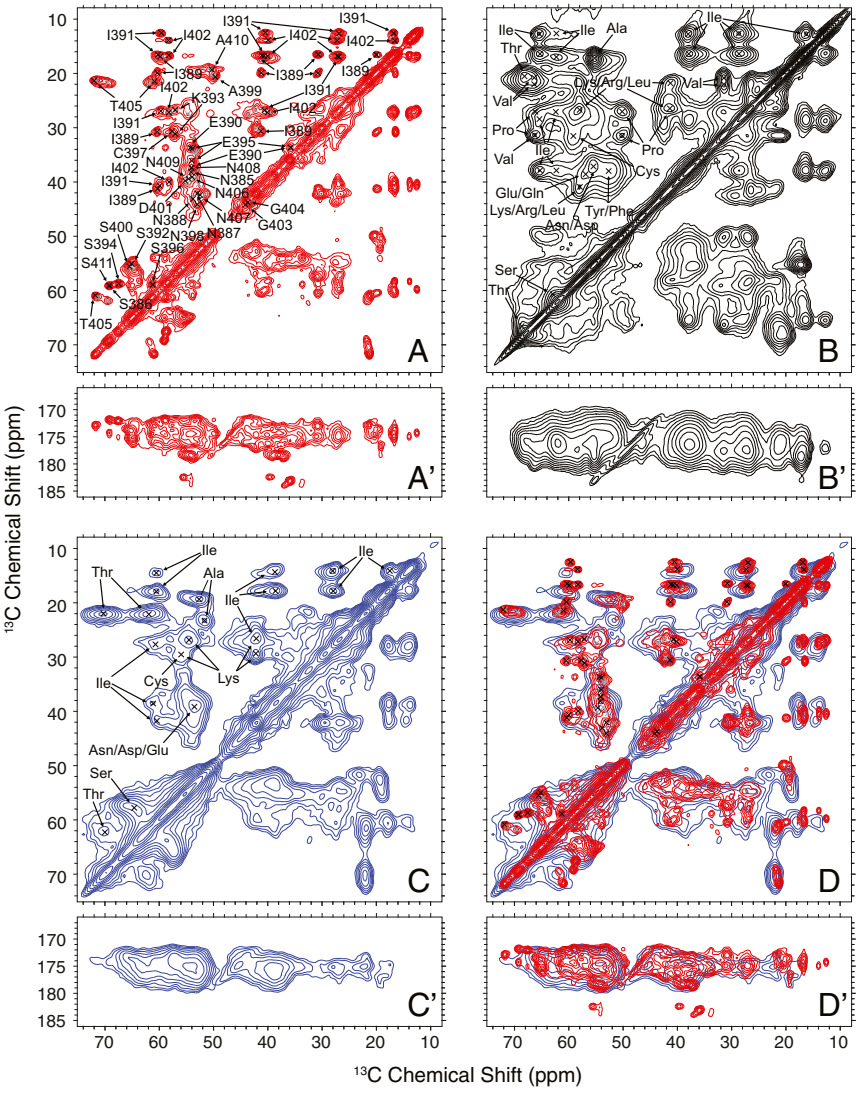

Fig. 2. Two-dimensional CP-based ss-NMR spectra of the Tm1-I/C tail domain. (*A* and *A′*) ^{13}C-^{13}C CP-DARR spectrum with 50-ms mixing time of ^{13}C/^{15}N-labeled Tm1-I/C tail domain-only polymers recorded at 16 °C. Sequence-specific resonance assignments determined in this work are indicated by amino acid type and location within the full-length Tm1-I/C protein. (*B* and *B′*) The same spectrum of an ether-precipitated sample of the Tm1-I/C tail domain as analyzed in its polymeric form in *A*. Amino acid residues specified in the spectrum of ether precipitated protein were not assigned sequence-specifically. (*C* and *C′*) ^{13}C-^{13}C CP-DARR spectrum with 50-ms mixing time of segmentally ^{13}C/^{15}N-labeled Tm1-I/C protein that had been assembled into intermediate filaments recorded at −19 °C (*SI Appendix*, Fig. S2). Amino acid residues specified in the spectrum of segmentally labeled Tm1-I/C intermediate filaments were not assigned sequence-specifically. (*D* and *D′*) An overlay of the spectra shown in *A* and *A′* and *C* and *C′*. The extensive overlap of signal intensity in the two samples indicates that the Tm1-I/C tail domain exists in highly similar structural states in the two assemblies. Sequence-specifically assigned amino acids from the data shown in *A* and *A′* are indicated by small X's in the overlay of *D* and *D′*. The spectra of the Tm1-I/C tail domain within assembled intermediate filaments (*C* and *C′*) overlap poorly or not at all with the spectra of ether-precipitated Tm1-I/C tail domain (*B* and *B′*). An overlay of the latter two spectra is shown in *SI Appendix*, Fig. S3.

such peaks overlapped with those observed for tail domain-only polymers. The most noticeable differences in the respective spectra shown in Fig. 2 *A* and *B* are significant reductions of serine and threonine β-strand signals, the appearance of a set of proline and valine signals, and a shift in the CA position of isoleucine signals.

The 2D ^{13}C-^{13}C CP-DARR spectrum of intermediate filaments prepared from the full-length, segmentally labeled sample of Tm1-I/C exhibited broad NMR resonances, with FWHM line widths >300 Hz at −19 °C (Fig. 2*C*). From this spectrum alone, it could not be determined whether the intensity in the spectrum arose from (1) multiple, single-site resonances with relatively narrow line widths arising from multiple sites with well-defined conformations or (2) a single site with a broad line shape, indicating a significant degree of conformational heterogeneity. For both of these cases, the center positions of the signal intensities were consistent with β-strand conformations for the residues giving rise to the signals.

