Structural and hydrodynamic properties of an intrinsically disordered region of a germ cell-specific protein on phase separation

Jacob P. Brady[a,b,c,1], Patrick J. Farber[d,1], Ashok Sekhar[a,b,c], Yi-Hsuan Lin[c,d], Rui Huang[a,b,c], Alaji Bah[d], Timothy J. Nott[e], Hue Sun Chan[a,c], Andrew J. Baldwin[e], Julie D. Forman-Kay[c,d,2], and Lewis E. Kay[a,b,c,d,2]

[a]Department of Molecular Genetics, University of Toronto, Toronto, ON, Canada M5S 1A8; [b]Department of Chemistry, University of Toronto, Toronto, ON, Canada M5S 1A8; [c]Department of Biochemistry, University of Toronto, Toronto, ON, Canada M5S 1A8; [d]Division of Molecular Medicine, The Hospital for Sick Children Research Institute, Toronto, ON, Canada M5G 0A4; and [e]Physical and Theoretical Chemistry Laboratory, University of Oxford, Oxford OX1 3QZ, United Kingdom

Edited by Michael F. Summers, Howard Hughes Medical Institute, University of Maryland, Baltimore County, Baltimore, MD, and approved July 31, 2017 (received for review April 13, 2017)

Membrane encapsulation is frequently used by the cell to sequester biomolecules and compartmentalize their function. Cells also concentrate molecules into phase-separated protein or protein/nucleic acid "membraneless organelles" that regulate a host of biochemical processes. Here, we use solution NMR spectroscopy to study phase-separated droplets formed from the intrinsically disordered N-terminal 236 residues of the germ-granule protein Ddx4. We show that the protein within the concentrated phase of phase-separated Ddx4, $Ddx4_{cond}$, diffuses as a particle of 600-nm hydrodynamic radius dissolved in water. However, NMR spectra reveal sharp resonances with chemical shifts showing $Ddx4_{cond}$ to be intrinsically disordered. Spin relaxation measurements indicate that the backbone amides of $Ddx4_{cond}$ have significant mobility, explaining why high-resolution spectra are observed, but motion is reduced compared with an equivalently concentrated nonphase-separating control. Observation of a network of interchain interactions, as established by NOE spectroscopy, shows the importance of Phe and Arg interactions in driving the phase separation of Ddx4, while the salt dependence of both low- and high-concentration regions of phase diagrams establishes an important role for electrostatic interactions. The diffusion of a series of small probes and the compact but disordered 4E binding protein 2 (4E-BP2) protein in $Ddx4_{cond}$ are explained by an excluded volume effect, similar to that found for globular protein solvents. No changes in structural propensities of 4E-BP2 dissolved in $Ddx4_{cond}$ are observed, while changes to DNA and RNA molecules have been reported, highlighting the diverse roles that proteinaceous solvents play in dictating the properties of dissolved solutes.

phase separation | Ddx4 | NMR | protein dynamics | membraneless organelles

Cellular organization is most often thought to involve membrane-delimited organelles, such as the nucleus and mitochondria. Recently, many cellular features that play critical roles in biology have been revealed to be dense proteinaceous bodies, termed membraneless organelles (1–4). These include nucleoli, speckles, and Cajal bodies in the nucleus and stress, germ, and neuronal granules in the cytoplasm (5–9) as well as signaling puncta found beneath activated plasma membrane receptors (10–12). These cellular compartments often contain nucleic acids and are involved in a variety of biochemical functions, including RNA processing and signaling (11, 13, 14).

The interactions driving both the phase separation and the disassembly of these proteinaceous bodies as well as the physical properties underlying their function are currently the focus of intense investigation (15). Their assembly and disassembly are dependent on dynamic changes in the cell and are sensitive to the cell cycle as well as to environmental stressors (15–17). Formation of these bodies is driven by multivalent weak interactions, including by multiple modular binding domain motif interactions or by self-association of intrinsically disordered proteins (IDPs) (11, 15, 16, 18–26). Rosen and coworkers (10, 12) have shown that repeats of disordered proline-rich motifs in NCK bind to multiple SH3 domains of N-WASP, leading to liquid droplet formation. In this case, multivalency of the binding partners is essential; increasing the number of both SH3 domains and proline-rich motifs reduces the protein concentration required for phase separation. Similarly, the nucleolar protein nucleophosmin interacts in a multivalent manner with arginine-rich motifs, promoting phase separation (27). IDP regions of EWS, FUS, and TAF-15 containing repetitive [G/S]Y[G/S] motifs have been shown to self-associate to form liquid droplets or hydrogels in vitro (23, 24). Additionally, the disordered P granule protein Laf-1, which is highly enriched in RGG motifs, undergoes phase separation into dense liquid droplets (18), and phase separation of low-complexity regions of hnRNPA1 and TDP-43 underlies the formation of stress granules (21).

Many of the cell's proteinaceous bodies have liquid-like properties, including germ granules or nuage, P granules, the nucleolus of *Caenorhabditis elegans*, and stress granules (9, 18, 21).

Significance

The cell is divided into compartments where specific biochemical functions are performed. These compartments can be delineated by membranes or through phase separation of proteins or protein and nucleic acids to form membraneless organelles. The latter situation occurs with an intrinsically disordered region of Ddx4, a major constituent of germ granules. The nature of the interior of membraneless organelles is poorly understood. Here, we use NMR to show that the intrinsically disordered Ddx4 region remains disordered and highly dynamic in the phase-separated state, while diffusing as slowly as a particle the size of a bacterial cell. Ddx4 molecules form a network of interactions on phase separation, providing an alternative environment to that found in membrane-encapsulated organelles.

Author contributions: J.P.B., P.J.F., A.S., J.D.F.-K., and L.E.K. designed research; J.P.B., P.J.F., R.H., T.J.N., and L.E.K. performed research; J.P.B., P.J.F., A.S., Y.-H.L., R.H., A.B., T.J.N., H.S.C., A.J.B., J.D.F.-K., and L.E.K. analyzed data; and J.P.B., P.J.F., H.S.C., J.D.F.-K., and L.E.K. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]J.P.B. and P.J.F. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: forman@sickkids.ca or kay@pound.med.utoronto.ca.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1706197114/-/DCSupplemental.

RNA can promote the liquid-phase separation of RNP granule proteins, including Whi3 in yeast and the RNA binding protein hnRNPA1, tuning the viscoelastic properties of liquid droplets (20, 28). However, there remains much to be learned about the physical properties of the proteins that comprise diverse phase-separated compartments, in particular with regard to their dynamics and the resulting solution hydrodynamics, as well as the atomic-level interactions that drive the phase transition. Additionally, little is known about solute molecules encapsulated within phase-separated protein states, including small molecules, proteins, and nucleic acids.

We recently observed the liquid-phase separation of the germ-granule protein Ddx4 into membraneless organelles in cells and into dense protein bodies in vitro (15). Ddx4 is a germ cell-specific protein that is the major constituent of the nuage/chromatoid body, a membraneless organelle present in the cytoplasm of spermatocytes and spermatids (29). Ddx4 contains an intrinsically disordered N terminus of ∼250 residues, a DEAD-box RNA helicase, and a disordered C terminus of ∼50 residues. A construct of Ddx4 retaining the disordered regions but substituting the DEAD-box helicase domain for YFP forms liquid membraneless organelles in nuclei when expressed in HeLa cells (15). The dense micrometer-sized spherical bodies in the nucleus are morphologically similar to liquid droplets formed in vitro from the N-terminal disordered region of the protein, Ddx4(1–236). Fluorescence recovery after photobleaching (FRAP) data from Ddx4 bodies in cells and in vitro are very similar, consistent with a 100-fold reduced mobility inside the droplet but retention of liquid dynamics. Phase separation of Ddx4(1–236), referred to as Ddx4 below, can be induced by lowering the temperature or by reducing the salt concentration and is highly reversible. Charge patterning with conserved F residues within positively charged blocks and conserved RG/GR and FG/GF motifs are important for formation of Ddx4 organelles and for in vitro and in-cell phase separation (15). These observations provide considerable insight as have analyses of the dilute concentration region of phase diagrams using Flory–Huggins (15) and random-phase approximation (RPA) (30) theories to obtain measures of entropy and enthalpy contributions to phase separation and to rationalize the dependence of phase separation on the charge pattern along the Ddx4 sequence. Nonetheless, a detailed understanding of Ddx4 phase separation in terms of the specific atomic-level interactions involved as well as the structural and dynamic features of Ddx4 and its solution hydrodynamic properties in the phase-separated state remain to be elucidated. Information on such properties is essential, however, for deciphering the biological functions of the membraneless organelles underpinned by Ddx4 phase separation.

Here, we take several substantive steps toward characterizing the molecular properties of the phase-separated state of Ddx4 by first mapping protein concentrations in both the condensed, $Ddx4_{cond}$, and dilute, $Ddx4_{dil}$, phases as a function of temperature and added salt. These data are used to construct full coexistence curves for Ddx4 by fitting to a simple Flory–Huggins model, wherein the salt dependence informs on the role of electrostatic interactions in driving phase separation. Solution NMR spectroscopy is then used to explore structural, hydrodynamic, and motional features of the Ddx4 molecules in $Ddx4_{cond}$. ^{1}H-^{15}N and ^{1}H-^{13}C correlation spectra establish that $Ddx4_{cond}$ is intrinsically disordered, with extensive mobility at both backbone and sidechain sites; however, nuclear Overhauser experiments measuring intermolecular contacts show an extensive network of interactions. Moreover, Ddx4 molecules in $Ddx4_{cond}$ diffuse ∼100-fold more slowly than for $Ddx4_{dil}$, and small molecules with hydrodynamic radii varying from 1 to 5 Å, including water, have significantly reduced translational diffusion rates in $Ddx4_{cond}$, all pointing to a highly viscous environment in membraneless organelles. When the intrinsically disordered but highly compact eukaryotic translation initiation factor 4E binding protein 2 (4E-BP2) is incorporated into $Ddx4_{cond}$, there is a significant reduction in translational mobility, as observed for small molecules, with little change in compaction or structural propensities. In contrast, a number of RNA and DNA solutes are significantly affected on dissolution into $Ddx4_{cond}$ (31), suggesting that the concentrated phase-separated milieu can have diverse effects on solutes that play different functional roles.

