LIQUID-LIQUID PHASE SEPARATION: MOLECULAR MECHANISMS AND INFLUENCE ON THE mRNA DECAPPING MACHINERY

DISSERTATION ZUR ERLANGUNG DES
DOKTORGRADES DER NATURWISSENSCHAFTEN (DR. RER. NAT.)
DER FAKULTÄT FÜR BIOLOGIE UND VORKLINISCHE MEDIZIN
DER UNIVERSITÄT REGENSBURG

vorgelegt von

STEFAN SCHÜTZ

aus

QUEDLINBURG

im Jahr

2019

Das Promotionsgesuch wurde eingereicht am:

14. Juni 2019

Die Arbeit wurde angeleitet von:

PROF. DR. REMCO SPRANGERS

Unterschrift:

Stefan Schütz

List of Publications and Manuscripts

Published

1. Schütz, S.; Nöldeke, E. R.; Sprangers, R. A Synergistic Network of Interactions Promotes the Formation of in Vitro Processing Bodies and Protects mRNA against Decapping. *Nucleic Acids Res.* **2017**, *45* (11), 6911–6922. https://doi.org/10.1093/nar/gkx353.

2. Damman R; Schütz, S; Luo, Y; Weingarth, M; Sprangers, R; Baldus, M. Atomic-level insight into the maturated state of mRNA processing bodies by combining solid and solution-state NMR spectroscopy. *Nat Commun.* **2019**, 10 (1), 4536. https://doi.org/10.1038/s41467-019-12402-3.

Accepted

3. Schütz, S.; Sprangers, R. Methyl TROSY spectroscopy: A versatile NMR approach to study challenging biological systems. *Prog. Nucl. Magn. Reson. Spectrosc.,* **2019**, in press, https://doi.org/10.1016/j.pnmrs.2019.09.004

In preparation

4. Schütz, S; Sprangers, R. Deciphering the contributions of molecular interactions that lead to liquid-liquid phase separation of the conserved DEAD-box protein Dhh1.

The contributions of the individual authors to the listed publications and manuscripts are indicated at the beginning of the respective chapters.

TABLE OF CONTENT

SUMMARY

Cellular liquid-liquid phase separation (LLPS) results in the formation of dynamic membrane-less granules that play an important role in many biological processes. On a molecular level, the clustering of proteins into a confined space results from an indefinite network of intra- and intermolecular interactions.

Here, we introduce and exploit a novel high-throughput bottom-up approach to study how the interactions between RNA, the Dcp1:Dcp2 mRNA decapping complex and the scaffolding proteins Edc3 and Pdc1 result in LLPS and the formation of processing bodies (P-bodies). We find that the LLPS boundaries are close to physiological concentrations upon inclusion of multiple proteins and RNA. Within *in vitro* P-bodies the RNA is protected against endonucleolytic cleavage and the mRNA decapping activity is reduced, which argues for a role of P-bodies in temporary mRNA storage.

Interestingly, the intrinsically disordered region (IDR) in the Edc3 protein emerges as a central hub for interactions with both mRNA and mRNA decapping factors. In addition, the Edc3 IDR plays a role in the formation of irreversible protein aggregates that are potentially detrimental for cellular homeostasis.

Until now, a detailed structural characterization of the intrinsically heterogeneous LLPS process has been challenging. Here, we combine solid- and solution-state NMR spectroscopy to obtain atomic-level insights into the assembly and maturation of *in vitro* P-bodies. Our results reveal that Edc3 domains exhibit diverse levels of structural organization and dynamics after LLPS. In addition, we find that interactions between the different Edc3 domains and between Edc3 and mRNA in solution are largely preserved in the condensed protein state, allowing P-bodies to rapidly form and dissociate upon small alterations in the cellular environment.

Additionally, we aim at unraveling the role of the conserved helicase Dhh1 in the formation of (*in vitro*) P-bodies. We found that the LLPS process of Dhh1 contains contributions from the RNA, the IDRs at the N- and C-terminal regions and the folded helicase core domains. Based on mutants of the enzyme, we identified residues in the C-terminal part of the second helicase core domain to be crucial for LLPS of Dhh1. In addition, we found that ATP enhances Dhh1 phase separation, even in the absence of RNA. Our results will allow us to conclude to what degree the above interactions contribute in a constructive manner to LLPS and, by employing NMR spectroscopic methods, which residues are involved in the phase separation process.

In summary, our work sheds light on both the molecular mechanisms that underlie liquid-liquid phase separation and provides clues about how this influences cellular processes.

SELECTED ABBREVIATIONS

ADP/ATP	Adenosine di/triphosphate
CSA	Chemical shift anisotropy
CSP	Chemical shift perturbation
CV	Column volume
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
Dcp(S)	(Scavenger) Decapping protein
DDX	DEAD-box protein
Dhh	DEAD-box helicase homolog
Edc	Enhancer of decapping
eIF	eukaryotic Initiation Factor
FUS	Fused in Sarcoma
GDP/GMP	Guanosine di/monophosphate
GST	Glutathion-S-transferase
HEPES	2-(4-(2-Hydroxyethyl)-1-piperazinyl)-ethanesulfonic acid
HLM	Helical leucine-rich motif
HMQC	Heteronuclear multiple quantum coherence
hn	Heterogeneous nuclear
HSQC	Heteronuclear single quantum coherence
IDR	Intrinsically disordered region
IPTG	Isopropyl β-D-1-thiogalactopyranoside
ITC	Isothermal titration calorimetry
LLPS	Liquid-liquid phase separation
LSm	Like Smith
MAGIC	Methyl assignment by graphing inference construct
MBP	Maltose binding protein
MES	2-(*N*-morpholino)-ethanesulfonic acid
(ss)NMR	(solid-state) Nuclear magnetic resonance
NOE(SY)	Nuclear Overhauser effect (spectroscopy)
PAGE	Polyacrylamide gel electrophoresis
PABP	Poly(A) binding protein
Pat	Protein associated with topoisomerase II
P-body	Processing body
RecA	Recombinase A
(m)RNA	(messenger) Ribonucleic acid
RNP	Ribonucleoprotein
Scd	Suppressor of clathrin deficiency
SDS	Sodium dodecylsulfate
SEC	Size exclusion chromatography
SH	Src-homology
SUMO	Small ubiquitin-related modifier
TEV	Tobacco etch virus
TROSY	Transverse relaxation optimized spectroscopy
Xrn	Exoribonuclease

Amino acids are abbreviated by their one- or three-letter code. Nucleobases are abbreviated by their one-letter code.

CHAPTER 1 *General Introduction*

1.1 *A SHORT VIEW ON THE LIFE OF AN mRNA*

Eukaryotic messenger RNA (mRNA) is transcribed from its DNA template in the 5' to 3' direction by RNA polymerase II (Pol II) in the nucleus[1]. Immediately after the first nucleotides emerge from the polymerase, the nascent transcript is protected at its 5' end by a cap structure[2–4] (**Figure 1.2 A**). Binding of the capping enzyme to the C-terminal domain of Pol II ensures that only Pol II transcripts are capped[5–7].

The simplest eukaryotic 5' cap structure, the so-called cap 0, consists of an N7-methylated guanosine (m^7G) that is linked to the first transcribed nucleotide via a 5'-5' triphosphate bridge (**Figure 1.1**)[2]. This unusual linkage results in a free 3' OH group at the m^7G, which confers stability of the nascent transcript against 5'-3' exonucleases[8,9]. In higher eukaryotes, additional methylation at the 2'-O ribose position of the first and second transcribed nucleotide result in cap 1 and cap 2 structures, respectively[10]. Higher methylated cap structures exist for minor RNA species and trypanosomal mRNAs[10]. Notably, Pol II has been shown to accept nicotinamide adenine dinucleotide (NAD) and 3'-dephospho coenzyme A as non-canonical nucleotides for transcription initiation *in vitro*[11]. Recent evidence confirms the existence of NAD caps for a subset of yeast mRNAs and underscores the potential that also other adenine-containing nucleotide caps could exist[12].

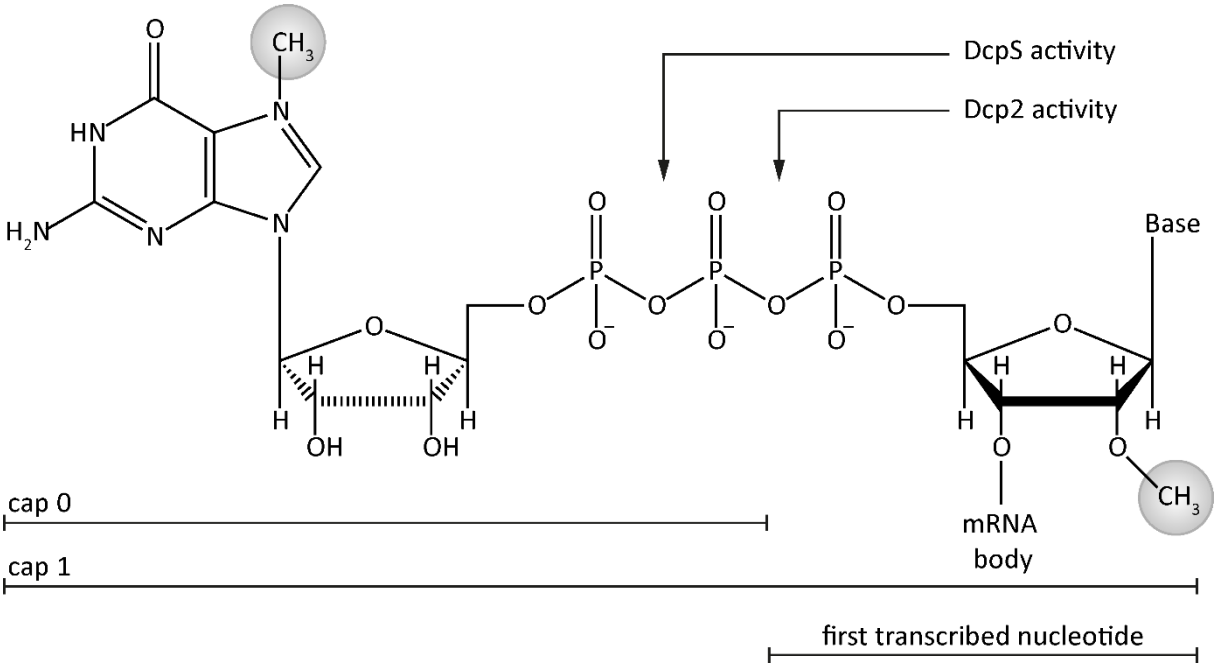

Figure 1.1: Cap structures protect the mRNA from premature 5'-3' exonucleolytic degradation. The cap 0 is characterized by an N7-methylated guanosine that is linked via a triphosphate to the first transcribed nucleotide, while the cap 1 has an additional methyl group at the 2'-O ribose position of the first transcribed nucleotide. Higher order cap structures exist for minor RNA species and in higher eukaryotes. The cleavage sites of the decapping enzymes DcpS and Dcp2 are indicated. Notably, DcpS is inhibited by the Dcp2 decapping product m^7GDP.

The capped, premature mRNA is further processed by splicing events that remove introns[13,14] and by protecting the 3' end by polyadenylation (**Figure 1.2 A**)[15]. The length of the polyadenosine (poly(A)) tail differs among species: in yeast, the poly(A) tails reach lengths of around 50-80 nucleotides while in mammals they can be up to 250 nucleotides long[16–18]. Thereby, long poly(A) tails are usually associated with high mRNA stability[19]. Notably, mRNAs with (not too) short poly(A)-tails can have higher translation rates than long-tailed mRNAs[20]. Replication-dependent histone mRNA is the only eukaryotic mRNA species that lacks a poly(A) tail; instead, these histone mRNAs contain a protective 3' stemloop structure[21].

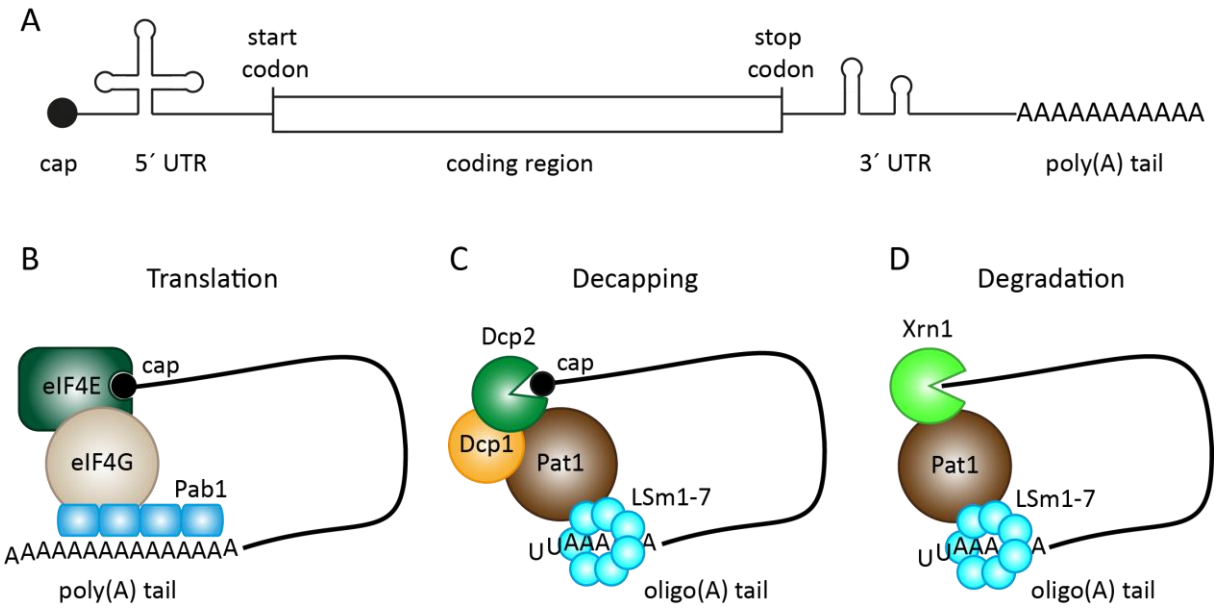

Figure 1.2: Characteristics of mRNA. (A) Structural features of a mature mRNA. The transcript is protected at its 5' end by an N7-methyl guanosine cap that is linked via a triphosphate to the first transcribed nucleotide. The 3' end of an mRNA is protected by a polyadenosine (poly(A)) tail. The coding region starts with the initiation codon AUG and terminates with one of three possible stop codons. The 5' and 3' untranslated regions (UTRs) can contain highly structured segments and regulate mRNA maturation, translation and degradation. **(B)-(D)** Schematic representation of closed-loop messenger ribonucleo-protein (mRNP) complexes during translation initiation **(B)**, decapping by Dcp2 **(C)** and 5'-3' degradation by Xrn1 **(D)**. These three processes are assumed to be enhanced by bridging the 5' and 3' end of the mRNA, which is shown in a simplified manner compared to **A**. Closed-loop structures of translational repressed mRNPs are not shown.