Despite the aforementioned uncertainty, the spectrum in Fig. 2*C* is consistent with a dynamic β-strand conformation for the Tm1-I/C tail domain in the intermediate filament assemblies analyzed at 16 °C, with slowed molecular motions as recorded at the reduced temperature of –19 °C. In Fig. 2*D*, an overlay of the 2D ^{13}C-^{13}C CP-DARR spectra of the isolated Tm1-I/C tail domain-only polymers recorded at 16 °C with that of the tail domain as recorded at −19 °C in segmentally labeled intermediate filaments shows highly similar positions of CP signal intensities in the two spectra. These similarities indicate that the Tm1-I/C tail domain in the intermediate filament assemblies, when evaluated at reduced temperature, converges toward a

rigid structural conformation similar to that of the Tm1-I/C tail domain-only polymers. Differences between the two spectra include one missing set of isoleucine peaks and slight differences in the ^{13}C chemical shifts of the CG1 and CG2 cross-peaks for isoleucine residues. In stark contrast, an overlay of the spectra of the segmentally labeled Tm1-I/C protein in its assembled state (Fig. 2*C*) with that in the ether-precipitated state of the Tm1-I/C tail domain (Fig. 2*B*) shows dramatic differences (*SI Appendix*, Fig. S3). Despite being chemically identical, the Tm1-I/C tail domain exhibits obvious structural differences in these two states.

Structural Characterization of the Cross-β Core-Forming Segment of the Tm1-I/C Tail Domain. A set of 2D and three-dimensional (3D) CP-based ^{15}N-^{13}C correlation spectra with 50-ms ^{13}C-^{13}C DARR mixing collected on the Tm1-I/C tail domain-only polymer sample at 16 °C allowed identification and structural characterization of the core-forming segment of the polymers. Fig. 3 *A* and

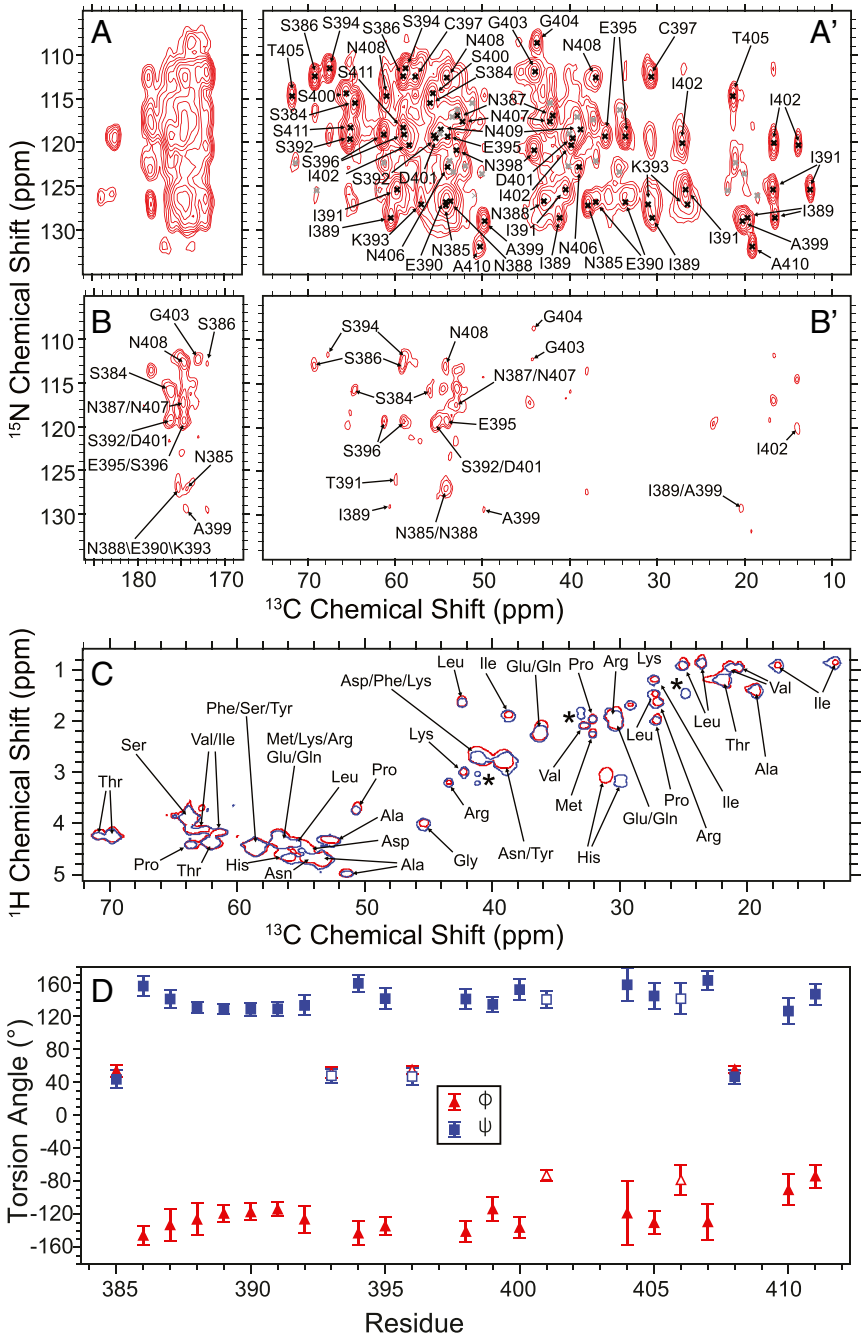

Fig. 3. Structural characterization of Tm1-I/C tail domain polymers. (*A* and *A′*) 2D CP-based NCACX spectrum of $^{13}C/^{15}N$-labeled Tm1-I/C tail domain-only polymers measured at 16 °C. (*B* and *B′*) 2D ^{15}N-^{13}C zf-TEDOR spectrum of mixed ($^{13}C/^{14}N$)- and ($^{12}C/^{15}N$)-labeled Tm1-I/C tail domain-only polymers measured at 5 °C. (*C*) 2D ^{1}H-^{13}C INEPT-MAS spectra of $^{13}C/^{15}N$-labeled Tm1-I/C tail domain-only polymers (red contours) overlapped with segmentally $^{13}C/^{15}N$-labeled Tm1-I/C intermediate filaments (blue contours). Both spectra shown in *C* were recorded at 16 °C. (*D*) NMR chemical shift-based TALOS-N torsion angle predictions. Solid triangles and squares represent "strong" β-strand predictions, and open triangles and squares represent "generous" β-strand predictions determined according to published criteria (31).