Results

Constructing a Homogenous NMR Sample of $Ddx4_{cond}$. To directly probe residue-specific interactions and local motions within $Ddx4_{cond}$, we set out to create a homogeneous solution of $Ddx4_{cond}$ for study by solution NMR. As has been shown previously, Ddx4 droplets are liquid-like and therefore, are able to coalesce together (15). We, therefore, generated a homogeneous NMR sample of Ddx4, in which all droplets are coalesced within the NMR tube (*Materials and Methods*). Specifically, $Ddx4_{cond}$ droplets (Fig. 1*A*) were coalesced at 4 °C to a volume of ∼250–500 μL, transferred to an NMR tube, and appeared turbid. Samples were then incubated at 30 °C for several hours until transparent, as in Fig. 1*B*, *Left*. NMR samples of

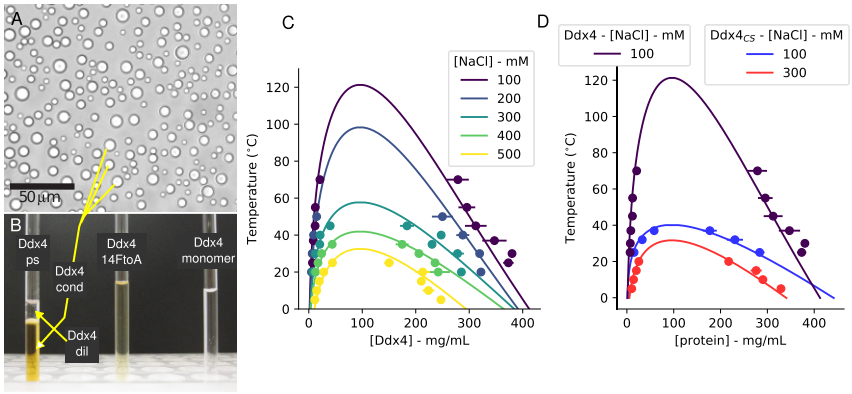

Fig. 1. A phase-separated Ddx4 sample for NMR analysis. (*A*) Bright-field microscopy image of $Ddx4_{cond}$ droplets (200 μM protein, 200 mM NaCl) at 25 °C. (*B*) NMR samples of (*Left*) phase-separated Ddx4, $Ddx4_{ps}$; (*Center*) $Ddx4_{14FtoA}$, 250 mg/mL; and (*Right*) $Ddx4_{monomer}$, 2.5 mg/mL. Note the separation of phases in $Ddx4_{ps}$, with the concentrated phase at the bottom (380 mg/mL) studied here. $Ddx4_{ps}$ and $Ddx4_{14FtoA}$ are dissolved in 20 mM NaPi, buffer, and 100 mM NaCl, while the buffer for $Ddx4_{monomer}$ contains 20 mM NaPi and 400 mM NaCl. (*C*) Phase diagrams for Ddx4 at NaCl concentrations between 100 and 500 mM. Fits to Flory–Huggins model (binodal) are shown with solid lines (*SI Materials and Methods*). Liquid–liquid demixing occurs at Ddx4 concentrations within the regions beneath the curves. (*D*) Phase diagrams for $Ddx4_{CS}$, 100 or 300 mM NaCl and Ddx4, 100 mM NaCl.

phase-separated Ddx4 can be visually identified as having two discrete phases (Fig. 1*B*, *Left*) and are distinct from samples of either the mutant $Ddx4_{14FtoA}$ (see below) at 250 mg/mL (Fig. 1*B*, *Center*) or dilute concentrations of Ddx4 in high salt (2.5 mg/mL), $Ddx4_{monomer}$ (Fig. 1*B*, *Right*), both of which are not phase-separated. Note that, in Fig. 1*B*, *Left*, the boundary between phase-separated compartments of Ddx4 is readily apparent; the top component, $Ddx4_{dil}$, is dilute in protein concentration (~7 mg/mL; 100 mM NaCl; 30 °C; see below), while the bottom component, $Ddx4_{cond}$, of interest here is a condensed Ddx4 solution (~380 mg/mL; see below). NMR experiments (*SI Materials and Methods*) were performed on Ddx4 samples where the majority of the supernatant containing dilute protein was removed, and care was taken to ensure that only $Ddx4_{cond}$ was within the volume of the NMR coil. Samples could be incubated for extended periods (at least 1 wk) at 30 °C without noticeable changes in appearance or in NMR spectra. We did not observe irreversible maturation of droplets in the case of Ddx4 (32, 33), and old samples (>3 wk), although sometimes gel-like, could be readily dissolved into buffer containing fresh reducing agent [100 mM NaCl, 20 mM sodium phosphate (NaPi), 10–20 mM DTT, pH 6.5]. Critically, negative stain EM and Thioflavin T fluorescence experiments were carried out on dilutions of the condensed phase after prolonged incubations and did not show evidence for aggregation or amyloid fibril formation (*SI Materials and Methods* and Fig. S1).

Generating a Highly Concentrated Monomeric Ddx4 Control. We have previously shown through bioinformatics and mutational analysis that conserved RG and FG motifs as well as charge patterning are important for phase separation of Ddx4 (15). Since a mutant in which nine highly conserved Phe residues were mutated to Ala, $Ddx4_{9FtoA}$, does not phase separate in HeLa cells on overexpression, we intended to use it here as a control sample for comparison with $Ddx4_{cond}$. Samples were concentrated in an attempt to reach the protein concentration of the phase-separated sample (~400 mg/mL) but still remain in a mixed state. $Ddx4_{9FtoA}$ remained a single phase to a concentration of approximately 100 mg/mL in NaPi at pH 6.5 and 100 mM NaCl (30 °C) before phase separation occurred at a protein concentration roughly fourfold more dilute than $Ddx4_{cond.}$ We next created a construct where all 14 Phe were mutated to Ala, $Ddx4_{14FtoA}$. This protein remained a single phase through concentrations in excess of 400 mg/mL in 20 mM NaPi, 100 mM NaCl, and 5 mM Tris(2-carboxyethyl)phosphine (TCEP) at pH 6.5 (30 °C) and was, therefore, chosen as an appropriate control. Unlike Ddx4 that spontaneously concentrates itself by ~50-fold under the conditions used here, as described below, high concentrations of $Ddx4_{14FtoA}$ (~400 mg/mL) at temperatures in excess of 0 °C are only achievable via extensive centrifugation. It is worth noting, however, that $Ddx4_{14FtoA}$ is able to phase separate spontaneously by incubating samples on ice; raising the temperature to 5 °C leads to mixing of the two phases. We have also used a second mutant as a control in a number of studies in which all 24 Arg in Ddx4 are replaced by Lys, such that the resulting $Ddx4_{24RtoK}$ preserves the overall protein charge (−3.4 at pH 6.5). Under identical conditions to $Ddx4_{14FtoA}$, $Ddx4_{24RtoK}$ remained as a single phase up to concentrations in excess of 400 mg/mL. Finally, to prevent WT Ddx4 from phase separating, both high salt (400 mM NaCl) and dilute protein concentration (2.5 mg/mL; 30 °C) were required for the $Ddx4_{monomer}$ state. As described above, it is possible to work with $Ddx4_{dil}$, in which case NMR samples are prepared with only a very small amount of condensed phase at the bottom of the tube that does not contribute to the signal.

A Phase Diagram for Ddx4. For proteins involved in the formation of membraneless organelles, phase diagrams that map the temperature and salt dependence of not only the dilute phase but also, the condensed phase are currently lacking. As far as Ddx4 is concerned, a preliminary attempt was made (15), but attention was focused exclusively on the dilute side of the coexistence curve. Here, Fig. 1 *A*–*C* clearly indicates that Ddx4 can phase separate into condensed and dilute phases, distinct from solutions of $Ddx4_{monomer}$ and $Ddx4_{14FtoA}$. Fig. 1*C* shows coexistence curves for Ddx4 at five salt concentrations ranging from 100 to 500 mM obtained by measuring protein concentrations in both the dilute and condensed phases under equilibrium conditions (*Materials and Methods* and *SI Materials and Methods*). The upper critical solution temperatures (UCSTs), corresponding to the temperatures above which the solution is always mixed regardless of protein concentration as estimated by the fitted curves, are strongly dependent on ionic strength and display an approximately 100 °C decrease from 100 to 500 mM NaCl, highlighting the ability of salt to disrupt the interactions driving the phase separation of Ddx4. Consistent with the simple Flory–Huggins theory used to fit the data, the concentration of protein in the condensed phase reduces with increasing temperature. Our experimental data indicate that addition of salt also reduces the concentration of protein in the condensed phase. Notably, as observed previously (15) and rationalized by RPA theory (30), sequence details, such as the distribution of charged blocks, are important determinants of phase behavior. This feature of sequence-dependent phase separation is reinforced by the coexistence curves in Fig. 1*D* for a charge scrambled version of Ddx4, $Ddx4_{CS}$ (Fig. S2) [denoted previously as $Ddx4^{N1}CS$ (15)], showing significant decreases in UCSTs and a lowering of the protein concentration in the condensed phase relative to the WT protein. This sequence sensitivity is consistent with RPA predictions (*Discussion*).