The cap and the poly(A) tail are parts of the 5' and 3' untranslated regions (UTRs), respectively, that flank the protein coding region of a mature mRNA (**Figure 1.2 A**)[22]. The UTRs function in the control of mRNA maturation, localization, stability, translation efficiency and also plays a role in various disease[23–27]. Stable secondary structures that are found in the UTRs can interfere with translation by preventing the ribosome from scanning for the start codon[28]. Additionally, they provide internal ribosome entry sites (IRES)[29,30] or serve as binding sites for regulatory proteins[31,32]. In higher eukaryotes, small RNA species like short interfering RNAs (siRNAs) or micro RNAs (miRNAs) can bind to the 3' UTR, which provides an additional level to regulate gene expression[33–35].

The 5' cap and the 3' poly(A) tail mediate nuclear export of the mature mRNA[36,37]. In the cytosol, the eukaryotic translation initiation factor (eIF) 4E binds to the m^7G cap[38], while the poly(A) tail is bound by the poly(A) binding protein (PABP1 in human and Pab1 in yeast)[39]. Both proteins are bridged by eIF4G, which leads to a closed-loop structure of the messenger ribonucleoprotein (mRNP) complex (**Figure 1.2 B**)[40]. The circular closed-loop is associated with efficient translation[41–43] and protection of the mRNA against decapping and subsequent degradation[44]. Further, the circularization serves as a quality control mechanism to ensure that only properly transcribed and processed mRNAs are translated, as a missing cap or poly(A) tail would prohibit circularization[30]. Notably, closed-loop structures are supposed to exist not only for actively translated mRNAs[45], but also for repressed mRNAs[46,47] or during mRNA degradation (**Figure 1.2 C** and **D**)[48,49].

1.2 *mRNA DEGRADATION IN EUKARYOTES*

Cellular mRNA levels depend on the equilibrium between transcription and mRNA degradation. The amount of actively translated mRNA in a cell must be tightly regulated in a spatiotemporal manner to allow for adaption to environmental changes and to different stages in cell cycle or development. Thus, some mRNA species are turned over rapidly, while others are kept for longer periods[50]. The half-life of mRNAs varies considerably among different species, between minutes and a few hours in yeast[51] and up to several days in mammals[52,53].

Degradation is the final step in the life of an RNA and provides the last possibility for a cell to control gene expression on the RNA level. mRNA degradation not only serves the purpose of routine mRNA turnover but also of differential gene expression. Additionally, aberrant transcripts must be removed from the cell to prevent their potentially dangerous accumulation. Dedicated quality control pathways exist for the degradation of mRNAs that contain premature stop codons (nonsense-mediated decay, NMD)[54], that lack a stop codon (non-stop decay, NSD)[55,56] or that are trapped in stalled ribosomes (no-go decay, NGD)[57,58]. Besides these minor surveillance mechanisms two major mRNA decay pathways exist, a 5' to 3' and a 3' to 5' degradation pathway (**Figure 1.3**)[59,60].

Both pathways rely on the shortening of the poly(A) tail by deadenylation complexes[61,62], whereby deadenylation is the rate-limiting step in mRNA turnover[61]. Deadenylation occurs in a biphasic manner[62], where the Pan2/Pan3 complex trims initially very long poly(A) tails of mature mRNAs[63–65], while further deadenylation by the CCR4-NOT complex leaves only a few adenines on the mRNA (**Figure 1.3**, top)[65–67]. In many eukaryotes, a stretch of uridine nucleotides is attached to the oligo(A) remnant, which finally marks the mRNA for degradation[68–70].

Interestingly, the poly(A)-binding protein PABP plays an ambivalent role by not only promoting translation, but also by recruiting the deadenylation machinery that finally displaces PABP from the mRNA when the poly(A) tail is shortened below a critical length[71,72]. It was found that poly(A) tails with a high occupancy of PABP are deadenylated slowly by Ccr4[72].

On the contrary, poly(A) tails free of PABP are rapidly deadenylated by Ccr4 and Caf1, another deadenylase of the CCR4-NOT complex[72]. Remarkably, low PAPB occupancy on the poly(A) tails was found to correlate with a high degree of sub-optimal codon usage in the coding region of the mRNA[72]. Moreover, mRNAs with poor codon optimality were reported to be occupied also with Dhh1, a DEAD-box RNA helicase involved in translational repression and mRNA degradation[73–75]. Thus, PABP and Dhh1 link codon optimality to mRNA turnover: efficiently translated mRNAs are protected against deadenylation and degradation, while slow translation triggers mRNA decay[76].

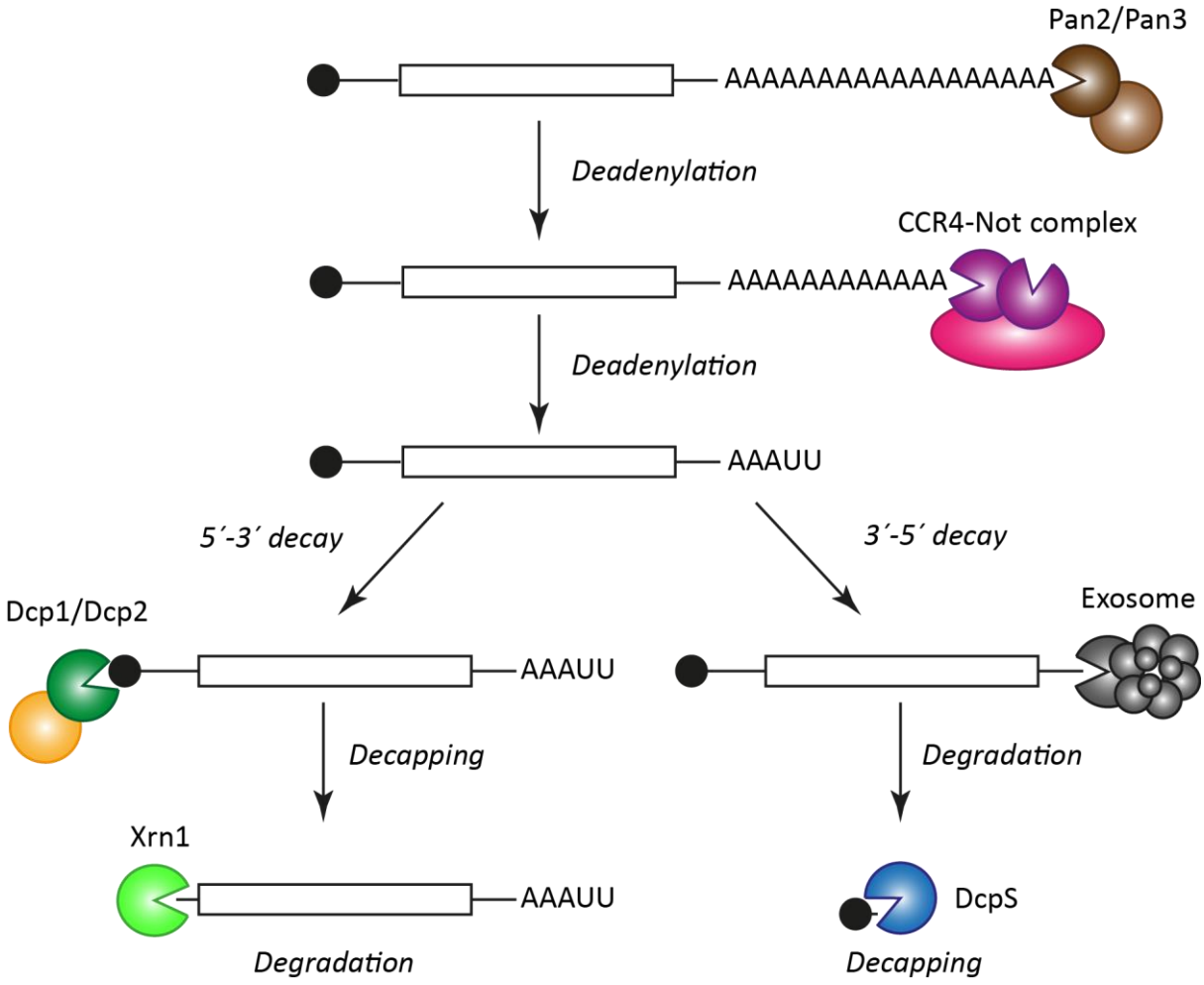

Figure 1.3: Schematic representation of the two major eukaryotic mRNA degradation pathways. Most mRNAs are turned over in a deadenylation-dependent manner. The poly(A) tail is removed in a biphasic process by the Pan2/Pan3 and CCR4-NOT deadenylation complexes. Subsequently, the deadenylated mRNA is subjected to one of two different decay pathways: in 5'-3' decay, irreversible decapping is followed by exoribonucleolytic degradation, while in 3'-5' decay the mRNA is first degraded from its 3' end before the short remnants are decapped. The decapping enzymes Dcp2 and DcpS produce m^7GDP and m^7GMP respectively. The cellular fate of the methylated nucleotides is unknown. The monophosphorylated nucleosides produced by Xrn1 and exosome activity can be recycled in the cell to transcribe new RNA molecules.

1.2.1 *3´-5´ decay*

In 3'-5' decay (**Figure 1.3**, bottom right), the cytosolic exosome complex degrades deadenylated mRNA in a processive, hydrolytic manner into monophosphorylated nucleosides (NMPs)[77–80]. In the cytosol, the exosome is accompanied by the Ski-complex that assists in mRNA recruitment and possesses helicase activity[80,81]. The short mRNA remnants of exosome activity are subsequently decapped by the scavenger decapping protein DcpS[82–84]. Thereby, hydrolysis of the triphosphate linkage between the cap structure and the first transcribed nucleotide releases N7-methyl GMP (m^7GMP) as a product (**Figure 1.1**)[85].

1.2.2 *5´-3´ decay and the mRNA degradation machinery*

In 5'-3' mRNA decay (**Figure 1.3**, bottom left), deadenylation-dependent decapping by the Dcp1/Dcp2 complex precedes exonucleolytic degradation[86]. Removal of the cap interferes with translation initiation, which usually requires the recognition of the m^7G-cap by the eukaryotic initiation factor 4E (eIF4E)[87,88]. As decapping is irreversible, Dcp2 activity inevitably leads to complete degradation of the mRNA. It is thus crucial for a cell to tightly regulate decapping factors to prevent premature mRNA degradation. Dcp2 hydrolyses the cap structure to release 5' monophosphorylated mRNA and m^7GDP[89,90], in contrast to m^7GMP that is produced by DcpS in 3'-5' decay (**Figure 1.1**). Subsequently, the decapped mRNA is hydrolyzed to NMPs in a processive manner by the conserved exoribonuclease Xrn1[91–93]. Notably, DcpS is inhibited by the Dcp2 decapping product m^7GDP, which provides a means to down-regulate the 3'-5' decay pathway if degradation in the 5'-3' direction is highly active[94].

Dcp2 is part of a larger mRNA degradation machinery, whose components increase the low intrinsic decapping activity of Dcp2[95,96]. This degradation machinery arises from a plethora of protein:RNA and protein:protein interactions. Although individual components of the mRNA degradation machinery and their specific interactions are not strictly conserved among different species, the basic principles of mRNA decapping and degradation are found to be similar from yeast to humans.

The Dcp1 protein is the main decapping activator and forms a tight complex with Dcp2 in yeast[97,98]. Dcp1 recruits other decapping factors such as the enhancer of decapping 1 (Edc1), Dhh1 (human DDX6) and Pat1 as well as Xrn1[99,100]. Further, it could be shown that Dcp2 directly interacts with the decapping activators Edc3 and Scd6 (human LSm14)[101–104]. In turn, Edc3 and Scd6 as well as Pat1 are bound via conserved peptide motifs by the DEAD-box helicase Dhh1[105–108].

In yeast, the 5’ cap-recognizing Dcp2 enzyme binds directly to the scaffolding protein Pat1[49]. In turn, Pat1 interacts strongly with the LSm1-7 complex that binds at the 3’ end of the mRNA with a strong preference for oligo(A) over poly(A) sequences[109,110]. Thus, the Pat-LSm complex specifically recognizes deadenylated mRNA and links deadenylation to decapping[48].

The Dcp2:Pat1:LSm1-7 interaction bridges the 5’ and the 3’ end of the mRNA to form a closed-loop structure, which is thought to further enhance decapping (**Figure 1.2 C**)[59,111]. As Pat1 also binds the exoribonuclease Xrn1, Dcp2 can be replaced by Xrn1 after decapping succeeded, thereby maintaining the closed-loop structure to facilitate 5’-3’ degradation (**Figure 1.2 D**)[49]. In human, the scaffolding protein Edc4 adopts the role of yeast Pat1 by mediating the contact between Dcp2 and Xrn1[112], but leaving the closed-loop mechanism untouched[49].