A′ shows the CP-based 2D NCACX spectrum of the Tm1-I/C tail domain polymers (24). The black and gray X's on the spectrum in Fig. 3*A′* indicate 39 sets of signals observed in the 3D implementation of the NCACX experiment. Signals that have been sequence-specifically assigned are indicated by black X's and labels in Fig. 3*A′* (*Materials and Methods* and *SI Appendix*, Table S1).

These assignments were obtained using four CP-based spectra: 2D and 3D NCACX and 2D and 3D NCOCX (24). Representative 2D slices from the 3D NCACX spectrum are shown in *SI Appendix*, Fig. S4. The 2D NCOCX spectrum and representative slices from the 3D NCOCX spectrum are shown in *SI Appendix*, Fig. S5. Of the 39 sets of signals observed in these spectra, 38 can be unambiguously assigned to a single amino acid type. Complete lists of the observed NCACX signals are shown in *SI Appendix*, Tables S1 and S2. Two alanine, three isoleucine, eight asparagine, one lysine, two glutamic acid, seven serine, one threonine, one cysteine, and two glycine residues were sequence-specifically assigned to residues 384 to 411 of the Tm1-I/C tail domain. These amino acid residues represent roughly 40% of the total amino acid content of the Tm1-I/C tail domain sequence.

Our assignment procedure was performed computationally using the MCASSIGN algorithm (25) in an iterative fashion comparable to our previous work on fused in sarcoma (FUS) and heterogeneous ribonucleoprotein A2/B1 (hnRNPA2) LC domain polymers (26, 27) to yield statistically significant and unambiguous assignments for residues 384 to 411. The assignment procedure is described in detail in *Materials and Methods*. Of note, no signals were consistently assigned to any region of the N-terminal His-tag sequence throughout the assignment calculations. The remaining unassigned signals have either relatively low signal-to-noise (values of ∼10 compared with the assigned signals, which mostly range from ∼30 to ∼60) or correspond to strong signals from asparagine residues ambiguously assigned to the stretch of five consecutive asparagines located between positions 379 and 383. The low signal-to-noise peaks arise from alanine or threonine residues present in the unassigned region immediately following residue 411. *SI Appendix*, Tables S1 and S2 list the assigned and unassigned signals with chemical shift and signal-to-noise values from the 3D NCACX spectrum.

A ^{15}N-^{13}C zf-TEDOR spectrum (28) of the Tm1-I/C tail domain polymers at 5 °C is shown in Fig. 3*B*. For this experiment, the Tm1-I/C tail domain polymers were prepared from a solution of protein in which 50% of the molecules were uniformly ^{15}N isotopically labeled and 50% of the monomers were uniformly ^{13}C isotopically labeled. The Tm1-I/C tail domain polymers resulting from this preparation should have on average alternating ^{13}C- and ^{15}N-labeled molecules. Signals in this spectrum arise from ^{13}C and ^{15}N nuclei that are ≤5 Å apart in space (28). In the aliphatic carbon region, we observed signals with ^{13}C and ^{15}N chemical shifts consistent with the assignments for S384, N385, S386, N387, N388, I389, E390, T391, S392, S394, E395, A399, D401, I402, G403, G404, N407, and N408. In the carbonyl carbon region, we observed ^{13}C and ^{15}N chemical shifts consistent with the assignments for residues S384, N385, N387, N388, E390, S392, K393, E395, S396, A399, S400, D401, G403, T405, N407, and N408. The observed signal-to-noise ratios between 5 and 15 are consistent with the signals arising from long-range, 4- to 5-Å ^{13}C-^{15}N distances during the 6.0-ms mixing time used in this experiment (21, 28). Of these signals, the only set of ^{13}CA, ^{13}CO, and amide ^{15}N chemical shift values that were ambiguous between two different residues were for S392 and D401. The unambiguous signals observed in the spectrum in Fig. 3*B* are consistent with the same residues in adjacent monomers in the Tm1-I/C polymers with ^{15}N and ^{13}C nuclei being separated by 4 to 5 Å (28) (i.e., the ^{13}C and ^{15}N chemical shifts are assigned to the same amino acid). Slices illustrating the signal-to-noise in the zf-TEDOR spectrum are shown in *SI Appendix*, Fig. S6.

Fig. 3*C* shows an overlay of 2D INEPT-based ^{1}H-^{13}C spectra of the Tm1-I/C tail domain at 16 °C (red) and the assembled, segmentally labeled Tm1-I/C intermediate filaments also analyzed at 16 °C (blue). Only a single contour is shown, to clarify the peak positions in the two spectra. Residue type assignments are indicated based on random coil chemical shift values for CA and CB sites (29) and on the average chemical shifts from the Biological Magnetic Resonance Bank for CG, CD, and CE sites (30). We observed signals from every residue type in the tail domain outside of the region assigned to the immobilized core of the polymers, including the 6xHis tag and linker.