Translational Diffusion of Ddx4. To better understand how hydrodynamic properties couple with phase separation, we next performed pulsed field gradient NMR experiments to measure the diffusion constants, *D*, of $Ddx4_{cond}$ and $Ddx4_{dil}$ at 30 °C (Fig. 2*A*) using a standard single-quantum longitudinal encode–decode ^{13}C-edited diffusion experiment (34). Because signal

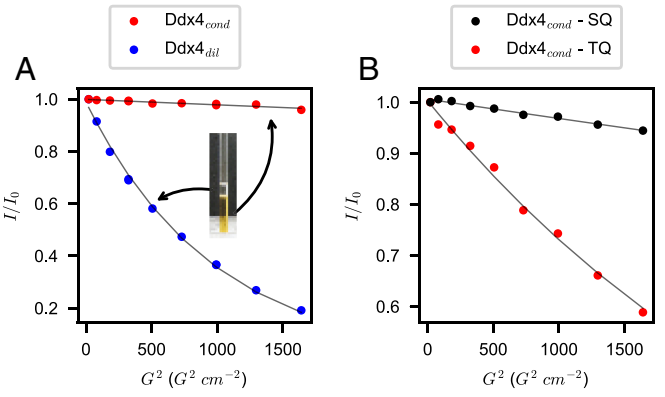

Fig. 2. (*A*) Translational diffusion as measured by 1D single-quantum (SQ) ^{13}C-edited longitudinal encode–decode pulsed field gradient diffusion experiments recorded on $Ddx4_{dil}$ (blue; 7 mg/mL) and $Ddx4_{cond}$ (red; 380 mg/mL) with encoding/decoding gradient times of 4 ms and a diffusion interval of 100 ms. (*B*) Comparative $Ddx4_{cond}$ (380 mg/mL) diffusion profiles obtained from 1D ^{13}C-edited SQ (black) vs. triple-quantum (TQ; red) sequences with 4-ms gradient encoding/decoding durations and a diffusion interval of 400 ms. A decay of ~40% is observed for the TQ sequence, enabling accurate extraction of diffusion constants. In *A* and *B*, *I* is the intensity of integrated 1D spectra measured using encode/decode gradient strengths squared of G^2, while I_0 is the extrapolated value to 0 gradient strength. Experimental points are indicated by circles, and best fitted exponentials are shown in curves.

intensity scales proportionally to $\exp(-\alpha D)$, where α takes into account experimental settings, large intensity losses are associated with faster diffusing molecules. In contrast to the $Ddx4_{dil}$ trace that decays significantly with encode/decode gradient strength squared (G^2), the profile for $Ddx4_{cond}$ (380 mg/mL) shows only a very slight decrease of signal ($\sim 5\%$), making reliable estimation of diffusion constants difficult. To surmount this technical obstacle, we have developed a pulse sequence that selects for methyl triple-quantum ^{1}H magnetization while still retaining close to 75% of the sensitivity of the single-quantum experiment (35), neglecting relaxation effects. Application of encoding/decoding gradients to a triple-quantum state enhances the effective gradient strength by threefold (36, 37), enabling accurate measurement of diffusion constants that are an order of magnitude slower than is possible with traditional single-quantum approaches (Fig. 2*B*). Accordingly, the diffusion coefficient of $Ddx4_{cond}$ was measured to be $7.5 \pm 0.4 \times 10^{-9}$ cm^2/s, which is comparable with values obtained using FRAP on in vitro droplets of Ddx4, where a value of 4×10^{-9} cm^2/s was reported (15). The NMR-based diffusion coefficient of $Ddx4_{dil}$ (7 mg/mL) was determined to be $8.8 \pm 0.2 \times 10^{-7}$ cm^2/s, a factor of 100 times faster than for $Ddx4_{cond}$.

Ddx4 Is Disordered in the Liquid Droplet Form. We have previously shown by solution NMR spectroscopy that $Ddx4_{dil}$ is intrinsically disordered (15). In the case of $Ddx4_{cond}$ (380 mg/mL) (Fig. 3*C*, red), there is extensive line broadening in the ^{1}H-^{15}N heteronuclear single-quantum coherence (HSQC) spectrum, although the amide backbone resonances overlap well with $Ddx4_{dil}$ (Fig. 3*A*, blue and Fig. S3*A*) with only minor chemical shift changes, indicating that Ddx4 molecules in $Ddx4_{cond}$ are also disordered. ^{1}H-^{15}N spectra of $Ddx4_{14FtoA}$ (370 mg/mL) (Fig. 3*B*, green) are very similar to those of $Ddx4_{cond}$ and $Ddx4_{dil}$ (Fig. 3 and Fig. S3*B*), with line broadening that is intermediate between the two. Similar to what was observed in ^{1}H-^{15}N spectra, a comparison of the corresponding ^{13}C datasets also establishes that $Ddx4_{cond}$ is intrinsically disordered, with methyl resonances of Ile, Leu, and Val residues showing little deviation from random coil chemical shift values (Fig. 3, *Right*) (38). Analysis of backbone chemical shift data also revealed little to no propensity for secondary structure formation (Fig. S3*C*). Interestingly, there is no significant line broadening of resonances in ^{1}H-^{13}C spectra of $Ddx4_{cond}$ relative to $Ddx4_{dil}$ and only minor chemical shift changes. Thus, despite the increased effective viscosity of the $Ddx4_{cond}$ solution, the side chains of the protein remain locally highly dynamic, leading to an efficient averaging of magnetic interactions that would otherwise contribute to resonance broadening.

The small diffusion constant of $Ddx4_{cond}$ ($7.5 \pm 0.4 \times 10^{-9}$ cm^2/s; 30 °C) is predicted from the Stokes–Einstein relation (39) for a protein of hydrodynamic radius ~550 nm in dilute aqueous solution. The slow tumbling of such a large molecule would normally preclude recording of NMR spectra. Thus, the observation of backbone amide resonances, albeit broadened, and sharp ^{13}C side-chain resonances indicates that Ddx4 undergoes rapid local motions in the condensed phase, as discussed further below.

Fast Timescale Dynamics of Ddx4. To better understand the reasonably sharp resonances in NMR spectra of $Ddx4_{cond}$, despite the extremely slow translational diffusion, we recorded ^{15}N R_1, R_2 rates and ^{15}N{^{1}H} NOE values (40) of $Ddx4_{dil}$ (~7 mg/mL), $Ddx4_{14FtoA}$ (370 mg/mL), and $Ddx4_{cond}$ (380 mg/mL). The histograms in Fig. 4 *A*–*C* and *E* show R_2 and ^{15}N{^{1}H} NOE values for residues that are sufficiently well-isolated, such that cross-peak intensities could be quantified; in the case of $Ddx4_{cond}$, data for 39 resonances could be obtained of 231 non-Pro residues. Of

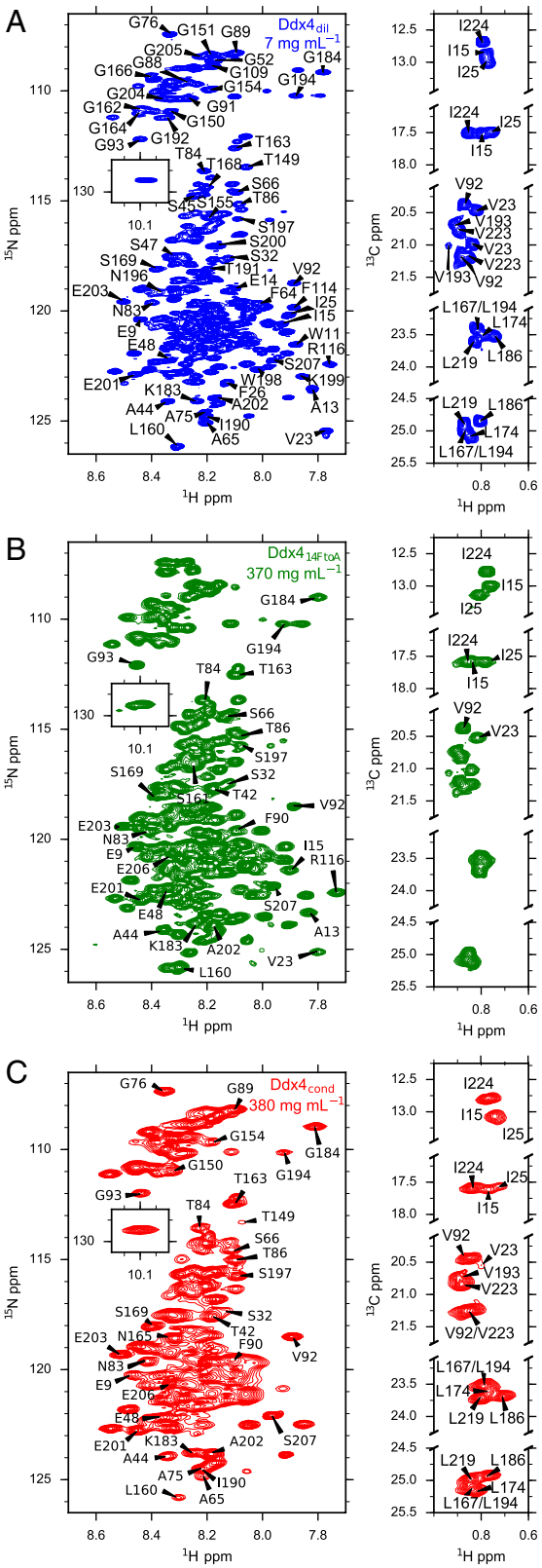

Fig. 3. Ddx4 is intrinsically disordered in the condensed phase-separated state. ^{1}H-^{15}N HSQC spectra (*Left*) and selected methyl region of constant time ^{1}H-^{13}C HSQC spectra (*Right*) of (*A*) $Ddx4_{dil}$ (7 mg/mL), (*B*) $Ddx4_{14FtoA}$ (370 mg/mL), and (*C*) $Ddx4_{cond}$ (380 mg/mL) at 30 °C and 18.8 T. Isoleucine, leucine, and valine (ILV) methyl groups were assigned by mutagenesis in $Ddx4_{dil}$, and assignments were transferred to spectra of $Ddx4_{cond}$ and $Ddx4_{14FtoA}$. Amides were partially (40%) assigned using standard triple-resonance experiments.