1.3 *LIQUID-LIQUID PHASE SEPARATION*

Given the many interactions between mRNA degradation factors, it is not surprising that they were found to co-localize in the cytosol. These foci that are enriched in mRNA degradation factors were named processing bodies (P-bodies) and appeared to be membrane-less compartments[113,114]. P-bodies or other cytosolic and nuclear foci are thought to arise from a process referred to as liquid-liquid phase separation (LLPS)[115–117]. Over the recent years, LLPS evolved into a widely accepted mechanism for subcellular compartmentalization[118]. Concepts from polymer physics have been used and extended to provide the theoretical framework to describe cellular and reconstituted phase separation processes involving very heterogeneous biological polymers such as proteins and RNA[119].

The physical properties of phase-separated cellular bodies have first been described for P granules, germ line-specific RNPs in *Caenorhabditis elegans*[117]. P granules exhibit properties of liquid droplets. As such, they are spherical in shape, they fuse, and deform under shear stress[117]. Fluorescence recovery within seconds revealed highly dynamic granule components and a viscosity similar to that of glycerol[117]. Additionally, the surface tension between the P granules and the cytoplasm was found to be quite small, which facilitates rapid and reversible dissolution and condensation of P granules that is required for proper *C. elegans* embryo development[117]. In the last years, liquid-like behavior has been demonstrated for a variety of phase separated droplets[103,120–124].

RNP containing cellular foci can be grouped into nuclear and cytosolic granules[118,125]. The first group comprises for example nucleoli[126], Cajal bodies[127], Para speckles[128], Histone locus bodies[129], PML bodies[130] and nuclear pore complexes[131], while P-bodies[113], stress granules[132], germ (P) granules[117] and Balbiani bodies[133] are in the cytosol. Additionally, signaling complexes[134–136] and biosynthetic clusters such as purinosomes[137] can also form by phase separation processes.

Cellular phase transitions are a result of supersaturation of proteins and nucleic acids[119]. In a cell, this can be achieved for example by regulating gene expression or the charge state of proteins. Indeed, methylation, acetylation and phosphorylation interfere with LLPS[120,138–141]. Changes in temperature do not only influence gene expression but also have direct effects on phase separation, as some RNPs undergo phase separation at elevated temperatures while others phase separate in the cold[142]. Additionally, a cell reacts to environmental stress factors such as osmotic or pH shocks with phase transitions that result in compartmentalization of specific proteins and RNAs. Besides that, *in vitro* phase separations are influenced directly by changes in salt or proton concentration (pH).

Three main driving forces for liquid-liquid phase transitions have been determined for proteins: (1) interactions within low complexity regions, (2) multivalent interactions involving folded domains and (3) protein:RNA interactions[118]. Thereby, different interaction modes can act simultaneously in phase separated droplets to give rise to the high redundancy observed for many LLPS processes[103,143].

Low complexity regions are unfolded protein segments with limited compositional diversity that are often enriched in glycine, polar, aromatic or charged residues. These intrinsically disordered regions (IDRs) are found frequently in proteins undergoing LLPS. Thereby, the IDRs mediate intermolecular contacts via charge-charge, cation-π, dipole-dipole and π-π interactions: For example, the P granule protein LAF-1 undergoes homotypic phase separation due to interacting clusters of positive and negative charges. For the DEAD-box RNA helicase DDX4, phase separation is dependent on an overrepresentation of aromatic FG/GF repeats within clusters of positive charge[120]. Dipole-dipole interactions dominate phase separation of prion-like IDRs in LSm4, huntingtin, Whi3 and a set of mRNA degradation factors that harbor stretches of poly-glutamine or -asparagine[122,144–146]. Lastly, proteins related to amyotrophic lateral sclerosis (ALS) such as FUS, hnRNPA1 and hnRNPA2 form amyloid-like fibrils that are stabilized by ladders of aromatic side chains[123,147–149]. Along these lines, phenylalanine-to-serine mutations within the FG-repeat containing nuclear pore protein Nsp1p interfere with the formation of hydrogel-like assemblies, highlighting the importance of π-π interactions for cellular phase transitions[131]. Notably, phase separations that involve ionic interactions rely on the clustering of charge, while a more equal charge distribution was found to disfavor intermolecular interactions[120,143].

Multivalency is an instrumental aspect of cellular and *in vitro* reconstituted liquid-liquid phase separation. Many phase separations rely on weak but multivalent interactions between the involved binding partners. For example, the tripartite system of nephrin, NCK and N-WASP associates via a set of multivalent interactions that manifests in phase separation. First, nephrin contains three phospho-tyrosine sites, which are recognized by the NCK SH2 domain. And second, the three SH3 domains in NCK can be bridged by N-WASP that contains six proline-rich motifs (PRMs).

The importance of multivalency has also been demonstrated for artificial two-component LLPS systems. One such system consists of multiple SH3 and PRM repeats on two separate polypeptide chains, where the degree of *in vitro* phase separation is directly dependent on the number of SH3 and PRM modules within the two proteins[134]. Additionally, $(SH3)_5$ and $(PRM)_5$ proteins were found to co-localize in liquid-like compartments in living cells, indicating that multivalent interactions are sufficient to induce cellular phase transitions[134]. In another multivalent two-component LLPS system, that has been engineered from multiple copies of SUMO and SUMO-interaction motifs (SIMs), phase separation and the strength of interaction also scaled with the number of compatible modules[150]. Other multivalent interactions important for physiological LLPS processes are found between the Edc3 LSm domain and several helical-leucine rich motifs (HLMs) in Dcp2[103] or between multiple RNA-recognition motifs (RRMs) in the polypyrimidine tract binding protein (PTB) and UCUCU repeats in RNA[134].

RNA is a key component of many cellular granules[125]. Interactions between RNA and IDRs of several proteins such as FUS, hnRNPA1 or LSm4 have been shown to promote LLPS *in vitro*[151]. Analogously, mRNA binding to a folded RRM enhances poly(Q)-driven phase separation of recombinant Whi3[122]. In hnRNPA1, RNA-binding to the two RRMs induces phase separation even in the absence of the low complexity region[123] and in the case of PTB, binding of pyrimidine-rich clusters to the four RRMs is essential for LLPS[134]. Thus, disruption of RNA-binding can result in decreased phase separation and cellular foci formation, as has been shown exemplarily for Pat1, where phosphorylation of the C-terminus interferes with RNA-binding *in vitro* and P-body formation *in vivo*[152,153].

It has been observed frequently that liquid-liquid phase separated proteins and RNPs can undergo a second phase transition to a more solid- or gel-like state[145,154–156]. This second transition, also referred to as maturation, can result in the formation of dissolution- and salt-resistant structures with non-spherical morphology[122,151,157]. In some cases, droplet maturation was found to be driven by the formation of amyloid-like fibrils that are associated with neurodegenerative disorders such as frontotemporal dementia (FTD) or amyotrophic lateral sclerosis (ALS) and disease-related mutants often show enhanced fiber formation[121,123,148]. Notably, *in vivo* maturation processes can also lead to functional instead of pathological states. For example, Balbiani bodies in *Xenopus leavis* oocytes, yeast stress granules or nuclear pore complexes behave more like solids or hydrogels than like liquids[116,133,158].

Remarkably, high concentrations of RNA were found to prevent fibrillization and to slow down phase transitions in some cases[121,122,159], although lower RNA concentrations frequently promote LLPS (see above). The RNA-dependent reduction in fiber formation has been linked to the charge screening properties of the polyanionic RNA[122].

Similarly, sub-physiological ATP levels facilitate phase separation of FUS, while physiological ATP concentrations between 5 and 10 mM result in droplet dissolution and frequently in protein solubilization and stabilization[160–162]. The effect of ATP on LLPS has been attributed to its hydrotropic properties[160]. These results hint at cellular mechanisms beyond posttranslational modifications (see above), autophagy-mediated clearance[163] or ATP-dependent chaperone and Dhh1 activity[12,164] to control LLPS and RNP homeostasis and to prevent pathological fibrilization.

In this thesis, I used the conserved decapping factors and P-body components Edc3 (CHAPTER 2 and CHAPTER 3) and Dhh1 (CHAPTER 4) to study LLPS mechanisms at an atomic level.

1.4 *NMR SPECTROSCOPY*

This chapter contains parts written for a review that I co-authored which has been accepted by "Progress in Nuclear Magnetic Resonance Spectroscopy" for publication and which is currently in press.

Until recently, biomolecular NMR spectroscopy studies of proteins with a molecular weight over 40 kDa were challenging and rare. For these systems rapid spin relaxation rates prevented the routine recording of high-quality NMR spectra[165]. Currently, this molecular weight limit of solution-state NMR spectroscopy has been shifted significantly and numerous reports demonstrated that complexes that are (far) over 100 kDa in size are amenable to detailed NMR studies. These advances can be ascribed to two important technological advances. On the one hand, sample preparation and isotope labeling methods have been established, where partial or complete deuteration has resulted in significant decreases in transverse relaxation rates by eliminating ^{1}H-^{1}H dipole-dipole coupling[166–173]. On the other hand, the exploitation of transverse relaxation optimized spectroscopy (TROSY) effects[174–177] has resulted in additional and significant sensitivity gains in protein NMR spectroscopy[178,179]. These TROSY approaches were initially introduced for ^{1}H,^{15}N-labeled proteins, and later adapted to aromatic ^{1}H-^{13}C spin systems[180] and $^{13}CH_3$-labeled methyl groups[181].

1.4.1 *The TROSY experiment*

The amide ^{1}H-^{15}N spin system possesses four different energy levels arising from the combination of α and β spin states of the ^{1}H and ^{15}N spins. The four energy levels can be described as magnetization terms (coherences), which are created by the pulse sequence of an NMR experiment. The four coherences each have an individual relaxation rate, whereof some coherences relax significantly slower than others due to the destructive interference of different relaxation mechanisms. For the amide ^{1}H-^{15}N spin system, transverse relaxation receives strong contributions from dipole-dipole coupling (DD) and chemical shift anisotropy (CSA).

^{1}H-^{15}N DD and ^{15}N CSA where found to interfere destructively for one of the four ^{1}H-^{15}N coherences in an heteronuclear single quantum correlation (HSQC) experiment, which leads to slow relaxation of this magnetization term[176].

In a traditional ^{1}H-^{15}N HSQC experiment[182], all four magnetization terms are mixed, which results in an averaged relaxation rate and a relatively broad resonance of medium intensity. Contrarily, NMR experiments that exploit the TROSY effect select the slowly relaxing coherence and keep it separated from fast relaxing terms throughout the pulse sequence[176]. Due to its slow relaxation, this magnetization term leads to a sharp and intense signal in the spectrum. As only one fourth of the equilibrium magnetization is finally used to record the NMR spectrum, TROSY-type experiments are preferably applied to large proteins and complexes, where relaxation is a severe issue. For small proteins with slow relaxation rates the sensitivity gains due to the TROSY effect do not compensate for the loss of three quarters of the initial magnetization.

In contrast to DD, the CSA for ^{1}H and ^{15}N nuclei in amide groups is dependent on the strength of the external magnetic field. For amide groups, optimal cancellation of DD and CSA and thus the most efficient TROSY effect was found to occur at field strengths of about 21 T, corresponding to a proton Larmor frequency of 900 MHz[176,183]. Due to the different CSA of ^{13}C and ^{15}N, aromatic CH-groups show an optimal TROSY effect at about 14 T (600 MHz proton frequency)[180]. The largest possible peak heights, in contrast to slowest transverse relaxation rates, are obtained with spectrometers operating at even higher proton frequencies of 900 MHz for ^{13}C-detected TROSY on aromatic CH-moieties and, theoretically, 1.5 GHz for ^{1}H-detected TROSY on amide groups[183]. Notably, the most powerful NMR spectrometers that are currently being developed will operate at a proton frequency of “only” 1.2 GHz.

1.4.2 *Methyl TROSY*

It has turned out that the combination of specialized methyl group labeling schemes and application of the methyl TROSY technique is one of the most successful approaches to make solution-state NMR spectroscopy amenable to assemblies that are far over 200 kDa[184]. Notably, amide TROSY and methyl TROSY rely on different principles as the former is effective on AX spin systems such as amide ^{15}N-^{1}H moieties (or aromatic CH-groups), while the latter requires an AX_3 spin system as found in $^{13}CH_3$-labeled methyl groups. In contrast to the amide ^{15}N or aromatic ^{13}C chemical shift anisotropy, the methyl ^{13}C CSA is very small and can thus not interfere with the large dipole-dipole couplings. Instead, proton-carbon and proton-proton dipolar interactions interfere destructively in the isolated ^{1}H-^{13}C spin system of methyl groups in high molecular weight proteins. As transverse relaxation in methyl groups is dominated solely by dipolar interactions, the methyl TROSY effect is independent on the magnetic field strength.

In methyl groups, the carbon atom is connected to three protons. As each ^{1}H and ^{13}C spin either adopts an α or a β spin state, the combination of all possible spin states results in 16 different energy levels. The energy levels are connected by 28 fast and slowly relaxing single-quantum proton, single-quantum carbon and heteronuclear double-/zero-quantum transitions. The group of Lewis Kay could show that the fast and slowly relaxing coherences never interconvert in a ^{1}H-^{13}C heteronuclear multiple quantum correlation (HMQC) experiment, which thus is an intrinsic TROSY experiment for methyl groups in high molecular weight proteins[181].

The ^{1}H-^{13}C HMQC experiment was found to be up to three-times more sensitive for methyl groups than the standard ^{1}H-^{13}C HSQC experiment[181]. In the ^{1}H-^{13}C HSQC pulse sequence, several 90° ^{1}H pulses interconvert fast and slowly relaxing methyl coherences multiple times, which results in broader and weaker signals for large proteins. Contrarily, the application of only a single 90° ^{1}H pulse in the ^{1}H-^{13}C HMQC pulse sequence prevents mixing of the differentially relaxing methyl coherences, which is essential for the gain in sensitivity. As a result, methyl resonances of large proteins are comparably sharp and intense, especially relative to amide resonances of the same protein.