The two spectra have nearly identical peak positions for all sites. There are subtle differences in the positions of the signals in the histidine CA and CB, aspartic acid CA, asparagine CA, and leucine CA regions. Five extra signals were observed in the spectrum of the assembled intermediate filament sample, as indicated with asterisks in Fig. 3*C*. The only signals in the 6xHis tag region not in the Tm1-I/C tail domain were methionine, glutamine, tyrosine, and phenylalanine. It was not possible to unambiguously assign signals to these residues. Regardless, the spectra in Fig. 3*C* are consistent with both the 6xHis tag and regions outside residues 384 to 411 being highly dynamic in the polymer assemblies. For the intermediate filament assemblies, the high degree of similarity with the spectrum of the Tm1-I/C tail domain polymer assemblies indicates that there are nearly identical regions of structural disorder in the two samples.

A TALOS-N prediction of backbone φ/ψ torsion angles based on the assigned ^{13}C and ^{15}N NMR chemical shifts from the CP-based spectra (31) is shown in Fig. 3*D*. The analysis uses a database of known protein structures to predict the torsion angles using amide ^{15}N, ^{13}CA, ^{13}CB, and ^{13}CO chemical shifts. Residues N387 to S392, N398 to A399, and G404 have β-strand secondary structure predictions and the extended regions of S386 to S392, S394 to E395, N398 to D401, G404 to N407, and A410 to S411 have backbone φ/ψ torsion angle predictions that are characteristic of extended β-strands. The zf-TEDOR measurements, the TALOS chemical shift analysis, and previous X-ray diffraction studies (7) confirm that Tm1-I/C tail domain-only polymers adopt a cross-β structural conformation.

Functional Assays of Deletion Mutations in the LC Tail Domain of Tm1-I/C. To study the functional importance of the 69 residues specifying the Tm1-I/C tail domain, we prepared and analyzed a series of deletion mutants. Our mutations systematically deleted sequences starting at the C terminus of the Tm1-I/C tail domain, with individual mutants sequentially removing five additional residues. Ten deletion mutants were assayed to measure both the formation of cross-β polymers made from the isolated tail domain and the formation of intermediate filaments made from the otherwise intact Tm1-I/C polypeptide.

Three changes in tail domain-only polymer formation were observed as a consequence of ever-greater deletion of C-terminal tail domain residues (Fig. 4). Crossing a boundary between residues 425 and 430 yielded polymers that changed the morphology from smooth to twisted. Crossing a second boundary between residues 405 and 410 yielded polymers that changed the morphology from twisted to spiky. Finally, crossing a boundary between residues 390 and 395 yielded no cross-β polymers at all.

It was likewise observed that Tm1-I/C intermediate filament assembly also suffered sequential impediments as a function of C-terminal deletion of the Tm1-I/C tail domain (Fig. 5). Crossing a boundary between residues 420 and 425 yielded proteins that changed filament morphology from normal to an aggregated, hairball configuration, and crossing a boundary between residues 400 and 405 yielded variant forms of the Tm1-I/C protein that failed to assemble into intermediate filaments.

Discussion

Here we report studies of an intermediate filament produced in the oocytes of *D. melanogaster*. Exons of the tropomyosin gene are alternatively spliced in oocytes, producing an unusual isoform of tropomyosin. Other than nuclear lamins, the Tm1-I/C protein is the sole intermediate filament protein found in fruit flies. The central, coiled-coil domain of the Tm1-I/C isoform of tropomyosin is flanked by LC sequences. Instead of interacting with troponin in a manner allowing for calcium-mediated regulation of muscle contraction, the C-terminal LC domain directs the Tm1-I/C protein to assemble into intermediate filaments.

Why is proper deposition of germ granules at the posterior pole of *Drosophila* oocytes dependent on Tm1-I/C-specified intermediate filaments? Visualization of a GFP-tagged form of Tm1-I/C in living oocytes has shown that the protein becomes precisely restricted to the posterior tip of oocytes well before the completion of polar granule deposition (8). We speculate that assembled Tm1-I/C intermediate filaments localized to the posterior pole of fly oocytes might constitute a Velcro-like landing pad for one or more of the constituent RNA-binding proteins specifying polar granules.

In an architectural sense, assembled intermediate filaments contain repeating collars of LC domain sequences circumferentially displayed along their axial length (*SI Appendix*, Fig. S1). In the case of vimentin intermediate filaments, we have observed that these LC domain collars represent repetitively organized binding sites for a GFP:FUS fusion protein, such that the latter protein can be iteratively bound to the filaments at 45 nm intervals as viewed by transmission electron microscopy (15). It is perhaps of importance that binding of the GFP:FUS protein to vimentin intermediate filaments is dependent upon the integrity of the LC tail domain of vimentin. As such, we speculate that Tm1-I/C intermediate filaments, on localization to the posterior tip of fly oocytes, may allow for the subsequent binding and organization of the germ granules themselves.

The experiments described in this paper were focused on the C-terminal LC domain of the Tm1-I/C protein, herein referred to as the tail domain. The Tm1-I/C tail domain has already been shown to be required for Tm1-I/C to assemble into intermediate filaments, and the isolated Tm1-I/C tail domain is known to form labile cross-β polymers (7). The primary focus of the present study was to determine whether the same structural forces leading to the formation of cross-β polymers observed in studies of the isolated Tm1-I/C tail domain might also be involved in the assembly of Tm1-I/C intermediate filaments.

We approached this central question via the use of ss-NMR spectroscopy. We compared ss-NMR spectra observed from isotopically-labeled tail domain-only polymers with spectra derived from fully assembled intermediate filaments. In the latter case, we used intein chemistry to restrict isotopic labeling to only the tail domain of the intact Tm1-I/C protein. The $^{13}C/^{15}N$ isotopic labels segmentally introduced into the intact Tm1-I/C protein were identical to those introduced into tail domain-only polymers.