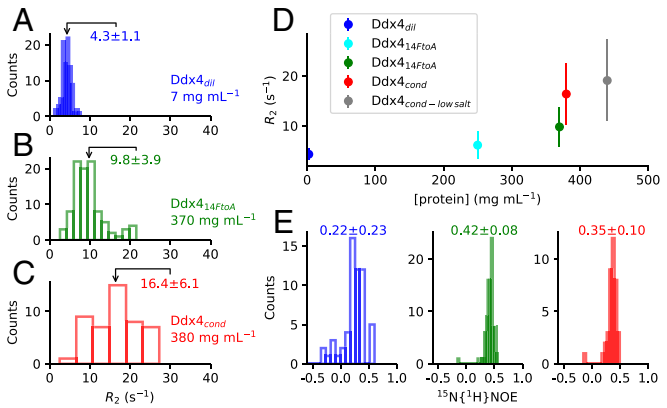

Fig. 4. Extensive internal dynamics in $Ddx4_{cond}$. Histogram plots showing the distribution of R_2 rates for $Ddx4_{dil}$ (*A*), $Ddx4_{14FtoA}$ (*B*), and $Ddx4_{cond}$ (*C*). (*D*) Mean R_2 values and distribution (vertical lines) for $Ddx4_{dil}$, $Ddx4_{cond}$, and $Ddx4_{14FtoA}$ are plotted as a function of protein concentration. (*E*) $^{15}N\{^1H\}$ NOE values of $Ddx4_{dil}$ (blue), $Ddx4_{14FtoA}$ (green), and $Ddx4_{cond}$ (red). Mean and SDs for each dataset are indicated in *A*–*C* and *E*.

these, 23 resonances were sufficiently isolated that they could be unambiguously paired with the corresponding peaks in $Ddx4_{dil}$ and $Ddx4_{14FtoA}$, although a great many more could be correlated in spectra of $Ddx4_{dil}$ and $Ddx4_{14FtoA}$.

The ^{15}N R_2 relaxation rates for $Ddx4_{dil}$ and $Ddx4_{14FtoA}$ measured at a number of concentrations (7, 250, and 370 mg/mL) (Fig. 4*D*) are consistent with values for an IDP, with rates ranging from 1 to 6 s^{-1} and from 2 to 14 s^{-1}, respectively, at 18.8 T and 30 °C. R_2 rates for both $Ddx4_{cond}$ and $Ddx4_{14FtoA}$ become significantly larger at high concentrations, to some extent reflecting increased sample viscosities (Fig. 4*D*). Strikingly, at equivalent concentrations, average R_2 values for $Ddx4_{cond}$ are almost twofold larger than for $Ddx4_{14FtoA}$, with a broader distribution ($16.4 \pm 6.1 s^{-1}$ for $Ddx4_{cond}$, 380 mg/mL, vs. $9.8 \pm 3.9 s^{-1}$ for $Ddx4_{14FtoA}$, 370 mg/mL). Of interest, relaxation dispersion studies (41) showed small (<1 s^{-1}) dispersion profiles for both $Ddx4_{cond}$ and $Ddx4_{14FtoA}$ that were not quantified in this study. Steady-state $^{15}N\{^1H\}$ NOE values, sensitive to local motions on the picosecond to nanosecond timescale, range from −0.7 to 0.6 and do not show a significant difference between samples.

To compare picosecond to nanosecond timescale motions between the different Ddx4 samples, we have recast the R_1, R_2, and $^{15}N\{^1H\}$ NOEs in terms of the product $S^2\tau_C$ (*SI Materials and Methods*). The parameters S^2 and τ_C correspond to the square of the order parameter that describes the amplitude of the amide bond vector motions and the residue-specific tumbling time, respectively (42, 43). Their product provides a measure of the extent of motion at each backbone site in the protein. Values of $S^2\tau_C$ were calculated for the resolved resonances that could be quantified in samples of $Ddx4_{dil}$ (79 residues), $Ddx4_{14FtoA}$ (78 residues), and $Ddx4_{cond}$ (37 residues). On average, $S^2\tau_C$ values for amides in $Ddx4_{cond}$ are 8.5 ± 3 ns, a factor of 6.5 times larger than for $Ddx4_{dil}$ (1.3 ± 0.6 ns) and 1.8 times higher than for $Ddx4_{14FtoA}$. This suggests an additional slowing of backbone motions in $Ddx4_{cond}$, potentially the result of transient intermolecular contacts that mediate phase separation.

Intermolecular Contacts in $Ddx4_{cond}$. Intermolecular contacts between Ddx4 monomers in $Ddx4_{cond}$ were probed by recording F_1 ^{13}C-filtered, F_2 ^{13}C-edited NOESY spectra (44) of samples prepared by mixing 10% ^{13}C-labeled Ddx4 with 90% unlabeled Ddx4. In these spectra, the observed NOEs derive from intermolecular contacts between pairs of protons (Fig. S4). Fig. 5*A* shows strips through 3D datasets extracted at a pair of ^{13}C frequencies, highlighting intermolecular contacts between Phe and Arg residues that are thought to be particularly important in establishing both cation–pi and pi–pi interactions that can play significant roles in stabilization of $Ddx4_{cond}$ (15). Fig. 5*B* tabulates the number of intermolecular NOEs vs. amino acid type. In this regard, it is noteworthy that there are more intermolecular NOEs involving Phe than any other residue in $Ddx4_{cond}$, with a large number for Arg as well. In this analysis, NOEs could not be assigned to specific residues but only to residue types, and quantification was only possible in cases where well-resolved cross-peaks that report on a given residue type could be identified. The data are thus biased by these limitations. Nevertheless, a comparison of NOE spectra recorded for $Ddx4_{cond}$ and $Ddx4_{FtoA}$ (Fig. S4) establishes, for example, that there are a large number of NOEs involving Phe, highlighting the importance of Phe interactions. Interestingly, an extensive network of intermolecular NOEs was also observed for $Ddx4_{14FtoA}$ at an equivalent concentration to $Ddx4_{cond}$ (Fig. S4). However, the contacts in $Ddx4_{14FtoA}$ are not sufficient to drive phase separation. Rather, such contacts for the FtoA mutant are a direct result of the high protein concentration used that is only possible through extensive centrifugation (see above).

Diffusion of Small Molecules and Compact Proteins in $Ddx4_{cond}$. In cells, membraneless organelles are often composed of a mixture of multiple proteins, RNA, and small molecules (1, 3). The motion of these molecules depends on the solvent environment of the organelle and is important for regulating the kinetics and thermodynamics of molecular interactions within the droplet (15) and for controlling chemical properties within the organelle, such as enzymatic activity (17, 45, 46). In addition to providing insight into how binding and reactivity might be affected by the phase-separated environment, quantifying rates of motion of molecules inside the droplet also inform on the intrinsic physical nature of the concentrated phase of the protein that provides the matrix of the proteinaceous organelle.

Pulsed field gradient NMR diffusion measurements were performed to obtain D_o/D values for nine small probe molecules

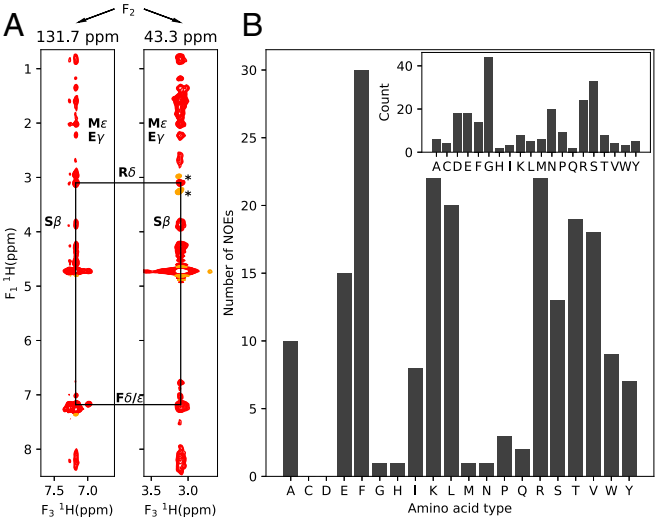

Fig. 5. An extensive network of intermolecular contacts for $Ddx4_{cond}$. (*A*) Selected planes from 3D F_1 ^{13}C-filtered, F_2 ^{13}C-edited NOESY experiments (44) (150-ms mixing times) recorded on $Ddx4_{cond}$ (380 mg/mL) at 30 °C. NOEs between Arg Hδ and Phe Hδ/ϵ side-chain protons are highlighted by black lines between strips. Sample was prepared with 90% unlabeled and 10% ^{15}N,^{13}C-labeled protein. Artifacts from imperfect purging in F_1 are labeled with asterisks. (*B*) Number of NOEs as a function of amino acid type for $Ddx4_{cond}$. (*B*, *Inset*) Number of each amino acid type found in the sequence of Ddx4 (residues 1–236).