1.4.3 *Methyl labeling*

Methyl groups occur in around one third of the proteinogenic amino acids (alanine, threonine, valine, leucine, isoleucine and methionine) and are thus abundant probes to study protein structure, function and dynamics. Routine experiments that exploit the methyl TROSY effect are most efficient on fully protonated and ^{13}C-labeled methyl groups ($^{13}CH_3$) that are embedded in an otherwise uniformly deuterated background[181]. Deuteration eliminates dipolar interactions with non-methyl protons that would lead to additional relaxation mechanisms. However, as ^{13}C has a natural abundance of only 1.1 %, methyl TROSY experiments require the enrichment with NMR-active nuclei by either providing appropriate isotope sources during protein expression (see below) or by posttranslational modifications with isotope-labeled tags[185,186].

During cell-based protein expression in *Escherichia coli (E. coli)*, natural metabolic pathways of the expression host are utilized to selectively channel $^{13}CH_3$-labeled methyl groups into specific residues. To that end, isotope labeled amino acids or amino acid precursors are added to the growth medium of the cells. To ensure that these labeled compounds only end up in the target sites and are not “scrambled” into other amino acids, it might be required to add additional unlabeled metabolites[187] or to genetically modify the expression host[188,189].

Most of the commonly applied labeling schemes work well in growth media that are based on D_2O as a solvent and glucose as the main carbon source. In case other carbon sources such as glycerol are used, the dominant cellular metabolic pathways change, which can result in a situation where specific precursors are no longer solely used in the corresponding amino acid synthesis pathway.

Methyl labeling strategies have been developed for Ala-β[190,191], Ile-γ2[192,193], Ile-δ1[165,181,194–196], Met-ε[197–200] and Thr-γ2[188,201–203] methyl groups as well as for the simultaneous labeling of the Leu-δ and Val-γ positions[194,204–207]. To reduce spectral overlap in the crowded region of Leu and Val methyl resonances, Leu[208] or Val[188,189,209] can be labeled separately or stereo-specific labeling of the pro-*(S)* (Leu-δ2 and Val-γ2)[210,211] or pro-*(R)* (Leu-δ1 and Val-γ1)[211] methyl groups can be applied.

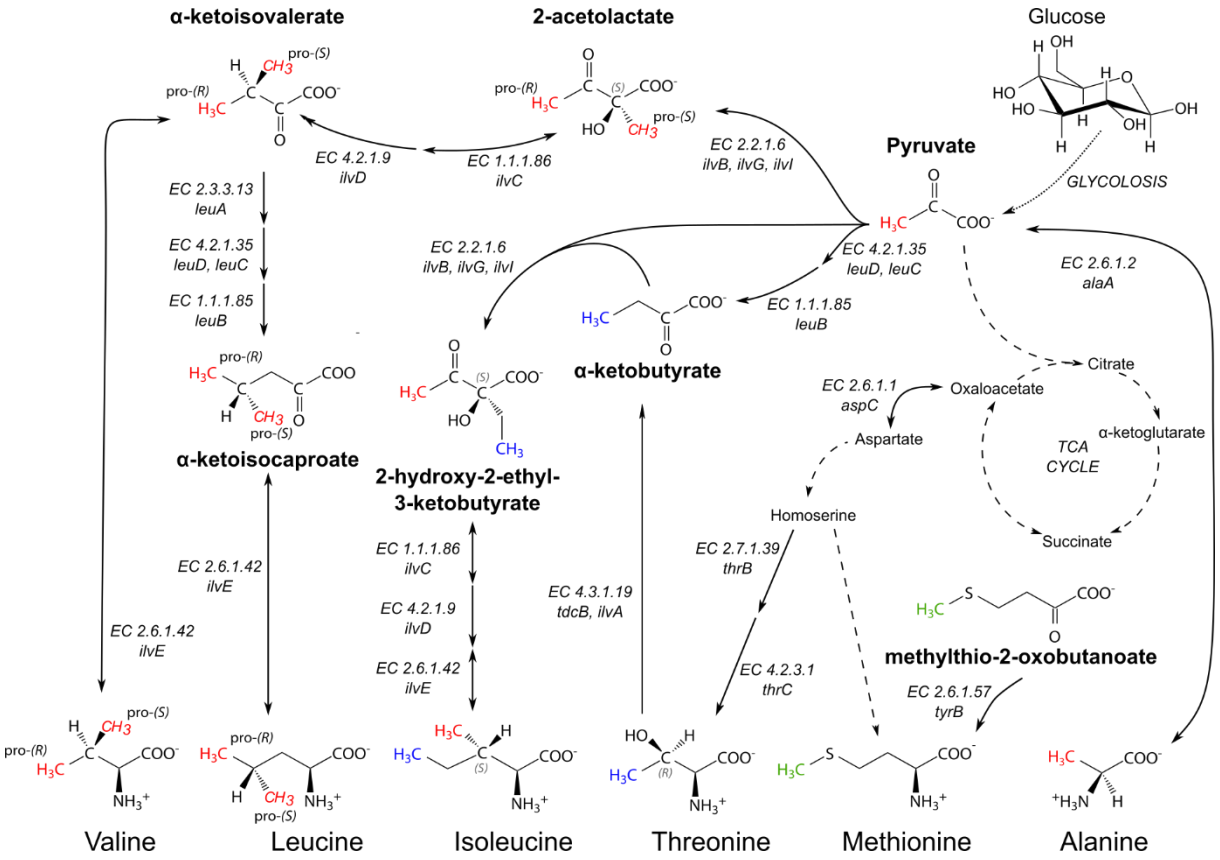

Figure 1.4: Schematic biosynthetic pathways of the methyl-bearing amino acids (Ala, Ile, Leu, Met, Thr and Val) in *E. coli*. Amino acids and key metabolites, that can be used as precursors for methyl labeling, are depicted with their structural formulas. Full arrows indicate one-step reactions, while dashed arrows resemble multiple reactions, double-headed arrows indicate reversible reactions. The scrambling pathways of the Ala-β methyl group (red) into leucine, valine and isoleucine-γ2 and of the Thr-γ2 methyl group (blue) into isoleucine-δ1 are indicated. The methionine methyl group (green) does not scramble. The enzymes or enzyme complexes that catalyze the biosynthetic reactions are abbreviated with their EC number and gene names. EC 1.1.1.85: 3-isopropylmalate dehydrogenase, EC 1.1.1.86: ketol-acid reductoisomerase (KARI), EC 2.2.1.6: aceto-hydroxy-acid synthase (AHAS), EC 2.3.3.13: 2-isopropylmalate synthase, EC 2.6.1.1: aspartate aminotransferase, EC 2.6.1.2: glutamate-pyruvate aminotransferase, EC 2.6.1.42: branched-chain amino acid aminotransferase (BCAT), EC 2.6.1.57: aromatic-amino-acid transaminase, EC 2.6.1.66: alanine-valine transaminase, EC 2.7.1.39: homoserine kinase, EC 4.2.1.9: dihydroxyacid dehydratase, EC 4.2.1.35: 3-isopropylmalate dehydratase, EC 4.2.3.1: threonine synthase, EC 4.3.1.19: threonine deaminase. Further information on the biosynthetic pathways can be found online: https://www.genome.jp/kegg/

Usually, a combination of methyl groups is labeled[184,196,200,206,212]. Choosing the correct precursors and supplements, all possible combinations of methyl-labeling can be achieved. Although not done frequently, methyl labeling of all methyl-bearing amino acids (ILVMAT) has been shown[203,213].

In our group, we exploit advanced labeling schemes such as IM-[214,215], ILVM- (Damman, Schütz, et al., under revision) or ILVMA-labeling[96] (Schütz et al., in preparation). We also label subsets of amino acids such as IA and IV, if residue-type specific assignments are required, for example as a prerequisite for automated assignment algorithms. In the case of IA- and IV-labeling, label scrambling to Leu/Val and Leu methyl groups is suppressed by supplementing unlabeled α-ketoisovalerate and α-ketoisocaproate, respectively. Whenever feasible, we make use of D_2O with a deuteration level of less than 100% and of protonated supplements for suppression of label scrambling, without compromising the information content of our experiments. This strategy works well for proteins and complexes with a molecular weight of up to 100 kDa. However, we frequently observe that methyl-methyl NOE and protein dynamics experiments such as relaxation dispersion require the highest possible level of deuteration.

In this thesis, I applied ILVM-labeling to the Edc3 YjeF domain (see CHAPTER 3) and ILVMA-labeling to the helicase core or isolated RecA-like domains of Dhh1 (see CHAPTER 4). For residue-type specific assignments of the Dhh1 RecA-like domains, I also used IA-, IV-, IMV- and IMA-labeling (see CHAPTER 4).

1.4.4 *Methyl resonance assignment*

The assignment of methyl resonances to specific methyl groups in the protein is a prerequisite for the analysis of methyl TROSY NMR data. For methyl groups in large proteins and complexes, this process can be time-consuming and challenging, especially when traditional methyl resonance assignment strategies that rely on assigned backbone resonances fail.

Methyl assignment via through-bond correlations

For small proteins, it is usually possible to assign the backbone resonances through traditional methods[216]. In case the spectral quality allows, the methyl resonances can subsequently be assigned by correlating the methyl chemical shifts with assigned backbone and/or side chain resonances based on total correlated spectroscopy (TOCSY) transfer methods. Due to fast signal relaxation processes, this assignment approach will fail for larger proteins. We observe that proteins and complexes with a molecular weight exceeding 25 kDa require full deuteration and special methyl labeling schemes to assign methyl groups in this manner[217,218].

Significant magnetization losses, that occur during the TOSCY transfer times, can be prevented by using a series of correlated spectroscopy (COSY)-type magnetization transfer steps as these ensure that magnetization from the methyl groups is solely transferred to one or a few specific backbone nuclei[219]. The complete magnetization transfer via COSY is, however, insensitive due to the large number of transfer steps. Thus, more sensitive methyl-detected "out-and-back" experiments have been introduced.

In those experiments, the magnetization is transferred from the methyl groups to side chain or carbonyl carbons with known chemical shifts and subsequently back to the methyl protons for detection[206,220,221].

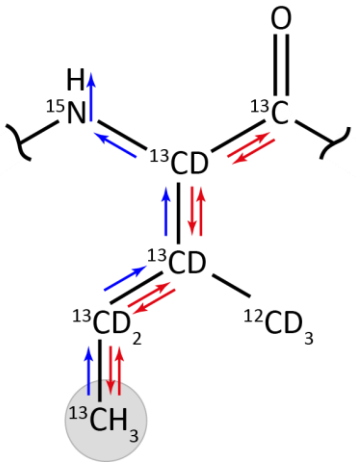

Figure 1.5: Methyl resonance assignment strategy based on J-couplings. Chemical shifts of methyl proton and carbon atoms are correlated with assigned backbone or side chain chemical shifts. Blue arrows indicate magnetization transfer pathways from the methyl groups to the backbone amide protons for detection. Red arrows indicate the flow of magnetization in "out-and-back" experiments, where magnetization is transferred from the methyl protons via the methyl carbon atoms to side chain and backbone carbons with known chemical shifts and back to the methyl protons for detection.

The assignment of methionine methyl resonances by spectroscopic methods is challenging as the methyl group is an isolated spin system that is separated from the other side chain atoms by an NMR-inactive sulfur atom. Nevertheless, for low molecular weight proteins, small ^{13}C-^{13}C and ^{1}H-^{13}C long-range J-couplings have been exploited to link the methyl group to the rest of the side chain[222]. For large proteins (> 20 kDa), this strategy will most likely not be of the required efficiency to provide any assignment information.

In this thesis, I initially aimed for an assignment of methyl resonances of the Dhh1 RecA-like domains based on through-bond correlations with backbone amide and side chain Cα and Cβ chemical shifts. However, the triple-resonance experiments required for the assignment of backbone resonances were of insufficient quality to assign methyl groups with this strategy. I thus turned to alternative assignment strategies that are discussed below.

Divide-and-conquer

For most proteins that have a molecular weight over 50 kDa, the assignment of the backbone becomes challenging[223]. This often prevents the assignment of methyl group resonances based on the backbone assignment. In the divide-and-conquer approach, a large complex or multi-domain protein is dissected into smaller building blocks. In case the fold of the building blocks is preserved in isolation, it is possible to transfer the traditionally obtained assignments from the small part onto the larger assembly (**Figure 1.6**).

This approach has turned out to be useful for symmetric multi-subunit assemblies, in case the individual subunits can be prepared in a monomeric form[184]. Changes in the chemical shifts between the monomer and the fully assembled complex are often limited, especially in the core of the protein building block. Hence, a straightforward transfer of the assignments from the subunit to the complex is possible. In addition, the divide-and-conquer approach has been successfully applied to complexes that contain more than one unique subunit[224] and to large multi-domain proteins[200,214,225,226].