We highlight four observations from these studies. First, highly similar spectra diagnostic of structural order were observed in both tail domain-only cross-β polymers and segmentally labeled Tm1-I/C intermediate filaments (Fig. 2). The portion of the Tm1-I/C tail domain specifying these cross-β interactions was mapped to a region of 28 amino acids spanning residues 384 to 411 of the Tm1-I/C polypeptide (Fig. 3). Second, evidence of robust molecular order within the tail domain was observed only upon cooling of the Tm1-I/C intermediate filaments (Fig. 1). Third, roughly 30 amino acid residues were observed to exist in a state of distinct structural disorder in both the tail domain-only cross-β polymers and the fully assembled, segmentally labeled intermediate filaments. Importantly, there was almost perfect overlap of the spectra reporting on these disordered residues in the two structures (Fig. 3). Fourth, the observed φ/ψ torsion angles based on ^{13}C and ^{15}N chemical shifts for almost all residues within the structurally ordered tail domain of Tm1-I/C were predictive of either β-strand secondary structure or an extended β-strand–like conformation. Many of these sites have 4- to 5-Å intermolecular distances appropriate for a parallel, in-register, cross-β structure (Fig. 3).

Our parsimonious interpretation of these observations is that Tm1-I/C intermediate filaments assemble via the formation of dynamically ordered cross-β interactions specified by the amino acid sequence of the C-terminal tail domain. When studied in the context of isolated, tail domain-only polymers, these cross-β interactions remain labile to disassembly but are far more stable than those formed in the context of biologically relevant Tm1-I/C intermediate filaments.

This interpretation posits an unusual form of protein structure. We hereby offer that what is being observed in this study is a form of protein structure that retains an ordered ground state sufficiently precise to dictate both molecular order and functional specificity. Curiously, however, the amino acid residues specifying molecular order are apparently in a state of continuous flux in and out of the structurally ordered state. This behavior should not be conflated with a transition of the dynamically ordered residues into disordered residues. The results presented here illustrate that these regions of differential order and disorder arise from entirely different, mutually exclusive parts of the C-terminal tail domain of Tm1-I/C. We speculate that what we have observed in studies of the LC domain-specifying function of the C-terminal tail domain of the Tm1-I/C protein may be instructive for the thousands of LC domains operative in many other aspects of cell biology.

Given the temperature-dependent behavior of the dynamically ordered region of the Tm1-I/C C-terminal tail domain, it is of interest to consider the behavior of the protein at more physiological temperatures for *Drosophila* (>16 °C). The data presented here for temperatures ranging from −19 °C to 16 °C indicate that the same conformation of the dynamically ordered segments is maintained as the temperature increases, with the protein acquiring increasingly rapid small-amplitude backbone and sidechain motions rather than a complete disassembly of the structured state. In the context of studies on protein dynamics at elevated temperatures (32–34), these motions would be expected to become increasingly more rapid above 16 °C. As would occur for most structured proteins, at some temperature a transition to a thermally denatured state is expected, but that temperature is likely higher than the physiological temperatures relevant for *Drosophila*.

It is also important to consider that the in vitro intermediate filaments studied here are in an environment that is relatively dilute compared to the inside of a cell. The crowded environment within the *Drosophila* embryo would likely favor the formation of the dynamically ordered state over a disordered or denatured one. In support of this concept, molecular crowding agents favor the formation of condensed structures in vitro for both LC domain proteins and amyloid forming proteins (35, 36). Therefore, we consider it reasonable to conclude that the dynamically ordered structure observed here would be present in the physiological environment of a living fly egg or embryo.

We close by emphasizing the fact that no more than half of the amino acids of the Tm1-I/C tail domain achieve any state of molecular order whether studied in either tail domain-only polymers or assembled intermediate filaments. As shown in Fig. 3*C*, precisely the same 30+ amino acid residues exist in a state of molecular disorder in both tail domain-only polymers and segmentally labeled intermediate filaments. That these disordered residues are likely to be of functional importance is supported by studies of systematically deleted variants of the Tm1-I/C polypeptide. Deletion of only 20 residues from the