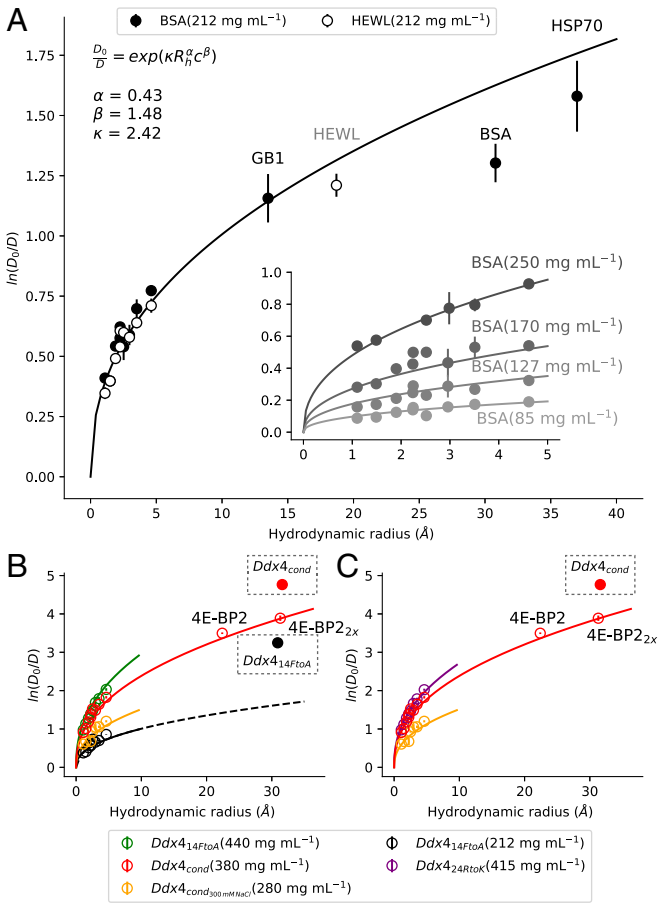

Fig. 6. The diffusion of small molecules and the compact disordered protein 4E-BP2 in $Ddx4_{cond}$ can be explained by an excluded volume model. Logarithm of the diffusion coefficient, *D*, of probe molecules in (*A*) 212 mg/mL HEWL and 212 mg/mL BSA, (*B*) $Ddx4_{cond}$ (280 and 380 mg/mL) and $Ddx4_{14FtoA}$ (440 and 212 mg/mL), and (*C*) $Ddx4_{cond}$ and $Ddx4_{24RtoK}$ (415 mg/mL) relative to diffusion in buffer, D_o, as a function of probe hydrodynamic radius (R_h; units of meters for the Phillies equation) at 30 °C. *A*, *Inset* shows similar plots as a function of BSA concentrations ranging from 85 to 250 mg/mL. Probes used in these experiments and their hydrodynamic radii include H_2O (1.1 Å), methanol (1.49 Å), ethanol (1.9 Å), acetate (2.24 Å), isopropanol (2.25 Å), glycerol (2.5 Å), ribose (2.96 Å), glucose (3.5 Å), sucrose (4.6 Å), 4E-BP2 (22.4 Å; $Ddx4_{cond}$ only), 4E-BP2$_{2x}$ (31.3 Å; $Ddx4_{cond}$ only), B1 domain from Ig binding protein G (13.5 Å; 212 mg/mL BSA only), HEWL (19 Å; 212 mg/mL HEWL only), BSA (31 Å; 212 mg/mL BSA only), and heat shock protein 70 (37 Å; 212 mg/mL BSA only). Diffusion data (circles) were fitted to the Phillies equation (36) (solid lines) as described in the text. Fitting parameters (κ, α, β) obtained from the diffusion series in concentrated samples of BSA and HEWL in *A* are used to generate the solid curves in *B* and *C*. The self-diffusion of $Ddx4_{cond}$ and $Ddx4_{14FtoA}$ (boxed points) is slower than predicted by the Phillies equation with these parameters, consistent with interactions between adjacent molecules.

varying in hydrodynamic radius, R_p, from 1.1 to 4.6 Å and dissolved in $Ddx4_{cond}$ (Fig. 6), where D and D_o are diffusion constants of the probe within a polymer solution of concentration C and in buffer, respectively. D_o/D values were initially quantified in the presence of two folded proteins, either BSA at concentrations of 85–250 mg/mL in steps of ∼50 mg/mL or hen egg white lysozyme (HEWL) at 212 mg/mL (Fig. 6*A* and *SI Materials and Methods*). In addition, diffusion measurements for a number of probe proteins have also been included, as indicated in Fig. 6*A*. The diffusion coefficient ratios increase monotonically with R_p, indicating that the diffusion of larger probes is slowed more than in smaller ones. Note that, in this analysis, the probe hydrodynamic radius is assumed to be invariant to the presence of the added proteins (Ddx4, BSA, HEWL).

While it remains a challenge to understand and predict the diffusion of both small and large molecules within a protein system from first principles, there are empirical models based on organic polymers that can be used to obtain insight (47, 48). Many of these models are based on the obstruction effect, whereby the "solvent" polymer chains, in this case protein molecules, are proposed to retard the motion of the "solvated" probe by reducing the free volume available for it to move (47). Here, we have used one such polymer model based on an equation proposed by Phillies (49, 50), $D_0/D = \exp(\kappa R_h^{\alpha} c^{\beta})$, where κ, α, and β are fitted constants; R_h is the hydrodynamic radius of the probe (meters); and c is the concentration of the polymer (milligrams per milliliter), to quantify the D_o/D vs. R_h profiles that have been obtained using BSA, HEWL, $Ddx4_{14FtoA}$, and $Ddx4_{cond}$ as the proteinaceous solvents. We refer to these proteins as solvents while recognizing that they are, in fact, cosolvents with significant hydration present in all cases. Even at a protein concentration of 380 mg/mL, approximately 73% of the volume of the $Ddx4_{cond}$ droplet is water based on prediction and verified using NMR experiments that quantify water in samples of the concentrated Ddx4 phase (*SI Materials and Methods*).

Fig. 6*A* shows that the Phillies equation is able to fit (black curve) the diffusion data (circles) acquired at all BSA and HEWL concentrations quantitatively with extracted values of κ, α, and β as indicated. The corresponding fits in Fig. 6*B* for probes dissolved in $Ddx4_{14FtoA}$ at 440 (green) and 212 mg/mL (black) or $Ddx4_{cond}$ at 380 (red) and 280 mg/mL (orange) as well as $Ddx4_{24RtoK}$ at 415 mg/mL (magenta in Fig. 6*C*) use the same fitting parameters as optimized for diffusion data recorded with HEWL or BSA polymers. Good fits are obtained despite the fact that Ddx4 is intrinsically disordered, while HEWL and BSA are folded proteins. That the same κ, α, and β parameters are relevant for these four different proteinaceous solvents, with essentially identical D_o/D profiles for a constant polymer concentration, is to be expected (*i*) if the protein obstructs probe motion through an excluded volume effect and (*ii*) if all protein solvents considered have similar densities, so that equivalent masses of protein would occupy similar volumes. In this regard, it is noteworthy that Ddx4 in the nonphase-separated state is highly compact, with an effective hydrodynamic radius of 30 Å (Table S1) that is much closer to what is predicted for a folded protein of the same number of amino acids (23.3 Å) than for an unfolded protein (50.4 Å) (51). Thus, individual molecules in $Ddx4_{cond}$ and $Ddx4_{14FtoA}$ are expected to have similar densities as folded proteins, such as HEWL and BSA. By means of comparison, we have also included $\ln(D_o/D)$ values measured at 30 °C for $Ddx4_{14FtoA}$ and $Ddx4_{cond}$ at concentrations of 370 and 380 mg/mL, respectively (boxes in Fig. 6 *B* and *C*). Unlike the case for the small molecules with diffusion properties that follow the Phillies equation and thus, behave as "hard spheres" in $Ddx4_{cond}$, the self-diffusion of Ddx4 molecules in the phase-separated milieu is slower than predicted, consistent with substantial interactions among Ddx4 molecules in the condensed phase.

Structural Properties of Intrinsically Disordered 4E-BP2 Within $Ddx4_{cond}$. Despite the obvious functional relevance, there is currently very little information on the structural properties of proteins solvated within a membraneless organelle. Proteins, such as GFP and YFP, have been used in photobleaching experiments and as markers for organelles (15, 25), suggesting that they retain photoactive, native-like conformations. Even less is known regarding IDPs solvated within organelles. The eukaryotic translation initiation factor 4E-BP2 is an IDP that plays an important role in the initiation of translation and is found in

neuronal RNA granules, membraneless organelles regulating local mRNA translation in brain function (52, 53). We study it here as a model for IDPs that do not phase separate but are found inside membraneless organelles by characterizing the structural properties of 4E-BP2 solvated within $Ddx4_{cond}$ (380 mg/mL; 30 °C). Assignments for ∼30% of the backbone resonances of 4E-BP2 dissolved in $Ddx4_{cond}$ were obtained, corresponding to ∼90% of the observed 4E-BP2 cross-peaks, based primarily on transfer from spectra recorded on 4E-BP2 dissolved in buffer (54). As illustrated in Fig. 7, the ^{1}H-^{15}N spectrum of 4E-BP2 solvated in $Ddx4_{cond}$ shows some amide shift differences relative to a dataset recorded in buffer (Fig. 7 *A* and *B*), with few changes in ^{13}CO shifts (Fig. 7 *C* and *D*). Taken together, the data are consistent with little change in fluctuating secondary structure of 4E-BP2 when solvated in $Ddx4_{cond}$, although it must be kept in mind that only 30% of the reporters of structure give rise to observable cross-peaks in spectra.

To further explore the effects of $Ddx4_{cond}$ on dissolved solutes, samples were doped with small amounts (∼3%) of ^{15}N, ^{13}C-labeled 4E-BP2. Although the small amount of label compromises the sensitivity of experiments for measuring intermolecular NOEs, diffusion coefficients for a number of 4E-BP2 constructs could be obtained using an HNCO-based longitudinal encode–decode experiment that we developed (Fig. S5). These include 4E-BP2 (120 residues) and a protein made by joining the N and C termini of a pair of 4E-BP2 molecules ($4E\text{-}BP2_{2x}$) (Fig. 6 *B* and *C*). Interestingly, the diffusion data for these proteins could be well-fit with the corresponding data recorded on the small molecules, suggesting that 4E-BP2 diffusion in $Ddx4_{cond}$ can be modeled as motion of hard spheres in the same way as that for the small molecules considered above (Fig. 6). Indeed, although 4E-BP2 is an IDP, it is compact, with an NMR diffusion-based measured hydrodynamic radius of 23 Å that is only slightly larger than that predicted for a folded protein (19 Å) using an empirical relation derived by Smith and coworkers (51). It appears from our measurements that $Ddx4_{cond}$ is no better (or worse) a solvent for 4E-BP2 than water, allowing 4E-BP2 to maintain structural tendencies that are present in buffer.

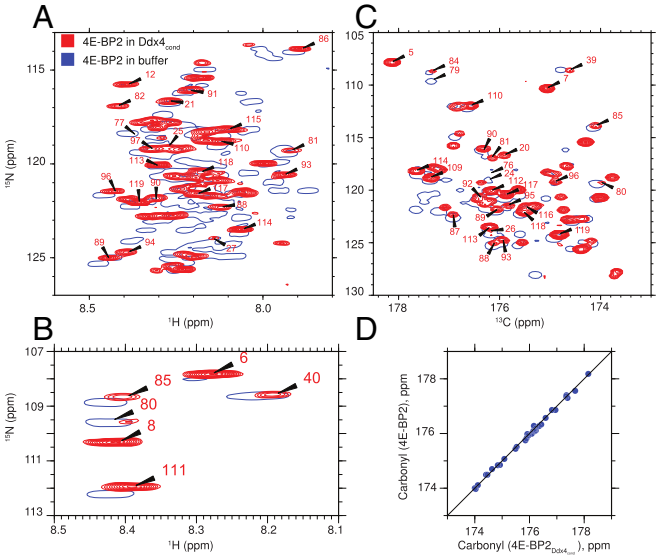

Fig. 7. NMR spectra of the intrinsically disordered 4E-BP2 protein in $Ddx4_{cond}$. (*A* and *B*) Overlay of ^{1}H-^{15}N HSQC spectra of 250 μM ^{15}N, ^{13}C-labeled 4E-BP2 in buffer (blue; single contour) and in $Ddx4_{cond}$ (red; 380 mg/mL) at 30 °C and 14.1 T. (*C*) Overlay of ^{15}N-^{13}CO projections from HNCO datasets recorded on the same samples. (*D*) Linear correlation plot of carbonyl shifts of 4E-BP2 dissolved in buffer (*y* axis) and in $Ddx4_{cond}$ (*x* axis).