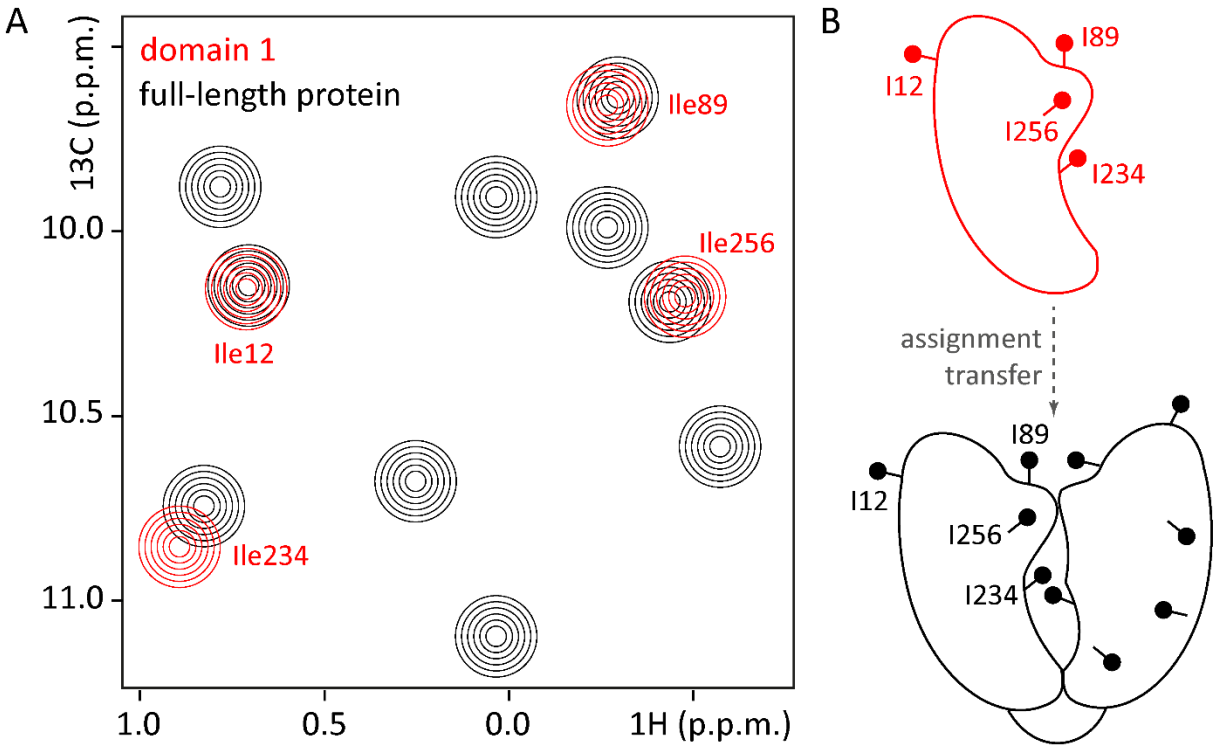

Figure 1.6: The divide-and-conquer approach is used to assign multi-domain proteins or multi-subunit complexes. **(A)** Schematic ^{1}H-^{13}C correlation spectrum of a hypothetical protein (black) and of one of its domains in isolation (red). Assignments from the isolated domain are easily transferred to the full-length protein due to limited chemical shift perturbations. Resonances corresponding to residues that experience a similar chemical environment in the isolated domain and in the full-length protein do not show chemical shift perturbations (see for example the resonance of Ile12 in **A**; Ile12 is remote from the domain interface in the full-length protein **(B)**). However, residues that experience a different chemical environment in the isolated domain compared to the full-length protein give rise to resonances that exhibit (small) chemical shift perturbations (see for example the Ile234 resonance in **A**; Ile234 is in the interface of the two domains in the full-length protein **(B)**). **(B)** Schematic representation of an isolated domain (red), whose assignments are transferred to the hypothetical full-length protein (black).

After the transfer of the resonance assignments from the building block to the large assembly, it is required to validate that these are indeed transferred correctly. To that end, additional information is required that can, for example, be derived from the comparison of the chemical shifts of additional side chain carbon atoms through “out-and-back” J-based experiments (see above), or from NOE based experiments (see below).

In this thesis, the divide-and-conquer approach has been applied to assign the Edc3 YjeF methyl resonances in the Edc3 ΔLSm construct, that comprises the IDR and the YjeF-domain (see CHAPTER 3), and to assign the ILVMA-methyl resonances of Dhh1 based on methyl resonance assignments of the isolated RecA1 and RecA2 domains (see CHAPTER 4).

Methyl assignment via site-directed mutagenesis

For very large or challenging protein complexes the assignment strategies mentioned above might not be applicable. In addition, in many cases, a full methyl group assignment might not be required to address the question at hand. In those cases, a limited or full methyl group assignment can be obtained through a mutagenesis approach. In this approach, a methyl-bearing residue of interest is mutated into a (closely related) other amino acid, without distorting the fold of the protein. Methyl TROSY spectra are subsequently recorded for the wild-type and for the mutant protein[227]. In the ideal case, both spectra are identical apart from one (alanine, methionine, threonine, isoleucine) or two (valine, leucine) resonances that are absent in the spectrum of the mutated protein. These resonance(s) then correspond to the methyl group(s) of the mutated residue (**Figure 1.7**).

In several cases, the mutagenesis approach has proven to be successful for obtaining methyl group assignments of large complexes[214,215,228,229]. In addition, the mutagenesis approach complements other assignment strategies, for example those that are based on the divide-and-conquer approach or the NOE-based approach for residues with other methyl groups in spatial proximity (see below). In practice, this assignment procedure can be complicated in case the introduced mutation results in severe chemical shift perturbations (CSPs) of other resonances so that the peak reporting on the mutation can no longer be unambiguously identified[184]. In that case, it might be necessary to include a large number of mutations to be able to distinguish between primary and secondary CSPs[228].

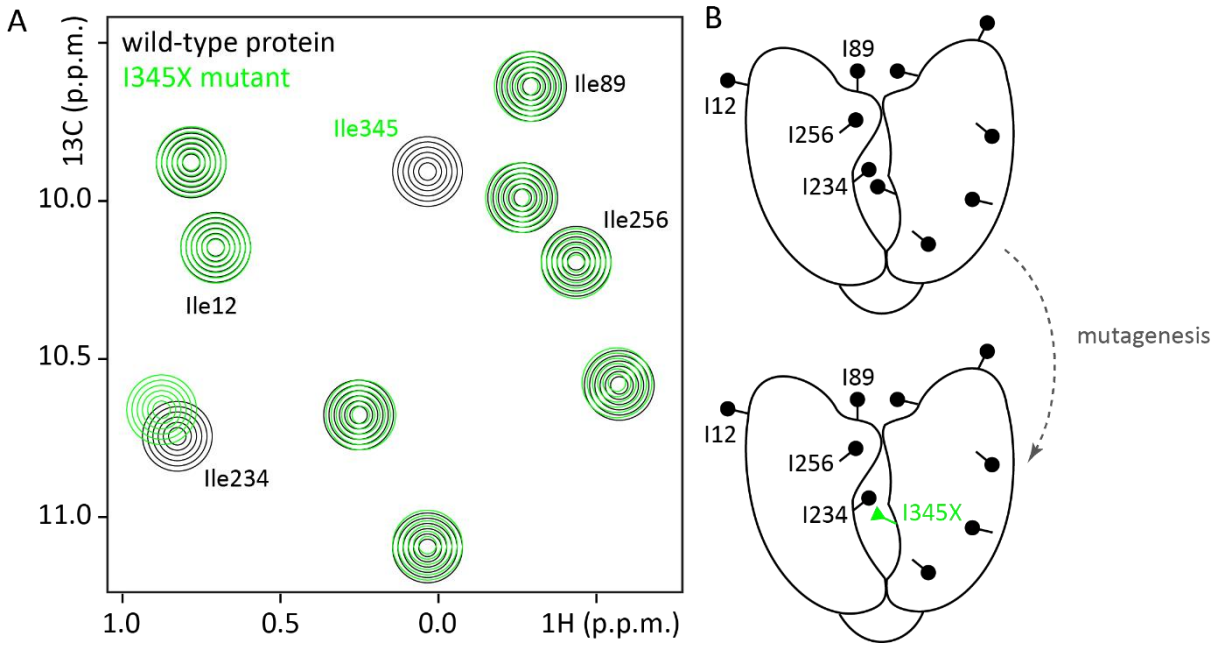

Figure 1.7: Methyl group assignment based on a mutagenesis approach. (A) Schematic ^{1}H-^{13}C correlation spectra of the hypothetical protein (black) and of a mutant (green), where one Ile residue (Ile345) has been mutated into another amino acid X. The Ile345 resonance is thus missing in the spectrum of the mutant protein, while the resonances of all other Ile methyl groups superpose well in the wild-type protein (black) and in the mutant (green). An exemption is the Ile234 resonance, which experiences a small chemical shift perturbation. This is due to the proximity of Ile234 to the mutation site at position 345 that results in a different chemical environment for Ile234 in the wild-type compared to the mutant protein. **(B)** Schematic representation of the hypothetical protein (black), where Ile345 is mutated to another amino acid X (green).

Here, I utilized the mutagenesis-driven assignment approach to assign several methyl groups in the Edc3 YjeF domain as a prerequisite to apply the methionine scanning methodology (see 1.4.6 and CHAPTER 3).

Assignments based on NOEs

Even for very large complexes, it is possible to obtain methyl-methyl NOE contacts with high sensitivity[184]. Methyl-methyl NOE spectra reveal methyl resonances that are derived from methyl groups close in space. This data can be used to validate assignments obtained by the divide-and-conquer approach through comparison of NOE patterns (**Figure 1.8**). In case high resolution structural information of the complex is available, it is possible to directly compare experimental NOE cross-peak patterns with expected, back-calculated NOE patterns. Based on that, assignments of residues that are close in space to already assigned residues can be accomplished.

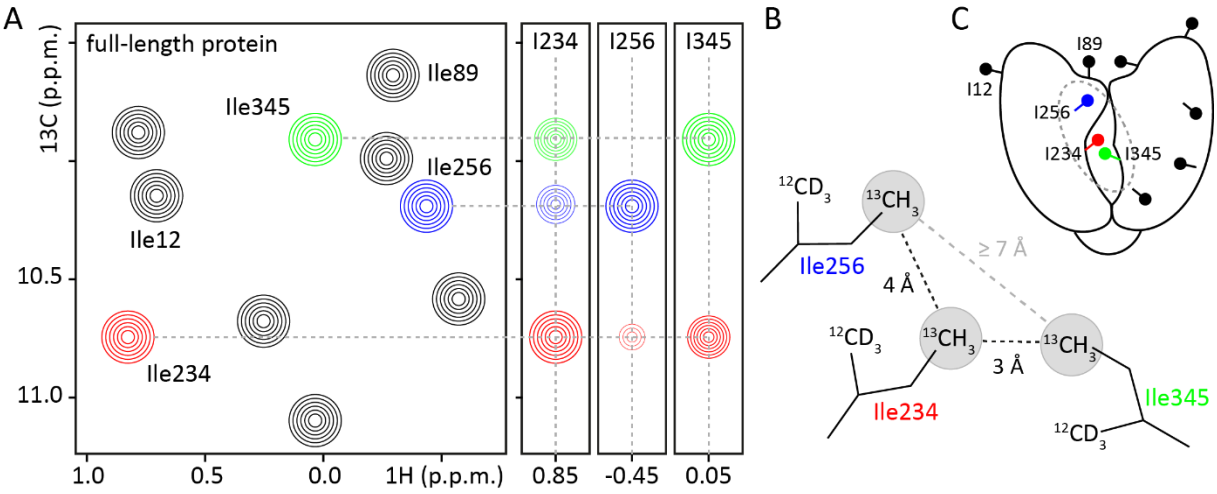

Figure 1.8: Methyl group assignment based on NOEs. (A) Schematic 2D methyl TROSY spectrum of a hypothetical protein labeled at the Ile-δ1 methyl groups (left) and three exemplary NOESY "strips" from a C-C-H experiment (right). Dashed lines indicate matching resonances in the 2D spectrum and the NOESY strip of Ile234. The NOESY strips show one intense peak for the respective Ile-δ1 methyl group and less intense cross-peaks for each Ile-δ1 methyl group that is close in space. Notably, the cross-peak intensity decreases with the distance between the methyl groups (r^{-6} dependence). For example, Ile234 (red, left strip) is in close proximity of Ile345 (green; see **B** and **C**) and the Ile345 cross-peak is thus quite intense. Compared to Ile345, Ile256 (blue) is more distant to Ile234 and its cross-peak is weaker. Above an inter-methyl distance of approximately 7 Å, NOE cross-peaks are no longer detectable using routine NOESY-experiments (see also panel **B**). For example, Ile345 shows an NOE to Ile234, but not to Ile256 (see **A**, right NOESY strip). In practice, the NOE transfer between two methyl groups can be more efficient in one than in the other direction, giving rise to cross-peaks of different intensity. For example, the cross-peak of Ile234 to Ile256 is more intense than the cross-peak of Ile256 to Ile234 (compare left and middle NOESY strip in **A**). **(B)** Schematic drawing of the three Ile side chains discussed above. Black dashed lines indicate observable NOEs, while a gray dashed line indicates an inter-methyl distance that is too large to detect NOEs. **(C)** Schematic representation of the hypothetical protein, where the assignment for Ile345, that has been obtained by a mutagenesis approach (see above), can be confirmed by an NOE to the proximal Ile234. The dashed ellipse indicates the area that is enlarged in **B**.

Experimentally, inter-methyl NOEs are readily obtained with the use of 3D HMQC-NOESY[230] or 4D HMQC-NOESY-HMQC[231,232] experiments. Due to the lower dispersion of the proton chemical shifts in methyl groups, H-C-H correlations are usually less informative than C-C-H correlations.

The recent development of non-uniform sampling (NUS)-based experiments[233,234] allow for the relatively fast recording of 4D H-C-C-H correlations that provide unambiguous information on methyl resonances that are close in space[231,232,235,236].

Computational methods for methyl assignments

Automated strategies try to facilitate the tedious processes that accompany the methyl group assignment process. These programs use experimental NMR data and a protein structure as input and ideally result in a complete and reliable assignment of all methyl group resonances. The programs MAP-XS[237] and FLAMEnGO[238] use a swapping procedure to reach convergence between measured and predicted NOEs. Both programs can include additional data such as paramagnetic relaxation enhancements (PREs), residual dipolar couplings (RDCs), and pseudo-contact shifts (PCS) as well as assignments from mutagenesis or chemical shift predictions. The programs MAGMA[239] and MAGIC[240] exploit graph theory to correlate experimental and back-calculated NOE patterns. Software to map experimental PRE- or PCS-based data onto known structures have also been introduced[241,242].