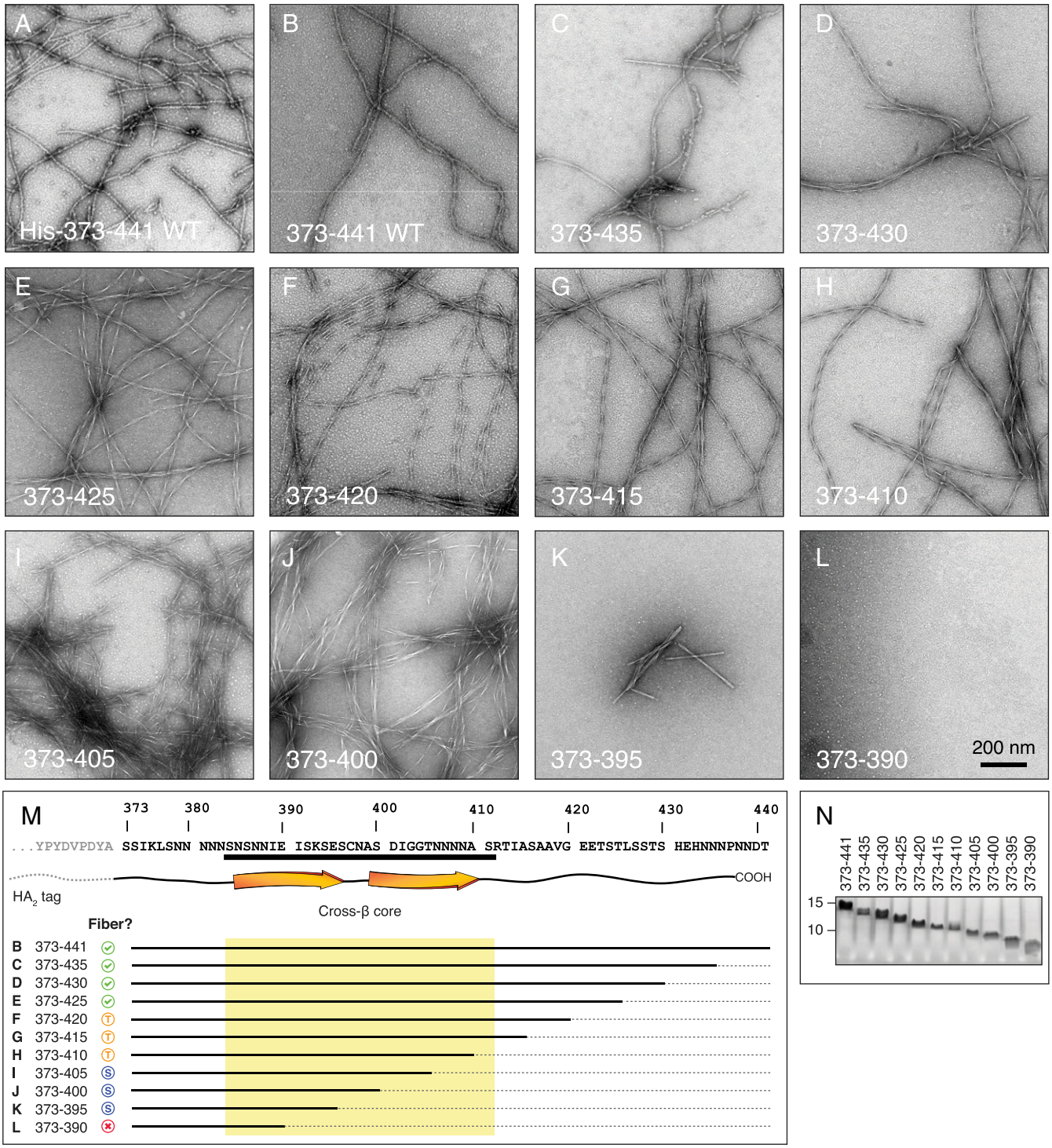

Fig. 4. Functional assays of Tm1-I/C tail domain-only polymer formation for 10 C-terminal deletion mutants. Tm1-I/C tail domain variants missing 6, 11, 16, 21, 26, 31, 36, 41, 46, or 51 residues from the C terminus were expressed in bacterial cells, purified, and incubated under conditions suitable for tail domain polymerization. Samples were negatively stained with uranyl acetate and imaged by electron microscopy (*Materials and Methods*). (*A* and *B*) Intact tail domain polymers bearing an N-terminal His tag (*A*) or an N-terminal HA_2 tag (*B*). (*C–E*) Variants missing up to 16 amino acids formed polymers similar to the intact HA_2-tagged Tm1-I/C tail domain. (*F–H*) Variants missing between 21 and 31 amino acids formed distinctly twisted (T) polymers. (*I–K*) Variants missing between 36 and 46 amino acids formed thin, spiky (S) polymers. (*L*) The variant missing 51 C-terminal amino acids did not form polymers. All electron micrographs were taken at the same magnification (Scale bar in *L*: 200 nm.) (*M*) Tm1-I/C tail deletion mutants relative to the amino acid sequence of the tail domain and the location of the cross-β core determined in this work (orange arrows and yellow highlight). (*N*) Coomassie brilliant blue-stained SDS polyacrylamide gel used to visualize the purified Tm1-I/C tail domain and 10 C-terminal deletion mutants. The numbers to the left of the SDS/PAGE gel indicate the migration positions of molecular weight markers at 15 and 10 kDa.