Discussion

The matrix of a membraneless organelle is composed of a dense network of proteins and often nucleic acids, creating a unique environment, in which many important biochemical processes can occur. For example, the half-time to completion of the N-WASP–stimulated actin assembly by the Arp2/3 complex is significantly decreased within a nephrin-NCK-N-WASP droplet (11). It has also been postulated that RNA processing is greatly increased within membraneless organelles (17, 45). Therefore, to ascertain the physical basis of the biological function of these organelles, it is of tremendous interest to understand the dynamic and hydrodynamic properties of phase-separated proteins and to quantify how solutes might move within phase-separated droplets.

As shown above, solution NMR spectroscopy is ideally suited for probing the motions of proteins forming phase-separated organelles as well as for monitoring the dynamics of both small molecules and protein solutes within these bodies. Unlike many other physical methods that require attachment of bulky reporters to the molecule of interest, NMR studies can focus on specific molecules that are enriched with NMR active nuclei, such as ^{15}N and ^{13}C, without structural perturbation. Here, we have used a range of different NMR experiments to explore the structure and dynamics of Ddx4 in a phase-separated environment. Rates of translational diffusion of Ddx4 molecules within this condensed phase are reduced by at least 100-fold relative to protein in the coexisting dilute phase (Table S1). Despite this large reduction in translational mobility, high-resolution NMR spectra could nevertheless be recorded. These findings establish that Ddx4 is intrinsically disordered and flexible in the condensed phase-separated state, as also evidenced by limited spectral dispersion in the amide proton region of ^{1}H-^{15}N datasets and in corresponding ^{1}H-^{13}C spectra and by the fact that spectra are very similar to those recorded on the IDP $Ddx4_{dil}$ or $Ddx4_{monomer}$ (15) (Fig. 3 and Fig. S3*A*). The reasonably sharp resonances in the NMR data establish further that both protein backbone and side chains retain considerable local dynamics. Specifically, despite an approximately 100-fold reduction in translational diffusion relative to $Ddx4_{dil}$, an average value of $S^2\tau_C$ of 8.5 ± 3 ns (380 mg/mL; 30 °C) was calculated from a set of backbone ^{15}N spin relaxation data that is only a factor of three larger than $<S^2\tau_C>$ for the 76-residue protein ubiquitin (Table S1). A similar conclusion regarding significant disorder and flexibility in the condensed phase-separated state of the FUS low-complexity region was made by Fawzi and coworkers (25).

Backbone amide ^{15}N R_2 relaxation rates are ∼1.7 times larger for $Ddx4_{cond}$ relative to an equivalently concentrated mixed state, $Ddx4_{14FtoA}$. NOE experiments further establish the presence of intermolecular interactions between neighboring molecules in the phase-separated environment that may contribute to the increased transverse relaxation rates observed for amide probes in $Ddx4_{cond}$. Assuming relatively compact proteins, the concentration ($C*$) at which the transition between a dilute and semidilute solution occurs, the latter corresponding to the regime where protein molecules are in contact with each other, can be estimated from the relation

$$C^* = \frac{3MW}{4\pi N_a R_H^3},$$

where MW and R_H are the protein molecular weight and hydrodynamic radius, respectively, and N_a is Avogadro's number (55). For Ddx4, using $R_H = 31.6$ Å, $C*$ is calculated to be ∼325 mg/mL, similar to the Ddx4 protein concentration in $Ddx4_{cond}$ (380 mg/mL; 100 mM NaCl). $Ddx4_{cond}$ is thus expected to approximate a semidilute protein solution in which Ddx4 molecules partially overlap with one another, leading to transient intermolecular contacts.

Insights into the driving forces for phase separation, summarized in Table S2, can be gained from the salt dependence of the phase diagrams (Fig. 1*C*). As the salt concentration is increased from 100 to 500 mM, the concentration of Ddx4 in the condensed phase decreases from ∼400 to 250 mg/mL, while the protein concentration in the dilute phase increases. When the data are fit to a simple Flory–Huggins model, the UCST decreases by ∼100 °C as the salt concentration is raised from 100 to 500 mM, underscoring the importance of electrostatic interactions to Ddx4 phase separation. In contrast, the mutant $Ddx4_{14FtoA}$ is only able to phase separate at 0 °C in the absence of salt (using a buffer of 5 mM NaPi, 2.5 mM TCEP, pH 6.5), while conditions for phase separation of $Ddx4_{24RtoK}$ have yet to be been found. This points to cation–pi and pi–pi interactions involving Arg–Phe pairs as important determinants in phase separation, consistent with the large numbers of intermolecular contacts involving these amino acid types as quantified in NOE spectra. Note that pi interactions are thought to involve induced quadrupolar electrostatic components that may also be expected to have salt dependence (56). Assuming that the simple Flory–Huggins mixing entropy provides a quantitatively reasonable account of conformational entropy, the Flory–Huggins χ parameter fitted to the data in Fig. 1 suggests that each amino acid in Ddx4 contributes, on average, about 0.3 kcal mol^{-1} effective energy to phase separation through favorable interactions with other Ddx4 residues. The highly approximate nature of this estimate notwithstanding, it is definitively weaker than the effective energy of $\approx$1.5 kcal mol^{-1} per residue that favors folding of globular proteins (57, 58) (*SI Materials and Methods*). This energetic consideration is consistent with the semidilute and dynamic nature of $Ddx4_{cond}$. Note that, while a simple Flory–Huggins theory is used here to model each coexistence curve with two parameters (*Materials and Methods* and *SI Materials and Methods*) and to estimate the average affinity between individual Ddx4 proteins, this fitting procedure does not by itself account physically for how charge blocks, aromatics, and other unique sequence features affect Ddx4 phase separation. Additional physical insights in that regard require more complex theoretical descriptions, such as RPA (30, 59), and explicit-chain simulations (60), the detailed investigation of which is beyond the scope of this work. The situation is best exemplified by the charge scrambled mutant $Ddx4_{CS}$ (Fig. S2) that phase separates poorly, with an UCST that is ∼80 °C lower relative to WT Ddx4 (Fig. 1*D*). Thus, the ordering of charged amino acids is critical for Ddx4 phase separation, as has been noted (15, 30, 59). Quite reassuringly, the general Ddx4 vs. $Ddx4_{CS}$ sequence dependence trend of these phase diagrams spanning both low and high Ddx4 concentrations can also be rationalized by RPA, although theoretical improvements are still needed to provide a quantitatively accurate account of the salt dependence (Fig. S6).

Having shown that $Ddx4_{cond}$ diffuses slowly but has significant amplitude picosecond to nanosecond timescale dynamics, we proceeded to investigate how small molecules and well-folded proteins would diffuse within the Ddx4 phase-separated environment. We found that the diffusion data for small molecules and the compact IDP 4E-BP2 dissolved in $Ddx4_{14FtoA}$ and $Ddx4_{cond}$ could be well-fit to the Phillies model using parameters obtained from the analysis of diffusion data measured in proteinaceous solvents comprising the folded proteins BSA or HEWL. This suggests that the diffusion of small molecules and compact proteins in $Ddx4_{cond}$ is well-approximated by the hard-sphere approximation in the case of minimal interactions between solute and Ddx4 solvent. Notably, ∼75% of the volume of the $Ddx4_{cond}$ (380 mg/mL) droplet is predicted to be water, and we have verified this by recording NMR experiments that quantify water in samples of phase-separated Ddx4. The effective relative permittivity ϵ_r (dielectric constant) of $Ddx4_{cond}$ is thus expected to be significantly lower than that of water. It has been estimated to be more like that of DMSO ($\epsilon_r \approx 47$) (15), and a more recent analytical treatment suggested that ϵ_r may be as low as 20 based on a 75% volume fraction of water (59). All in all, the phase-separated state provides a unique solvent environment that is predicted to enhance electrostatic interactions. This effect is expected to impact on the behavior of charged RNA molecules and many disordered proteins found in proteinaceous organelles.

A large number of distinct membraneless organelles and other cellular bodies (5–7, 9, 52) are formed via interactions between phase-separated proteins of low-complexity disordered regions (15, 16, 18–22, 25) or via dynamic multivalent interactions involving modular binding domains and motifs within disordered regions (10–12, 27). This raises the question of how specific the interactions driving phase separation are and what the rules for enabling or prohibiting specific cophase separation might be. Ddx4 disordered regions form membraneless organelles in the nuclei that do not colocalize with nucleoli, speckles, promyelocytic leukemia (PML), or Cajal bodies (15). In contrast, different cytoplasmic RNA processing bodies can have some similarities in protein composition (52), highlighting the diversity of protein cosolubility in defining cellular compartmentalization. FUS and hnRNPA1 phase-separated protein regions selectively recruit various low-complexity disordered proteins to different extents (20), and a recent study shows that proteins can be included or excluded from phase-separated Ddx4 depending on their charge properties, while structured RNA inside droplets is stabilized and DNA melted (31). Taken together with our finding that $Ddx4_{cond}$ is able to solvate the compact but disordered 4E-BP2 protein with little change to its secondary structural propensities, it is clear that a wide range of interactions is possible between the constituent biomolecules within a phase-separated environment. These results suggest that a detailed description of the interactions underlying phase separation in various cellular bodies will be essential for understanding how the cell organizes itself.