In this work, I used the NOE-based MAGIC algorithm for the computational assignment of ILVMA methyl resonances in the isolated Dhh1 RecA1 and RecA2 domains (see CHAPTER 4).

1.4.5 *CSP experiments*

The strength of NMR spectroscopy is its ability to pick up the subtle differences in the chemical environment of individual spins. Hence, in an NMR spectrum, each spin gives rise to a resonance with a characteristic chemical shift. Chemical shift perturbation (CSP) experiments exploit the sensitivity of a spin to changes in the chemical environment. CSP experiments are most frequently performed as 2D heteronuclear correlation experiments, either on amide (in ^{1}H-^{15}N correlations) or on methyl groups (in ^{1}H-^{13}C correlations) as NMR probes. In the latter case, each isotopically labeled methyl group within a protein gives rise to a resonance in ^{1}H-^{13}C correlation experiments such as the HMQC experiment which exploits the methyl TROSY effect[181]. **Figure 1.9 A** schematically shows the resonances of Ile methyl groups in a ligand-binding protein. Ligands include but are not limited to ions, small molecules such as drugs, metabolites or other proteins.

In presence of a ligand, the chemical environment of methyl groups in or close to the ligand binding site (**Figure 1.9 B**, e.g. Ile234 and Ile345) changes. This change results in a shift (or perturbation) of the respective methyl resonances in the ligand-bound state relative to the ligand-free state of the protein (**Figure 1.9 A**, compare black and magenta spectra).

On the contrary, methyl groups remote from the ligand binding site (**Figure 1.9 B**, e.g. Ile12 or Ile89) exhibit the same chemical environment in the ligand-bound and in the ligand-free protein. Consequently, the corresponding resonances in both states are the same (**Figure 1.9 A**, compare black and magenta spectra).

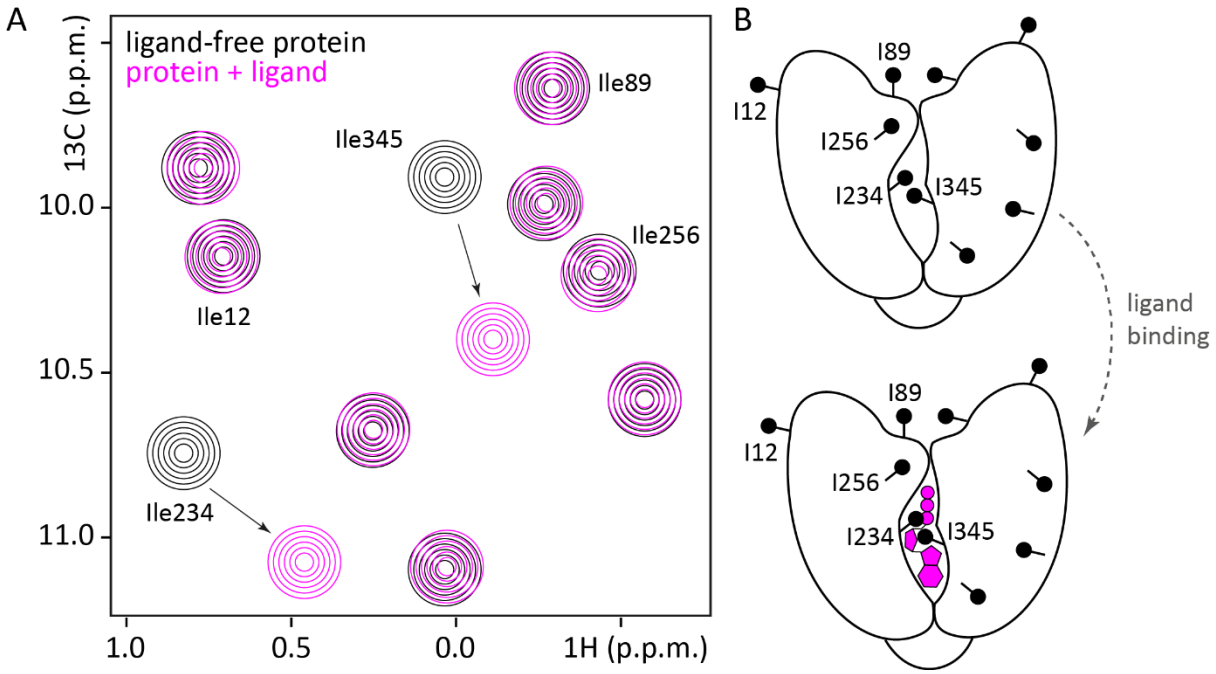

Figure 1.9 The principle of chemical shift perturbation (CSP) experiments. (A) Methyl groups in a hypothetical protein give rise to distinct resonances in a schematic ^{1}H-^{13}C-correlation spectrum (black) due to the unique chemical environment of each methyl group. Resonances of methyl groups that sense ligand binding undergo CSPs in the presence of a ligand (e.g. Ile234 and Ile345, magenta versus black), while resonances of methyl groups that are outside of the binding interface superpose perfectly in the apo (black) and in the ligand-bound state (magenta). **(B)** Schematic representation of a hypothetical protein, where a ligand (magenta shape) binds at an interface that comprises Ile234 and Ile345. All other Ile residues (black dots) are outside of the binding pocket.

CSP experiments are a versatile tool to study protein:ligand interactions since structural, thermodynamic and kinetic aspects of binding can be addressed. First, by comparing protein spectra in the absence and presence of a ligand, it is possible to determine if a ligand binds to a protein and in case resonance assignments are available, it is also possible to locate the ligand binding site. Second, analysis of CSP experiments that were recorded at different ligand concentrations enables the extraction of thermodynamic parameters including the ligand affinity (K_D). And third, determination of association (k_{on}), dissociation (k_{off}) and exchange rates ($k_{ex} = k_{on} + k_{off}$) between the ligand-free and ligand-bound state allow insights into the binding kinetics. The exchange rate between the ligand-free and the ligand-bound state (k_{ex}) as well as the chemical shift difference between these two states ($\Delta\omega$) determine the appearance of NMR resonances in a titration experiment. In the fast exchange regime ($k_{ex} >> \Delta\omega$), a ligand-sensitive resonance progresses with increasing ligand concentration in a linear direction from its position in the ligand-free state to its position in the saturated ligand-bound state. Notably, only one resonance is observed at every titration step for each ligand-sensing NMR probe.

Contrarily, if ligand binding is in slow exchange (k_{ex} << Δω), in every titration step two resonances per ligand-sensing probe are observed: one corresponds to the ligand-free and the other corresponds to the ligand-bound form. The intensities of both resonances then depend on the population of both states. Between the extreme cases of slow and fast exchange, intermediate exchange regimes show a more complex relation between changes in chemical shift and peak intensity. Thus, depending on the exchange regime of a protein:ligand interaction, resonances that report on ligand binding in CSP experiments are identified either by a peak shift or by a decrease in peak intensity (or a combination of both).

In this thesis, CSP experiments have been used to map binding sites for the Edc3 YjeF domain, RNA and the Dhh1 helicase on the Edc3 IDR (see CHAPTER 2), for the Edc3 IDR on the Edc3 YjeF domain (see CHAPTER 3) and to detect Dhh1 residues that are affected by LLPS (see CHAPTER 4). In the latter case, the "ligands" are other Dhh1 molecules.

1.4.6 *Methionine scanning*

The above-mentioned CSP experiments to investigate biomolecular interactions via methyl TROSY NMR spectroscopy rely on the natural occurrence of methyl-bearing side chains within the binding site. However, NMR is "blind" to areas devoid of appropriate probes and interaction surfaces may not harbor methyl groups in the quantity and distribution required for a detailed binding site mapping. To overcome this bottleneck, Stoffregen *et al.* proposed the methionine scanning approach[243], which extends the applicability, information content, and spatial resolution of standard methyl CSP experiments.

Methionine scanning involves the systematic, serial substitution of solvent-exposed residues with reporter methionines. This strategy is comparable to alanine-scanning[244,245], as it identifies functional epitopes via a mutagenesis-driven approach, but is highly advantageous in many ways. Since methionine is on average under-represented in proteins[246–248], the methionine region in a methyl TROSY spectrum is usually only sparsely populated and introduced methionines are instantaneously assigned by the appearance of a new resonance (**Figure 1.10**, bottom panels). Endogeneous methyl groups serve as internal probes for proper folding of methionine mutants (M456 in **Figure 1.10**). In addition, methionine labeling provides further advantages: First, the label can be stably incorporated during bacterial growth such that no post-translational modification is required. Second, label scrambling from methionine to other amino acids is not observed, which results in excellent incorporation rates[249,250].

Third, the high degree of rotational freedom on the long, unbranched methionine side chain results in very favorable NMR relaxation properties. This allows for the investigation of high-molecular weight complexes, unstable or dilute samples and can be combined with fast pulsing methods[249,250]. And fourth, methionine scanning can be applied to systems which are invisible in traditional HSQC-based backbone experiments or where backbone assignments are unavailable.

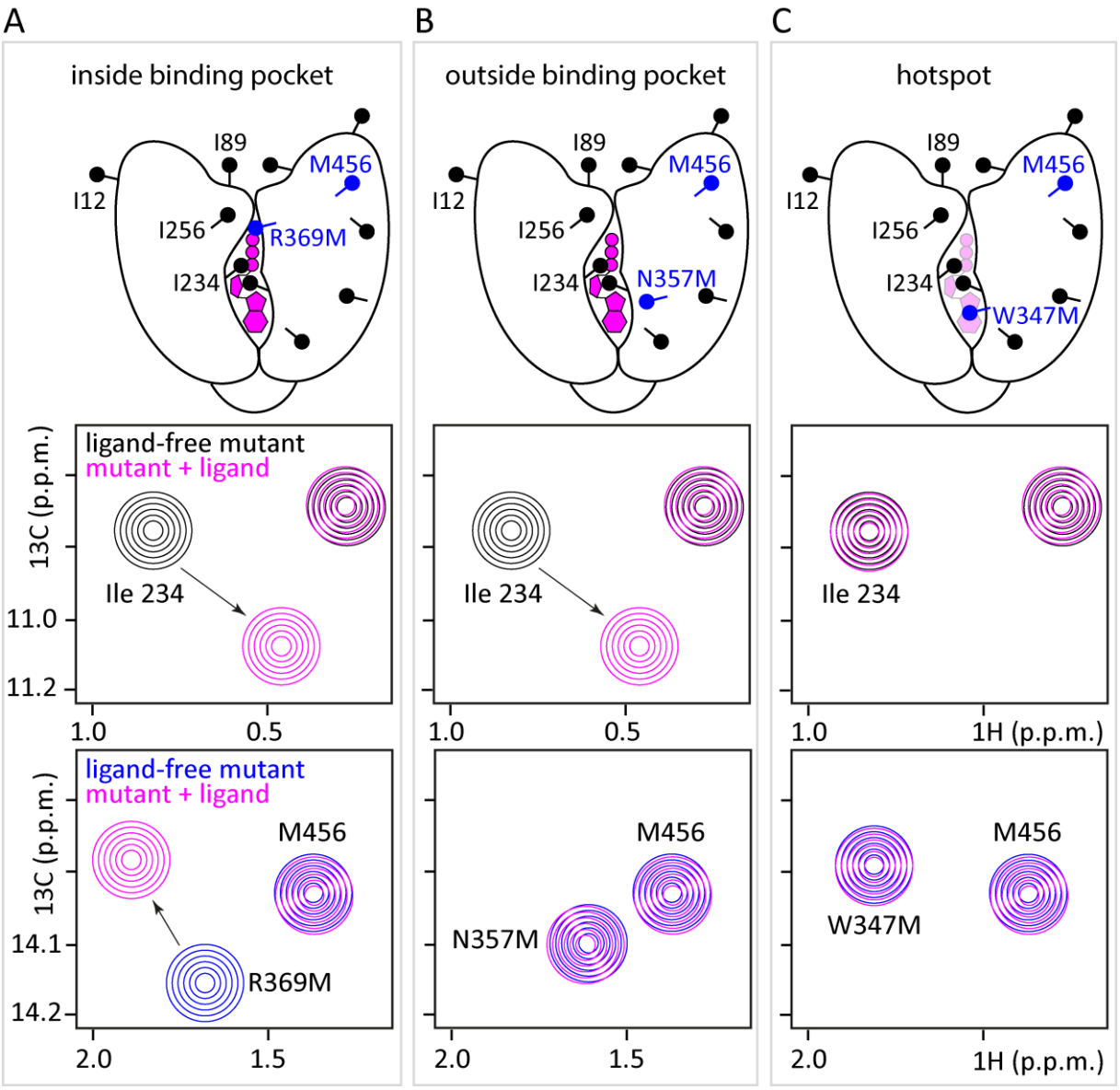

Figure 1.10: Schematic representation of the methionine scanning approach. (A) Top panel: The residue mutated to methionine (blue, R369M) is located inside the binding pocket. An endogenous methionine (blue, M456) reports on proper folding of the mutant. I234 is used as an internal reference for the ligand binding properties of the mutant protein. Middle and bottom panels: overlay of representative regions of ^{1}H-^{13}C correlation spectra in the presence (magenta) and absence (black and blue, respectively) of ligand (magenta shape in top panel). Binding of the internal reference is not affected by the mutation (middle panel). **(B)** The methionine mutant (blue, N357M) is located outside the binding pocket (top). No CSPs are observed for the introduced methionine (bottom), while the internal reference reports on ligand binding as in the WT protein (middle). **(C)** The methionine mutant (blue, W347M) impairs binding of the ligand (indicated by the transparent magenta shape in the top panel). In this case, no CSPs are observed for the resonances corresponding to the methionine mutant (bottom panel) and to the internal reference (middle panel).