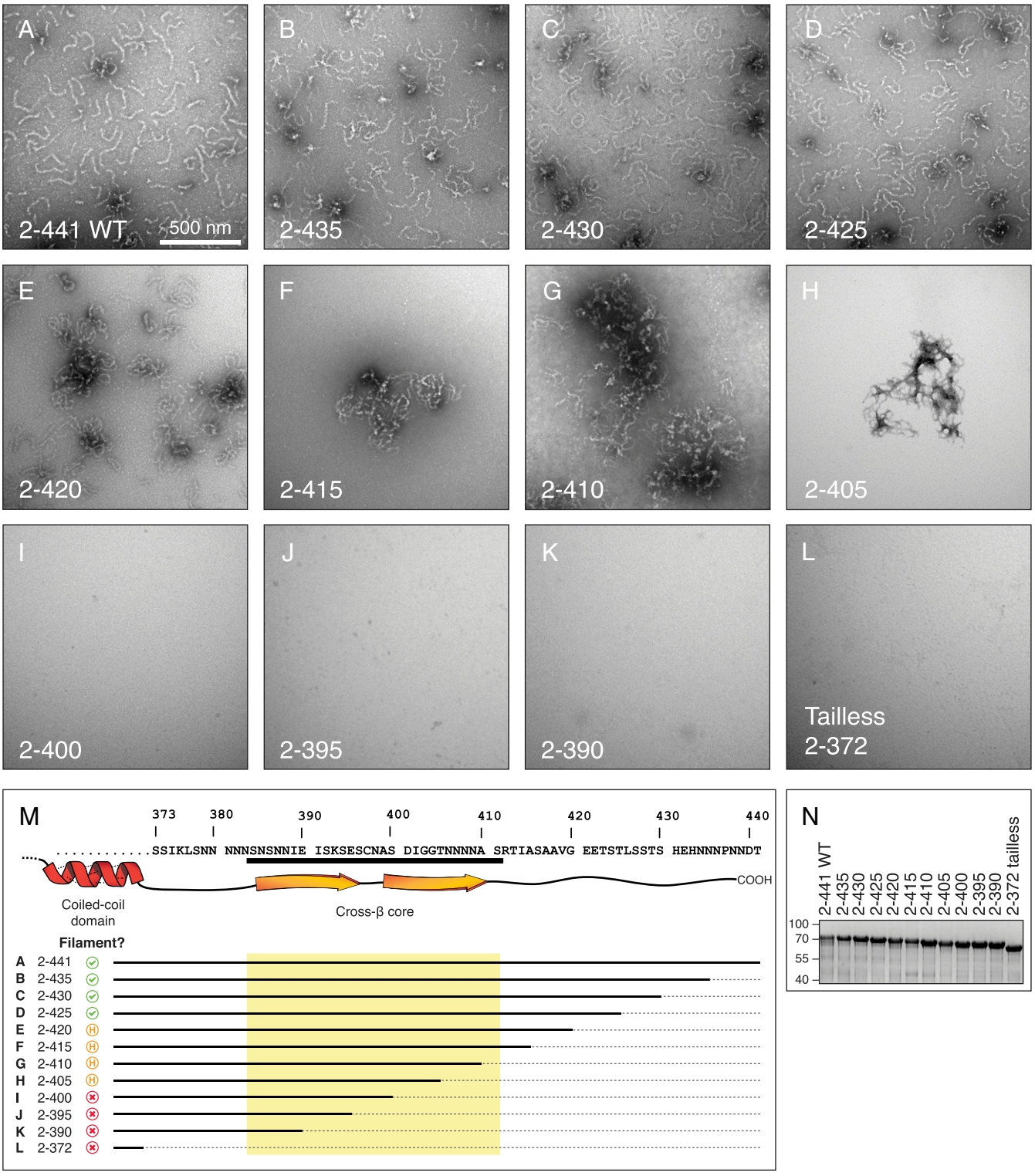

Fig. 5. Functional assays of the ability of Tm1-I/C C-terminal deletion mutants to assemble into intermediate filaments. Deletion variants of the otherwise full-length protein missing 6, 11, 16, 21, 26, 31, 36, 41, 46, 51, or 69 C-terminal residues were expressed in bacterial cells, purified, and incubated under conditions allowing for the formation of intermediate filaments (*Materials and Methods*). (*A–D*) Tm1-I/C variants missing up to 16 C-terminal residues of the protein formed intermediate filaments indistinguishable from those made from the intact protein. (*E–H*) Variants missing between 21 and 36 C-terminal residues yielded aberrantly assembled filaments of a tangled or hairball-like morphology. (*I–L*) Variants of the Tm1-I/C protein missing 41 or more C-terminal residues failed to form intermediate filaments. (*M*) Tm1-I/C tail domain deletion mutants relative to the amino acid sequence of the tail domain, the location of the C-terminal end of the coiled-coil domain (red), and the location of the cross-β core determined in this work (orange arrows and yellow highlight). (*N*) Coomassie brilliant blue-stained SDS polyacrylamide gel used to visualize purified samples of both intact (WT) Tm1-I/C and 11 C-terminal deletion mutants. The numbers to the left of the SDS/PAGE gel indicate the migration position of molecular weight markers at 100, 70, 55, and 40 kDa.

C terminus of the Tm1-I/C protein, although not impinging whatsoever on the structurally ordered region of the tail domain, significantly impedes the assembly of intermediate filaments and yields tail domain-only polymers with a distinctly twisted morphology (Figs. 4 and 5). That LC domains use no more than a modest fraction of their amino acid sequences to achieve transient structural order has also been reported for both the hnRNPA2 and FUS RNA-binding proteins (10, 26). Why LC domains are reliant on regions specifying both structural order and disorder remains an intriguing mystery.

Materials and Methods

All proteins were expressed in *Escherichia coli* and chromatographically purified. Proteins used for NMR studies were expressed in the presence of ^{13}C and ^{15}N isotopes. Segmentally labeled full-length Tm1-I/C was synthesized from an unlabeled N-terminal fragment and an isotope-labeled tail domain fragment in an intein reaction. Biological assemblies — cross-β polymers and intermediate filaments — were assembled in vitro and visualized by electron microscopy. Structures of isotope-labeled proteins were probed by ss-NMR spectroscopy. The study methodology is described in detail in *SI Appendix, Materials and Methods*.

Data Availability. All study data are included in the main text and *SI Appendix*.

ACKNOWLEDGMENTS. We thank Robert Tycko for thoughtful discussions regarding all aspects of the research described herein. We also thank Ivan Hung for assistance with the NMR spectrometers at the Tallahassee, Florida, facility, and staff of the Electron Microscopy Core Facility at UT Southwestern. S.L.M. was supported by National Institute of General Medical Sciences Grant 5R35GM13130358, as well as by unrestricted funding provided by an anonymous donor. D.T.M. was supported by faculty startup funding provided by the University of California, Davis. A portion of this work was performed at the National High-Magnetic Field Laboratory, which is supported by NSF Cooperative Agreement DMR-1644779 and the State of Florida.