It has been proposed that membraneless organelles are important hubs for reactions to occur (15, 16, 45) and that the high concentration of reactants, such as proteins and nucleic acids, may increase rates of catalysis (45). By contrast, slower translational diffusion would be expected to attenuate reaction rates that are influenced by collision frequencies. In this regard, it is noteworthy that the presence of small molecules and RNA can affect diffusion rates, as evidenced by the fact that the liquidity of Whi3 containing cytoplasmic RNA granules is tuned by RNA (28). Furthermore, formation of phase-separated mRNA processing bodies has been observed in vitro to reduce the catalytic activity of mRNA decapping enzymes (46). Additional experiments are needed to probe the properties of both biomolecules and small molecule solutes in a variety of different phase-separated droplets to better appreciate the unique environment that this milieu provides and how it affects the biochemistry that occurs within it. This study makes it clear that solution NMR spectroscopy is well-poised to contribute to this important goal.

Materials and Methods

Sample Preparation.

Monomeric samples. Genes for Ddx4, $Ddx4_{14FtoA}$, $Ddx4_{24RtoK}$, and $Ddx4_{CS}$ were synthesized and subcloned into pET Sumo vectors by Genscript. Isotopically labeled Ddx4 was overexpressed in *Escherichia coli* and purified as described in detail previously (15). Briefly, cell pellets were resuspended in buffer [50 mM Tris, pH 8.0, 6 M guanidinium (Gdm) HCl, 20 mM imidazole, 500 mM NaCl, 1 mM TCEP] and lysed. Proteins were purified by nickel affinity chromatography (His Trap HP 5-mL column; GE Healthcare Life Sciences) and eluted in buffer containing 50 mM Tris, pH 8.0, 400 mM imidazole, 500 mM NaCl, and 1 mM TCEP. The Sumo tag was removed with Ulp1 protease, and the sample was passed over the Ni column again to remove the cleaved tag after dialyzing out the imidazole. The protein was further purified and buffer-exchanged by size-exclusion chromatography into storage

buffer (20 mM NaPi, pH 6.5, 500 mM NaCl, 5 mM TCEP). 4E-BP2 and 4E-BP2$_{2x}$ were overexpressed in *E. coli* and purified as described previously (54). Ddx4$_{14FtoA}$ and Ddx4$_{24RtoK}$ could be concentrated to ∼250 mg/mL using a 3-kDa molecular mass cutoff (MWCO) Amicon Ultra 4-mL filters at 4,000 × *g* at room temperature. To reach concentrations in excess of 250 mg/mL, samples were centrifuged at 10,000 × *g* at room temperature in Sartorius VivaSpin 500 centrifugal concentrators (3.0-kDa MWCO). Samples were retrieved from 500 μL concentrators by centrifuging the inverted concentrator in a 1.5-mL tube. Sample concentrations were obtained from absorption measurements at 280 nm using an extinction coefficient of 23, 950 $M^{-1}cm^{-1}$ (*Phase Diagram of Ddx4*).

Phase-separated NMR samples. Phase separation was induced by dialysis of a high concentration of Ddx4 in 20 mM NaPi, 1 M NaCl, and 5 mM TCEP, pH 6.5 into a buffer containing 20 mM NaPi, 100 mM NaCl, and 5 mM TCEP, pH 6.5. Phase separation occurs in the dialysis container in the form of small droplets, which then coalesce into a single large drop that can be transferred to either 3-mm tubes (Wilmad) or 5-mm Shigemi tubes to perform NMR measurements. Typically, ∼50 mL of 200 μM Ddx4 was dialyzed against 4 L NMR buffer to produce one NMR sample. Samples were gently centrifuged in the NMR tube and allowed to equilibrate at 30 °C for a minimum of 24 h.

Small molecule diffusion in viscogen. BSA and HEWL samples were prepared by dissolving lyophilized protein powder (purchased from Sigma and BioBasic, respectively) in 20 mM NaPi, 100 mM NaCl, and 90% H_2O/10% D_2O, pH 6.5 (BSA) or 20 mM $Na_3C_6H_5O_7$ (sodium citrate) and 100 mM NaCl, pH 5 (HEWL). The different buffer systems reflect the poor solubility of HEWL in NaPi buffer. Small molecule probes were added at a concentration of 2–10 mM, and experiments were repeated at a pair of concentrations to ensure that there were no interactions between probes and protein viscogens; ^{13}C isotopically labeled small molecules (Fig. 6 shows the list) were generously donated by Cambridge Isotope Laboratories.

Phase Diagram of Ddx4. Phase diagrams or coexistence curves (temperature vs. concentration) were constructed for Ddx4 and Ddx4$_{CS}$ over a range of NaCl concentrations. Initially, the two phases of the demixed sample were dispersed by pipetting up and down until a milky dispersion was formed. Equal volumes (∼200 μL) of sample were then incubated in triplicate at the desired temperature using a PCR thermocycler with a heated lid (90 °C) for at least 1 h before measurement. Subsequently, samples were briefly centrifuged (<15 s at 10,000 × *g*) to ensure full coalescence of phase-separated droplets comprising the condensed phase at the bottom of the tube. For concentration measurements, dilutions were made into 6 M Gdm HCl and 20 mM NaPi, pH 6.5. Measurements of the dilute (top) phase were made on samples of 10 μL removed from each tube and diluted 10-fold. For the condensed (bottom) phase, 2 or 10 μL were diluted 250- to 500-fold. Absorbance at 280 nm was converted to concentration (milligrams per milliliter) using the Beer–Lambert law with a predicted extinction coefficient (ϵ) of 23,950 $M^{-1}cm^{-1}$ (61) using molecular masses of 25,833, 24,768, 25,161, and 25,833 Da for WT Ddx4, 14FtoA, 24RtoK, and charge scrambled Ddx4, respectively. The reported concentrations and errors are means and SDs of triplicate measurements, respectively. The procedure for fitting to the Flory–Huggins model is described in *SI Materials and Methods* and Fig. S7.

Details of the NMR experiments are provided in *SI Materials and Methods*.

ACKNOWLEDGMENTS. We thank Zev Ripstein for help with negative stain EM experiments and Professor Daniel Blair and Pasha Tabatabai (Georgetown University) for valuable discussions. This work was supported by grants from the Canadian Institutes of Health Research (to H.S.C., J.D.F.-K., and L.E.K.), the Natural Sciences and Engineering Research Council of Canada (to J.D.F.-K. and L.E.K.), and the Canadian Cancer Society (to H.S.C. and J.D.F.-K.). J.D.F.-K. holds a Canada Research Chair in Intrinsically Disordered Proteins, and L.E.K. holds a Canada Research Chair in Biochemistry.

1. Toretsky JA, Wright PE (2014) Assemblages: Functional units formed by cellular phase separation. *J Cell Biol* 206:579–588.
2. Hyman AA, Weber CA, Jülicher F (2014) Liquid-liquid phase separation in biology. *Annu Rev Cell Dev Biol* 30:39–58.
3. Mitrea D, Kriwacki R (2016) Phase separation in biology; functional organization of a higher order. *Cell Commun Signal* 14:1.
4. Chong PA, Forman-Kay JD (2016) Liquid-liquid phase separation in cellular signaling systems. *Curr Opin Struct Biol* 41:180–186.
5. Cajal SR, et al. (1903) Un sencillo metodo de coloracion seletiva del reticulo protoplasmatico y sus efectos en los diversos organos nerviosos de vertebrados e invertebrados. *Trab Lab Invest Biol (Madrid)* 2:129–221.
6. Ramón Y, Cajal S (1910) El núcleo de las células piramidales del cerebro humano y de algunos mamíferos. *Trab Lab Invest Biol* 8:27–62.
7. Montgomery TS (1898) Comparative cytological studies, with especial regard to the morphology of the nucleolus. *J Morphol* 15:265–582.
8. Brangwynne CP, Mitchison TJ, Hyman AA (2011) Active liquid-like behavior of nucleoli determines their size and shape in Xenopus laevis oocytes. *Proc Natl Acad Sci USA* 108:4334–4339.
9. Brangwynne CP, et al. (2009) Germline P granules are liquid droplets that localize by controlled dissolution/condensation. *Science* 324:1729–1732.
10. Banjade S, et al. (2015) Conserved interdomain linker promotes phase separation of the multivalent adaptor protein Nck. *Proc Natl Acad Sci USA* 112:E6426–E6435.
11. Li P, et al. (2012) Phase transitions in the assembly of multivalent signalling proteins. *Nature* 483:336–340.
12. Banjade S, Rosen MK (2014) Phase transitions of multivalent proteins can promote clustering of membrane receptors. *Elife*, 10.7554/eLife.04123.
13. Anderson P, Kedersha N (2009) RNA granules: Post-transcriptional and epigenetic modulators of gene expression. *Nat Rev Mol Cell Biol* 10:430–436.
14. Sheth U, Parker R (2003) Decapping and decay of messenger RNA occur in cytoplasmic processing bodies. *Science* 300:805–808.
15. Nott TJ, et al. (2015) Phase transition of a disordered nuage protein generates environmentally responsive membraneless organelles. *Mol Cell* 57:936–947.
16. Weber SC, Brangwynne CP (2015) Inverse size scaling of the nucleolus by a concentration-dependent phase transition. *Curr Biol* 25:641–646.
17. Brangwynne CP (2013) Phase transitions and size scaling of membrane-less organelles. *J Cell Biol* 203:875–881.
18. Elbaum-Garfinkle S, et al. (2015) The disordered P granule protein LAF-1 drives phase separation into droplets with tunable viscosity and dynamics. *Proc Natl Acad Sci USA* 112:7189–7194.
19. Patel A, et al. (2015) A liquid-to-solid phase transition of the ALS protein FUS accelerated by disease mutation. *Cell* 162:1066–1077.
20. Lin Y, Protter DS, Rosen MK, Parker R (2015) Formation and maturation of phase-separated liquid droplets by RNA-binding proteins article formation and maturation of phase-separated liquid droplets by RNA-binding proteins. *Mol Cell* 60:208–219.
21. Molliex A, et al. (2015) Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. *Cell* 163:123–133.
22. Jiang H, et al. (2015) Phase transition of spindle-associated protein regulate spindle apparatus assembly. *Cell* 163:108–122.
23. Kato M, et al. (2012) Cell-free formation of RNA granules: Low complexity sequence domains form dynamic fibers within hydrogels. *Cell* 149:753–767.
24. Han TW, et al. (2012) Cell-free formation of RNA granules: Bound RNAs identify features and components of cellular assemblies. *Cell* 149:768–779.
25. Burke KA, Janke AM, Rhine CL, Fawzi NL (2015) Residue-by-residue view of in vitro FUS granules that bind c-terminal domain RNA polymerase. *Mol Cell* 60: 231–241.
26. Fromm SA, et al. (2014) In vitro reconstitution of a cellular phase-transition process that involves the mRNA decapping machinery. *Angew Chem Int Ed Engl* 53:7354–7359.
27. Mitrea DM, et al. (2016) Nucleophosmin integrates within the nucleolus via multimodal interactions with proteins displaying R-rich linear motifs and rRNA. *Elife* 5:e13571.
28. Zhang H, et al. (2015) RNA controls PolyQ protein phase transitions. *Mol Cell* 60:220–230.
29. Kotaja N, Sassone-Corsi P (2007) The chromatoid body: A germ-cell-specific RNA-processing centre. *Nat Rev Mol Cell Biol* 8:85–90.
30. Lin Y-H, Forman-Kay JD, Chan HS (2016) Sequence-specific polyampholyte phase separation in membraneless organelles. *Phys Rev Lett* 117:178101.
31. Nott TJ, Craggs TD, Baldwin AJ (2016) Membraneless organelles can melt nucleic acid duplexes and act as biomolecular filters. *Nat Chem* 8:569–575.
32. Molliex A, et al. (2015) Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization article phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. *Cell* 163:123–133.
33. Lin Y, Protter DSW, Rosen MK, Parker R (2015) Formation and maturation of phase-separated liquid droplets by RNA-binding proteins. *Mol Cell* 60:208–219.
34. Choy W-Y, et al. (2002) Distribution of molecular size within an unfolded state ensemble using small-angle X-ray scattering and pulse field gradient NMR techniques. *J Mol Biol* 316:101–112.
35. Yuwen T, Vallurupalli P, Kay LE (2016) Enhancing the sensitivity of CPMG relaxation dispersion to conformational exchange processes by multiple-quantum spectroscopy. *Angew Chem Int Ed Engl* 55:11490–11494.
36. Zax D (1983) Study of anisotropic diffusion of oriented molecules by multiple quantum spin echoes. *J Chem Phys* 78:6333–6334.
37. Kay LE, Prestegard JH (1986) An application of pulse-gradient double-quantum spin echoes to diffusion measurements on molecules with scalar-coupled spins. *J Magn Reson* 67:103–113.
38. Wishart DS, Bigam CG, Holm A, Hodges RS, Sykes BD (1995) 1H, 13C and 15N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects. *J Biomol NMR* 5:67–81.
39. Einstein A (1905) Investigations on the theory of the Brownian Movement. *Ann Phys* 17:549–560.
40. Farrow NA, et al. (1994) Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by 15N NMR relaxation. *Biochemistry* 33:5984–6003.
41. Palmer AG 3rd, Kroenke CD, Loria JP (2001) Nuclear magnetic resonance methods for quantifying microsecond-to-millisecond motions in biological macromolecules. *Methods Enzymol* 339:204–238.