The methionine scanning approach consists of two steps. First, residues of interest are mutated to methionine and the new methyl resonance is assigned (**Figure 1.10**, bottom panels). In a second step, the ligand of interest is added to the methionine mutant (**Figure 1.10**, compare magenta with black and blue spectra in middle and bottom panels, respectively). There are three possible outcomes of this CSP experiment. First, if CSPs are observed for the reporter methionine, this residue is inside the binding pocket (**Figure 1.10 A**). Second, if the reporter is outside of the interface there are no CSPs observed for the mutated methionine probe (**Figure 1.10 B**). The same is true for the third scenario when the mutation hits a "hotspot" residue whose mutation abolishes binding (**Figure 1.10 C**). To differentiate between the latter two options, CSPs of an intrinsic indicator of binding are analyzed (**Figure 1.10**, middle panels). The intrinsic indicator of binding (internal reference) is a methyl-bearing residue whose resonance is not perturbed due to the introduced mutation and which reports on ligand binding of the protein (I234 in **Figure 1.10**; see also **Figure 1.9**). If the methionine mutation is outside the binding pocket, the internal reference still shows CSPs upon ligand binding (**Figure 1.10 B**, middle panel), while in case the mutation hits a "hotspot", disruption of the binding interface by the introduced methionine also abrogates CSPs of the internal reference (**Figure 1.10 C**, middle panel)[243].

Applicability of the methionine scanning method was showcased for the interaction of the E3 ubiquitin ligase Rsp5 HECT-domain with its cognate E2 ubiquitin conjugating enzyme, Ubc4, resulting in a 62 kDa complex[243]. Mutations at 19 positions allowed a detailed binding site mapping, whereby five hotspots and eight more residues involved in binding have been identified. Importantly, methionine scanning is not restricted to protein:protein interfaces but can be extended to protein:RNA interaction and complexes far beyond 100 kDa as has been impressively demonstrated by Cvetkovic *et al.*[214] for the archaeal exosome complex, an exoribonuclease that is involved in 3´-5´ mRNA decay[77,251].

In this thesis, I applied a slightly modified version of the methionine scanning approach to map the binding surface of the Edc3 IDR on the dimeric Edc3 YjeF domain (see CHAPTER 3).

1.5 *AIMS OF THIS THESIS*

Liquid-liquid phase separation (LLPS) is increasingly recognized as a process involved in the cellular organization of proteins and RNA into membrane-less compartments. One of these compartments are so-called processing bodies (P-bodies) which are cytosolic foci containing translationally repressed mRNA as well as proteins and enzymes involved in mRNA degradation. The cellular function of P-bodies is a matter of debate, as some results point to P-bodies as sites for active mRNA degradation, while other studies associate them with mRNA storage. In this thesis, I aim at elucidating the cellular role of P-bodies by performing enzymatic assays under phase separation and non-phase separation conditions.

In vivo and *in vitro* LLPS processes are often studied using fluorescence microscopy techniques. However, we observe some disadvantages when using light microscopy to study *in vitro* LLPS of proteins involved in P-body formation. We found that light microscopy experiments required high amounts of protein and RNA and that they are extremely time-consuming. One aim of this thesis was to develop a robust assay that allows for the rapid and reliable analysis of *in vitro* LLPS. With this assay at hand, I analyzed the molecular mechanisms that underly LLPS and the formation of *in vitro* P-bodies. Importantly, I aimed for a more quantitative analysis of the molecular contributions that lead to LLPS relative to what is achieved with microscopy-based approaches.

I focused on the conserved Edc3 and Dhh1 proteins that are tightly embedded in the mRNA degradation machinery. Edc3 is a scaffolding protein that interacts with the decapping enzyme Dcp2 and other decapping factors to undergo phase separation, while Dhh1 is a DEAD-box RNA helicase that has been reported to undergo *in vitro* LLPS in the presence of ATP and RNA. Although both proteins are important for P-body formation and some interactions facilitating LLPS of these proteins have been described, little is known about the intra- and intermolecular interactions at an atomic level. I aimed to shed light on these interactions by employing NMR spectroscopy that enables the detection of interactions with atomic resolution.

Many intrinsically disordered proteins that undergo LLPS are observed to form insoluble aggregates over time. These aggregates were found to contain amyloid-like fibers which are associated with neurodegenerative diseases. Here, I aim to investigate if Edc3 liquid droplets also undergo a second phase transition from a liquid-like to a solid-like state and which interactions mediate this maturation process. As structural data on phase-separated proteins that contain folded domains is scarce, we aim to reveal structural features of Edc3 after LLPS in collaboration with the group of Marc Baldus at Utrecht University (The Netherlands) by combining solid-state NMR spectroscopy with state-of-the-art labeling schemes and solution-state methyl TROSY NMR techniques.

CHAPTER 2 *A synergistic network of interactions promotes the formation of in vitro processing bodies and protects mRNA against decapping*

This chapter was published in Nucleic Acids Research[252]. Large parts of the text and figures are thus identical to the published manuscript. Remco Sprangers conceived the project. Experiments were designed by all authors. Stefan Schütz and Erik Nöldeke performed molecular cloning, protein expression and purification. E.N. established fluorescence microscopy and small scale, high-throughput liquid-liquid phase separation (LLPS) assays. S.S. and E.N. performed fluorescence microscopy experiments and LLPS assays. S.S. prepared samples for NMR spectroscopy and acquired NMR data with help from R.S. Isothermal titration calorimetry (ITC), RNA *in vitro* transcription and purification were done by S.S., who also performed RNA protection and decapping assays. All authors analyzed and interpreted data. R.S. wrote the paper with contributions from S.S. All authors commented on the manuscript.

2.1 *INTRODUCTION*

To facilitate and regulate biological reactions the intracellular space is partitioned into distinct compartments[253]. These compartments can be divided in those that are encapsulated in a lipid membrane and those that are devoid of a lipid membrane[125,254]. Membrane-less compartments, also referred to as intracellular bodies or granules, are dynamic cellular sub-structures that arise from a spontaneous liquid-liquid phase separation (LLPS) process, which results in a very high local concentration of specific components[119]. Both the nucleus and the cytoplasm of eukaryotic cells contain a number of functionally distinct granules, each of them typically harboring tens to hundreds of specific proteins and RNA. Examples of nuclear ribonucleoprotein (RNP) bodies include Cajal bodies, nucleoli and PML bodies, whereas cytoplasmic RNPs include stress granules, germ granules and processing bodies (P-bodies)[125]. The large number of RNA granules reflects the involvement of these assemblies in central cellular processes including mRNA translation, processing, localization, and turnover. Membrane-less compartments display a liquid-like behavior, where smaller granules can fuse together into larger assemblies[116,117,126]. The viscosity inside these cellular droplets has been determined to be three to six orders of magnitude above that of pure water, which can impact on the enzymatic processes taking place inside these foci. Interestingly, it has been shown that functionally different RNP granules, like P-bodies and stress granules, share certain components and that these foci can fuse or dock together[132,255].

Several *in vivo* and *in vitro* approaches have revealed insights into the molecular details that lead to cellular phase transition. Currently, three mechanisms have been shown to be important for the self-assembly process. First, intrinsically disordered regions (IDRs) are highly abundant in proteins that segregate into cellular foci[256]. These IDRs are characterized by the lack of tertiary structure and are sparse in hydrophobic amino acids. Indeed, interactions between the IDRs in DDX4, have been shown to be important for the self-assembly of the protein into germ granules[120]. In addition, an IDR in hnRNPA1 is important for the phase separation of the protein into stress granules[123,151], IDRs in LAF-1 have been shown to promote P-granule assembly[124] and a glutamine/ asparagine-rich domain in yeast LSm4 stimulates P-body formation[144]. The IDRs in these proteins are able to associate into large networks due to a combination of charge-charge, cation-pi, dipole-dipole and pi-pi stacking interactions[119]. The distribution of charged residues has been shown to be important for self-assembly, where an unequal distribution of charges stimulates intermolecular contacts[120,143]. Secondly, interactions between folded protein domains and short linear motifs in disordered regions are found in interaction networks that result in phase separations. Examples include the interactions between SH3 domains and proline-rich motifs[134] and those between the Edc3 LSm domain and helical leucine-rich motifs (HLMs) in Dcp2 and Pdc1[101,103] (**Figure 2.1 A**). Finally, interactions between folded RNA recognition motifs (RRMs) in granule proteins and RNA stimulate the formation of mRNP granules, which has been shown for hnRNPA1[123], PTB[134] and Whi3[122]. In addition, non-translating mRNAs have been shown to increase the number of cellular P-bodies[114]. On the other hand, RNA does not contribute to the phase separation of the P-granule protein LAF-1[124], indicating that RNA is not a general component important for RNP formation. Importantly, it has been shown that different intermolecular interaction modes can simultaneously take place within phase-separated granules[103,143], illustrating the high level of redundancy within LLPS processes.

The integrity of cellular granules is regulated and cellular foci can be disassembled by a number of processes, including phosphorylation by the DYRK3 kinase[139], the activity of the Hsp70 and Hsp40 chaperones[257], through autophagy and Cdc48/VCP function[163] and the activity of the helicase Dhh1[164]. As the cellular environment appears to be close to the phase separation boundary, small perturbations to the interaction network that underlies granule formation allow for a rapid formation and disassembly of cellular foci. Proteins in cellular foci can undergo further transitions to form fiber-like structures. Interestingly, the mechanisms that result in cellular phase transitions and those that are important for fibrillization are different[123]. The formation of fibers appears to be enhanced in cellular foci due to the high local protein concentration, which can result in pathological inclusions[115,121,151,258].

Processing bodies are dynamic cytoplasmic RNP granules that contain proteins that are involved in translational repression and mRNA degradation[114]. The main constituents are mRNA[259] and the mRNA decay machinery, including the Dcp1:Dcp2 mRNA decapping complex, the enhancer of decapping 3 (Edc3), the RNA helicase Dhh1, the Pat-Lsm1-7 complex and the exonuclease Xrn1[89,113,144,260]. The composition and the mechanism of assembly of P-bodies varies between different organisms, as an example, the protein Pdc1[261] plays a central scaffolding role in *Schizosaccharomyces pombe (S. pombe)*, but this factor appears to be absent in *Saccharomyces cerevisiae (S. cerevisiae)*. Despite the ubiquitous presence of P-bodies in eukaryotes their biological role remains a matter of debate. Based on the protein composition, they have been implicated in mRNA decapping and degradation. However, mRNA can leave processing bodies to re-engage in translation[262], which points towards a role of P-bodies in the temporary storage of non-translating mRNAs, e.g. during cellular stress conditions.

Here, we address the cooperative intermolecular interactions in the formation of processing bodies and maturation of these foci into a gel-like phase. To that end, we exploit an automated, fast and reliable *in vitro* bottom-up approach[103,134] using purified processing body components (Dcp1, Dcp2, Pdc1, Edc3, Dhh1 (Ste13) and RNA) and assess their ability to undergo LLPS and subsequent formation of gel-like structures. Using mRNA enzymatic assays, we reveal the first insights into the activity of the mRNA decapping complex within processing body-like foci that support a role for processing bodies in the storage of translationally inactive mRNA.

2.2 *MATERIALS AND METHODS*

2.2.1 *Protein expression and purification*

The genes for full-length Pdc1 and Edc3 proteins as well as for the isolated Edc3 LSm and YjeF_N domains and for the first RecA-like domain of Ste13 from *S. pombe* were cloned into modified pET vectors that carried an N-terminal TEV-cleavable His_6-tag. The genes for the intrinsically disordered region (IDR) of Edc3 and the second RecA domain of Ste13 (Dhh1 in *S. cerevisiae*) were cloned into modified pET vectors that carried an N-terminal TEV-cleavable His_6-GST- or His_6-GB1-tag, respectively. Dcp1:Dcp2 complexes were cloned in a modified pET vector, where only the Dcp1 protein carried an N-terminal TEV- cleavable His_6-tag. Point mutations, insertions and deletions were introduced into the genes using standard site-directed mutagenesis methods (**Table 2.1**).

Table 2.1: Protein constructs used in this study

Protein / complex	Residues	Purification tag	Internal reference
Edc3	1-454	N-His_6-TEV	#180
Edc3	1-71 (LSm)	N-His_6-TEV	#1380
Edc3	72-194 (IDR)	N-His_6-GST-TEV	#1293
Edc3	195-454 (YjeF_N)	N-His_6-TEV	#196
Edc3	1-194 (ΔYjeF_N)	N-His_6-TEV	#244
Edc3	1-454, Δ72-194 (ΔIDR)	N-His_6-TEV	#1291
Edc3	72-454 (ΔLSm)	N-His_6-TEV	#1301
Edc3	72-194, Δ90-110 (IDRΔ1)	N-His_6-GST-TEV	#1514
Edc3	72-194, Δ158-171 (IDRΔ2)	N-His_6-GST-TEV	#1512
Edc3	72-194, Δ90-110 + Δ158-171 (IDRΔ12)	N-His_6-GST-TEV	#1516
Edc3	72-194, F108A (FDK-to-ADK)	N-His_6-GST-TEV	#1688
Edc3	72-194, F99A, F101A (FDF-to-ADA)	N-His_6-GST-TEV	#1689
Dcp1:Dcp2 ΔMid	1-127 (Dcp1), 1-741 Δ290-554 (Dcp2)	N-His_6-TEV on Dcp1	#264
Pdc1	1-105 (Pdc1) – GST – 880-1076 (Pdc1)	N-His_6-TEV	#928
Ste13	43-249 (RecA1)	N-His_6-TEV	#183
Ste13	250-421 (RecA2)	N-His_6-GB1-TEV	#1675

Chemical competent *E. coli* BL21 (DE3) Codon Plus RIL (Stratagene) cells were transformed with the appropriate plasmid and grown at 37 °C to an OD_{595} of 0.8 in lysogeny broth (LB) medium, after which protein overexpression was induced with 0.5 mM IPTG at 20 °C. After expression overnight, cells were pelleted by centrifugation and lysed in buffer A (25 mM sodium phosphate, pH 7.4, 250 mM NaCl, 1 mM DTT) complemented with 10 mM imidazole, lysozyme, 0.1 % Triton X-100, 5 mM $MgCl_2$ and 0.2 U/mL DNase I. The cell lysate was cleared from insoluble debris by centrifugation and the supernatant was loaded on Ni^{2+}-nitrilotriacetic acid (NTA) resin. The resin was washed with buffer A that was complemented with 10 mM imidazole. The protein bound to the resin was eluted with buffer A complemented with 300 mM imidazole. 1 mg TEV protease was added to the eluted proteins to cleave the purification tag from the target protein.