1. T. Trcek, R. Lehmann, Germ granules in Drosophila. *Traffic* **20**, 650–660 (2019).
2. A. Ephrussi, R. Lehmann, Induction of germ cell formation by oskar. *Nature* **358**, 387–392 (1992).
3. E. R. Gavis, R. Lehmann, Translational regulation of nanos by RNA localization. *Nature* **369**, 315–318 (1994).
4. A. Bardsley, K. McDonald, R. E. Boswell, Distribution of tudor protein in the Drosophila embryo suggests separation of functions based on site of localization. *Development* **119**, 207–219 (1993).
5. B. Hay, L. Y. Jan, Y. N. Jan, Localization of vasa, a component of Drosophila polar granules, in maternal-effect mutants that alter embryonic anteroposterior polarity. *Development* **109**, 425–433 (1990).
6. M. Erdélyi, A. M. Michon, A. Guichet, J. B. Glotzer, A. Ephrussi, Requirement for Drosophila cytoplasmic tropomyosin in oskar mRNA localization. *Nature* **377**, 524–527 (1995).
7. A. Cho, M. Kato, T. Whitwam, J. H. Kim, D. J. Montell, An atypical tropomyosin in Drosophila with intermediate filament-like properties. *Cell Rep.* **16**, 928–938 (2016).
8. I. Gáspár, V. Sysoev, A. Komissarov, A. Ephrussi, An RNA-binding atypical tropomyosin recruits kinesin-1 dynamically to oskar mRNPs. *EMBO J.* **36**, 319–333 (2017).
9. M. Kato *et al*., Cell-free formation of RNA granules: Low-complexity sequence domains form dynamic fibers within hydrogels. *Cell* **149**, 753–767 (2012).
10. S. Xiang *et al*., The LC domain of hnRNPA2 adopts similar conformations in hydrogel polymers, liquid-like droplets, and nuclei. *Cell* **163**, 829–839 (2015).
11. C. Ader *et al*., Amyloid-like interactions within nucleoporin FG hydrogels. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 6281–6285 (2010).
12. S. Frey, R. P. Richter, D. Görlich, FG-rich repeats of nuclear pore proteins form a three-dimensional meshwork with hydrogel-like properties. *Science* **314**, 815–817 (2006).
13. S. Frey, D. Görlich, A saturated FG-repeat hydrogel can reproduce the permeability properties of nuclear pore complexes. *Cell* **130**, 512–523 (2007).
14. K. Y. Shi *et al*., Toxic PR_n poly-dipeptides encoded by the *C9orf72* repeat expansion block nuclear import and export. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E1111–E1117 (2017).
15. Y. Lin *et al*., Toxic PR poly-dipeptides encoded by the C9orf72 repeat expansion target LC domain polymers. *Cell* **167**, 789–802.e12 (2016).
16. A. Pines, M. G. Gibby, J. S. Waugh, Proton-enhanced NMR of dilute spins in solids. *J. Chem. Phys.* **59**, 569–590 (1973).
17. A. E. Bennett, C. M. Rienstra, M. Auger, K. V. Lakshmi, R. G. Griffin, Heteronuclear decoupling in rotating solids. *J. Chem. Phys.* **103**, 6951–6958 (1995).
18. G. A. Morris, R. Freeman, Enhancement of nuclear magnetic resonance signals by polarization transfer. *J. Am. Chem. Soc.* **101**, 760–762 (1979).
19. A. J. Stevens *et al*., A promiscuous split intein with expanded protein engineering applications. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 8538–8543 (2017).
20. K. D. Kloepper *et al*., Temperature-dependent sensitivity enhancement of solid-state NMR spectra of α-synuclein fibrils. *J. Biomol. NMR* **39**, 197–211 (2007).
21. M. D. Tuttle *et al*., Solid-state NMR structure of a pathogenic fibril of full-length human α-synuclein. *Nat. Struct. Mol. Biol.* **23**, 409–415 (2016).
22. K. Takegoshi, S. Nakamura, T. Terao, 13C–1H dipolar-assisted rotational resonance in magic-angle spinning NMR. *Chem. Phys. Lett.* **344**, 631–637 (2001).
23. K. Takegoshi, S. Nakamura, T. Terao, 13C–1H dipolar-driven 13C–13C recoupling without 13C rf irradiation in nuclear magnetic resonance of rotating solids. *J. Chem. Phys.* **118**, 2325–2341 (2003).
24. J. Pauli, M. Baldus, B. van Rossum, H. de Groot, H. Oschkinat, Backbone and side-chain 13C and 15N signal assignments of the α-spectrin SH3 domain by magic angle spinning solid-state NMR at 17.6 Tesla. *ChemBioChem* **2**, 272–281 (2001).
25. K.-N. Hu, W. Qiang, R. Tycko, A general Monte Carlo/simulated annealing algorithm for resonance assignment in NMR of uniformly labeled biopolymers. *J. Biomol. NMR* **50**, 267–276 (2011).
26. D. T. Murray *et al*., Structure of FUS protein fibrils and its relevance to self-assembly and phase separation of low-complexity domains. *Cell* **171**, 615–627.e16 (2017).
27. D. T. Murray *et al*., Structural characterization of the D290V mutation site in hnRNPA2 low-complexity-domain polymers. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E9782–E9791 (2018).
28. C. P. Jaroniec, C. Filip, R. G. Griffin, 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 (2002).
29. D. S. Wishart, Interpreting protein chemical shift data. *Prog. Nucl. Magn. Reson. Spectrosc.* **58**, 62–87 (2011).
30. E. L. Ulrich *et al*., BioMagResBank. *Nucleic Acids Res.* **36**, D402–D408 (2008).
31. Y. Shen, A. Bax, Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J. Biomol. NMR* **56**, 227–241 (2013).
32. D. Yang, Y.-K. Mok, J. D. Forman-Kay, N. A. Farrow, L. E. Kay, Contributions to protein entropy and heat capacity from bond vector motions measured by NMR spin relaxation. *J. Mol. Biol.* **272**, 790–804 (1997).
33. J. Evenäs, A. Malmendal, M. Akke, Dynamics of the transition between open and closed conformations in a calmodulin C-terminal domain mutant. *Structure* **9**, 185–195 (2001).
34. N. H. Pawley, S. Koide, L. K. Nicholson, Backbone dynamics and thermodynamics of Borrelia outer surface protein A. *J. Mol. Biol.* **324**, 991–1002 (2002).
35. L. A. Munishkina, E. M. Cooper, V. N. Uversky, A. L. Fink, The effect of macromolecular crowding on protein aggregation and amyloid fibril formation. *J. Mol. Recognit.* **17**, 456–464 (2004).
36. A. Molliex *et al*., Phase separation by low-complexity domains promotes stress granule assembly and drives pathological fibrillization. *Cell* **163**, 123–133 (2015).