42. Lipari G, Szabo A (1982) Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. *J Am Chem Soc* 104:4546–4559.
43. Lipari G, Szabo A (1982) Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 2. Analysis of experimental results. *J Am Chem Soc* 104:4559–4570.
44. Zwahlen C, et al. (1997) Methods for measurement of intermolecular NOEs by multinuclear NMR spectroscopy: Application to a bacteriophage λ N-peptide/boxB RNA complex. *J Am Chem Soc* 119:6711–6721.
45. Weber SC, Brangwynne CP (2012) Getting RNA and protein in phase. *Cell* 149:1188–1191.
46. Schütz S, Nöldeke ER, Sprangers R (2017) A synergistic network of interactions promotes the formation of in vitro processing bodies and protects mRNA against decapping. *Nucleic Acids Res* 45:6911–6922.
47. Amsden B (1998) Solute diffusion within hydrogels. Mechanisms and models. *Macromolecules* 31:8382–8395.
48. Masaro L, Zhu XX (1999) Physical models of diffusion for polymer solutions, gels and solids. *Prog Polymer Sci* 24:731–775.
49. Phillies GDJ (1986) Universal scaling equation for self-diffusion by macromolecules in solution. *Macromolecules* 19:2367–2376.
50. Phillies GDJ (2011) *Phenomenology of Polymer Solution Dynamics* (Cambridge Univ Press, Cambridge, UK).
51. Wilkins DK, Grimshaw SB, Dobson CM, Jones JA, Smith LJ (1999) Hydrodynamic radii of native and denatured proteins measured by pulse field gradient NMR techniques. *Biochemistry* 38:16424–16431.
52. Anderson P, Kedersha N (2006) RNA granules. *J Cell Biol* 172:803–808.
53. Wolozin B (2012) Regulated protein aggregation: Stress granules and neurodegeneration. *Mol Neurodegener* 7:56.
54. Lukhele S, Bah A, Lin H, Sonenberg N, Forman-Kay JD (2013) Interaction of the eukaryotic initiation factor 4E with 4E-BP2 at a dynamic bipartite interface. *Structure* 21:2186–2196.
55. Mohr P, Taylor B, Newell D (2012) CODATA recommended values of the fundamental physical constants: 2010. *Rev Mod Phys* 84:1527–1605.
56. Sherrill CD (2013) Energy component analysis of π interactions. *Acc Chem Res* 46:1020–1028.
57. Baxa MC, Haddadian EJ, Jumper JM, Freed KF, Sosnick TR (2014) Loss of conformational entropy in protein folding calculated using realistic ensembles and its implications for NMR-based calculations. *Proc Natl Acad Sci USA* 111:15396–15401.
58. Privalov PL, Gill SJ (1988) Stability of protein structure and hydrophobic interaction. *Adv Protein Chem* 39:191–234.
59. Lin Y-H, Song J, Forman-Kay JD, Chan HS (2017) Random-phase-approximation theory for sequence-dependent, biologically functional liquid-liquid phase separation of intrinsically disordered proteins. *J Mol Liq* 228:176–193.
60. Ruff KM, Harmon TS, Pappu RV (2015) CAMELOT: A machine learning approach for coarse-grained simulations of aggregation of block-copolymeric protein sequences. *J Chem Phys* 143:243123.
61. Gill SC, von Hippel PH (1989) Calculation of protein extinction coefficients from amino acid sequence data. *Anal Biochem* 182:319–326.
62. Delaglio F, et al. (1995) NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* 6:277–293.
63. Goddard TD, Kneller DG (2006) Sparky—NMR Assignment and Integration software (University of California, San Fransisco).
64. Vranken WF, et al. (2005) The CCPN data model for NMR spectroscopy: Development of a software pipeline. *Proteins* 59:687–696.
65. Altieri AS, Hinton DP, Byrd RA (1995) Association of biomolecular systems via pulsed field gradient NMR self-diffusion measurements. *J Am Chem Soc* 117:7566–7567.
66. Holz M, Weingartner H (1991) Calibration in accurate spin-echo self-diffusion measurements using 1H and less-common nuclei. *J Magn Reson* 92:115–125.
67. Griesinger C, Sattler M (1999) Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog Nucl Magn Reson Spectrosc* 34:93–158.
68. Panchal SC, Bhavesh NS, Hosur RV (2001) Improved 3D triple resonance experiments, HNN and HN(C)N, for HN and 15N sequential correlations in (13C, 15N) labeled proteins: Application to unfolded proteins. *J Biomol NMR* 20:135–147.
69. Abragam A (1961) *Principles of Nuclear Magnetism* (Clarendon, Oxford).
70. Kroenke CD, Rance M, Palmer AG (1999) Variability of the N-15 chemical shift anisotropy in Escherichia coli ribonuclease H in solution. *J Am Chem Soc* 121:10119–10125.
71. Sekhar A, Vallurupalli P, Kay LE (2013) Defining a length scale for millisecond-timescale protein conformational exchange. *Proc Natl Acad Sci USA* 110:11391–11396.
72. Colby RH, Rubinstein M (2003) *Polymer Physics* (Oxford Univ Press, Oxford).
73. Chan HS, Dill KA (1991) Polymer principles in protein structure and stability. *Annu Rev Biophys Biophys Chem* 20:447–490.
74. Patel A, et al. (2015) A liquid-to-solid phase transition of the ALS protein FUS accelerated by disease mutation. *Cell* 162:1066–1077.
75. Das RK, Ruff KM, Pappu RV (2015) Relating sequence encoded information to form and function of intrinsically disordered proteins. *Curr Opin Struct Biol* 32:102–112.
76. Camilloni C, De Simone A, Vranken WF, Vendruscolo M (2012) Determination of secondary structure populations in disordered states of proteins using nuclear magnetic resonance chemical shifts. *Biochemistry* 51:2224–2231.
77. Shaka AJ, Keeler J, Frenkiel T, Freeman R (1983) An improved sequence for broadband decoupling: WALTZ-16. *J Magn Reson* 52:335–338.
78. Kay LE, Ikura M, Tschudin R, Bax A (1990) Three-dimensional triple-resonance NMR spectroscopy of isotopically enriched proteins. *J Magn Reson* 89:496–514.
79. McCoy MA, Mueller L (1992) Selective shaped pulse decoupling in NMR: Homonuclear [13C] carbonyl decoupling. *J Am Chem Soc* 114:2108–2112.
80. Marion D, Ikura M, Tschudin R, Bax AD (1989) Rapid recording of 2D NMR spectra without phase cycling. Application to the study of hydrogen exchange in proteins. *J Magn Reson* 85:393–399.
81. Kay L, Keifer P, Saarinen T (1992) Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *J Am Chem Soc* 114:10663–10665.
82. Schleucher J, Sattler M, Griesinger C (1993) Coherence selection by gradients without signal Attenuation: Application to the three-dimensional HNCO experiment. *Angew Chem Int Ed English* 32:1489–1491.