Proteins carrying an N-terminal TEV cleavable GST- or GB1-tag were dialyzed overnight at 4 °C into buffer B (25 mM HEPES, 50 mM NaCl, 1 mM DTT) supplemented with 0.5 mM EDTA. The dialyzed proteins were applied to a HiTrap 5mL SP FF column (GE Healthcare) to separate the protein of interest from the cleaved solubility-tags using appropriate gradients from buffer B to buffer C (25 mM HEPES, 1 M NaCl, 1 mM DTT) over 100 mL. The pH of the ion exchange buffers varied depending on the used solubility tag (pH 7.3 for GB1 and pH 8.0 for GST). The buffer of the pooled fractions was exchanged to SEC buffer (25 mM HEPES, pH 7.3, 125 mM NaCl, 1 mM DTT) using centrifugal filters.

Proteins carrying an N-terminal His_6-tag were dialyzed overnight at 4 °C into SEC buffer supplemented with 0.5 mM EDTA. Purification to homogeneity was achieved by size exclusion chromatography on Superdex 75 or Superdex 200 columns (GE Healthcare) in SEC buffer.

Labeling of proteins with NMR-active nuclei was achieved by over-expression of the gene in M9 minimal medium. The growth medium was supplemented with 0.5 g/L $^{15}NH_4Cl$ as the sole nitrogen source and either 2 g/L $^{1}H^{13}C$-glucose for resonance assignment purposes or 4 g/L $^{1}H^{12}C$-glucose for NMR binding experiments.

2.2.2 *Protein fluorescence labeling*

For fluorescence microscopy studies that assess droplet maturation, the Edc3 protein was labeled with Oregon Green 488 (OG). Therefore, the protein was prepared in SEC buffer free of any reducing agent. A few micro grams of OG-maleimide were dissolved in DMSO and added in a two- to three-fold excess to the protein solution. The labeling reaction was allowed to proceed for up to two hours at room temperature in the dark. The reaction was stopped by addition of DTT to a final concentration of 4 mM. Unreacted dye was removed using a PD10 desalting column (GE Healthcare) in SEC buffer. The concentration of the labeled protein was determined photometrically at 280 and 492 nm, whereby the lower concentration was used for further calculations. The fluorescently labeled protein was mixed with non-labeled protein to obtain an appropriate signal-to-noise ratio in the fluorescence microscopy experiments.

2.2.3 *RNA in vitro transcription, purification and capping*

RNA was prepared using *in vitro* transcription with in-house purified T7 polymerase[263]. The DNA template was obtained from one of two different sources (**Table 2.2**). In the first case, two DNA primers were mixed, where the forward primer encodes for the T7 promoter and the reverse primer encodes for the desired RNA sequence plus the T7 promoter. In the second case, the DNA template was cloned into a vector that coded for the target RNA sequence followed by a 3´ HDV ribozyme that cleaves at the end of the target RNA sequence and results in a 3´ cyclic phosphate.

In vitro transcribed RNA was purified under denaturing conditions at 80 °C using anion exchange chromatography with a DNAPac PA100 column (22 x 250 mm, Dionex) (Buffer D: 20 mM Tris, pH 8, 5 M urea; Buffer E: as D, supplemented with 2 M NaCl)[99,263,264]. The target RNA was precipitated from the pooled fractions by addition of 0.7 volumes isopropanol and incubation at -20 °C. The precipitate was collected by centrifugation, washed with 70 % ethanol and resuspended in water.

Complete desalting was achieved using a PD10 column, followed by vacuum concentration to obtain pure dry RNA product. RNA was resuspended at concentrations required for subsequent experiments. The quality of the RNA was assessed using urea-PAGE with 1x TBE (89 mM Tris, pH 8.0, 89 mM boric acid, 2 mM EDTA) as a running buffer. RNA was visualized by methylene blue staining.

5mer RNA was produced from a purified GA-containing 30mer RNA with a single uracil nucleotide at position 5 that was subjected to preparative RNase A cleavage[263]. After phenol-chloroform extraction, the different RNA species were separated by anion exchange chromatography as described above. Fractions containing the 5mer RNA were pooled and supplemented with 3.5 volumes ethanol and 0.2 M NaCl. After incubation at -20 °C for at least 12 hours, the precipitate was treated further as described above.

Capped RNA was prepared as previously published by Fuchs *et al.*[263]. For a 21mer RNA the capping efficiency was directly analyzed by Urea-PAGE, while a 100mer RNA (single uridine at position 15) was treated with RNase A before gel analysis as described[263]. For labeling of RNA with a fluorescent dye, a GA-containing 30mer RNA with a single uridine at position 15 was transcribed in the presence of 4-thiouridine triphosphate instead of UTP. The transcript was purified and capped as described above. Fluorescence labeling was achieved using 5-(iodoacetamido)-fluorescein following published protocols[99,265].

Table 2.2: RNA used in this study

RNA	**Properties**	**Sequence**	**Internal reference**
15mer	GA-only	GGAGAAGAGAAGGAG	#4, primer
21mer	GA-only	GGAAGGAGAGGAAGGAAAGGA	#40, plasmid
30U5mer	GA-only, but single U at position 5	GGAGUGAGAGGAAGGAAGGGAAGAAAGAAG	#8, primer
30U10mer	GA-only, but single U at position 10	GGAGGAGAGUGAAGGAAGGGAAGAAAGAAG	#9, primer
30U15mer	GA-only, but single U at position 15	GGAGGAGAGGAAGGUAAGGGAAGAAAGAAG	#10, primer
63mer	GA-only	GGAAGGAGAGGAAGGAAGGGAAGAAAGAAGAGGA GAGGAAGGAAGGGAAGAAAGAAGGGAAGA	#5, primer
100U15mer	GA-only, but single U at position 15	GGGAAGGAAGGGAAUAAAGAAGGGAAGAGGAAGG AGAGGAGGGAAGAAAGAAGAGGAGAGGAAGGAAG GGAAGAAAGAAGAGGGAAGAGGAAGGAGAGGA	#6 primer

2.2.4 *Liquid-liquid phase separation experiments*

For *in vitro* phase separation experiments, samples were prepared in glass-bottom 1536-well plates (Greiner). Stock solutions of RNA, DNA, proteins and different buffer conditions were prepared in 96-well V-bottom plates (Greiner). A Mosquito pipetting robot (TTP Labtech) was used for transferring liquids from the 96-well source plates to the 1536-well assay plates. The general pipetting order was (1) buffer, (2) RNA or DNA and (3) proteins (Dcp1:Dcp2, Pdc1, Edc3). Automated pipetting routines were applied for a thoroughly mixing of the samples prior to analysis. The total sample volume in the 1536-well plates was 5 µL. Unless indicated otherwise, the salt concentration used was 125 mM.

1536-well plates were analyzed immediately after preparation in a Synergy Mx plate reader (BioTek) at 30 °C. To avoid droplet maturation during prolonged pipetting routines, only four out of 48 columns of the 1536-well plates were prepared in parallel. The absorption at 630 nm (as a good estimate of the OD_{600}) was used as a quantitative measure for phase separation.

For LLPS maturation experiments, Edc3 samples were prepared at 150 µM in 25 mM HEPES, pH 7.3 and 20, 30, 40, 50, 60, 70, 80, 90, 100 or 125 mM NaCl. At each time-point, the solution was mixed by pipetting and the absorption at 600 nm (OD_{600}) was measured. Subsequently, 8 µL of the sample were collected and mixed with 2 µL 2.5 M NaCl to dissolve liquid-like droplets. The protein concentration of this clear solution was measured, and the concentration of the sample was calculated, thereby considering the dilution factor. The fraction of insoluble protein (gel) was determined as the loss in the protein concentration.

2.2.5 *Microscopy*

Fluorescent droplets were observed at room temperature with a Zeiss Axio Observer Z1 equipped with an AxioCam MRM imaging system. Micrographs were processed with Fiji[266].

2.2.6 *NMR*

All NMR samples were in SEC buffer and contained 5 % D_2O. NMR spectra were recorded at 283 K on Bruker AVIII-600 and AVIII-800 spectrometers with room temperature probe-heads. Backbone and side chain resonances were assigned using HSQC-based HNCACB, HN(CO)CACB, HNCACO, HNCO and CC(CO)NH experiments[216]. NMR titration experiments were carried out with 0.01 to 0.1 mM ^{15}N-labeled protein (Edc3 IDR) and a 4-fold excess of unlabeled protein (Edc3 YjeF_N or Ste13 RecA2 domain) or an equimolar amount of RNA or DNA. NMR spectra were processed using the NMRPipe/NMRDraw software suite[267]. Figures displaying NMR spectra were prepared using NMRview (onemoonscientific.com).

2.2.7 *Decapping assays*

RNA decapping assays were performed as triplicates at 30 °C in 20 µL SEC buffer (supplemented with 0.1 mM EDTA) containing 1 µM Dcp1:Dcp2, 20 µM capped RNA and 15 µM BSA. The reaction mixture was complemented with 15 µM of either full-length Edc3, Edc3 LSm, Edc3 IDR or Edc3 YjeF_N or 15 µM of each Edc3 LSm, IDR and YjeF_N. The decapping reaction was started by addition of $MgCl_2$ to a final concentration of 5 mM. At different time-points 4 µL samples were collected and the reaction was quenched by addition of 26 µL stop solution (10 mM EDTA, pH 8.0, 500 mM NaCl) and 15 µL phenol-chloroform-isoamylalcohol. Depending on the length of the used RNA, the upper aqueous RNA-containing phase was either directly subjected to HPLC analysis or was treated with RNase A first, followed by another phenol-chloroform extraction step to remove the enzyme before HPLC analysis.

2.2.8 *HPLC analysis*

10 µL of the quenched reaction were automatically injected onto an analytical DNAPac PA200 RS column (4.6 x 250 mM, Dionex). Capped and decapped RNA were separated by anion exchange HPLC at 50 °C using appropriate gradients (buffer F: 25 mM Tris, pH 8; buffer G: as buffer F, supplemented with 1.5 M NaCl) at a flow rate of 0.45 mL/min. Eluting RNA was detected using the absorption at 260 nm (**Figure 2.4 D**).

2.2.9 *Analysis of degradation data*

Substrate and product peaks were integrated and corrected for different absorption coefficients of capped and decapped RNA to calculate the ratio of the two RNA species in the sample. The fractions of decapped RNA at different time-points were fitted to a first order curve using in-house written Matlab scripts. Based on the known concentrations of Dcp2 and RNA in the reaction mixture, the decapping rates (kcat in min^{-1}) were derived.

2.2.10 *RNase A protection assays*

RNase A protection assays were performed at 37 °C in 50 µL SEC buffer containing 25 µM of a 30mer GA-containing RNA with a single uracil nucleotide at position 10. The reaction was complemented with 15 µM Dcp1:Dcp2 and/or 30 µM full-length Edc3 or 30 µM of each Edc3 LSm, IDR and YjeF_N. A sample that only contained RNA, but no additional proteins, served as a control. The reaction was started by addition of 10 nmol RNase A per nmol RNA.

After ten minutes, 4 µL 5M NaCl were added to dissolve any liquid-like droplets and the reaction was quenched by addition of 54 µL phenol-chloroform-isoamylalcohol. 10 µL of the upper aqueous RNA-containing phase were mixed with an equal volume of 2x RNA loading dye. RNA species were separated on a 15 % urea-polyacrylamide gel and band intensities were analyzed with Fiji[266] to determine the fraction of cleaved RNA. A protection factor was calculated by dividing the fraction of cleaved RNA for each sample by the fraction of cleaved RNA for the control sample.

2.3 *RESULTS*

2.3.1 *High throughput approach for LLPS determination*

In previous work[103], we studied the LLPS process of the Dcp1:Dcp2 decapping complex and the Edc3 adaptor protein (**Figure 2.1 A**). In those experiments purified proteins were manually mixed in 96-well plates and the presence of phase separations was assessed using a wide field microscope. For each condition we used 80 µL protein solution and for a small phase diagram with 30 conditions we required tens of mg protein. Due to the long time it takes to manually prepare these grids and due to the high amount of required protein this approach is not feasible for larger LLPS screens. To be able to probe LLPS in a fast, reliable and high throughput manner, we reduced the volume per condition by a factor of 16. To pipette small volumes, we used a pipetting robot and 1536-well plates. Importantly, this process also increased the speed and accuracy of the pipetting procedure.

After mixing the individual protein components, the LLPS process results in the formation of droplets that fuse over time and then settle on the bottom of the well as a homogeneous layer. In our previous setup, it took on the order of hours to record stacked images of 30 conditions, during which the appearance of the droplets gradually changed. On a macroscopic level, LLPS is visible as a turbidity of the solution[134]. In our new setup, we measure this turbidity directly in a plate reader. Importantly, this turbidity correlates directly with the amount of LLPS as judged from manually interpreted microscopy images (**Figure 2.1 B**, **Figure 2.2**). From that we conclude that the turbidity provides an unbiased and quantitative measure of the LLPS process. The determination of the turbidity of hundreds of conditions takes under a minute in a plate reader, which ensures that all conditions are assessed at the same stage. In summary, we here established a fast method by which large phase separation diagrams can be reliably and quantitatively measured with minimal amounts of material. To test our novel strategy, we assessed how salt influences the LLPS potential of Dcp1:Dcp2 and Edc3.

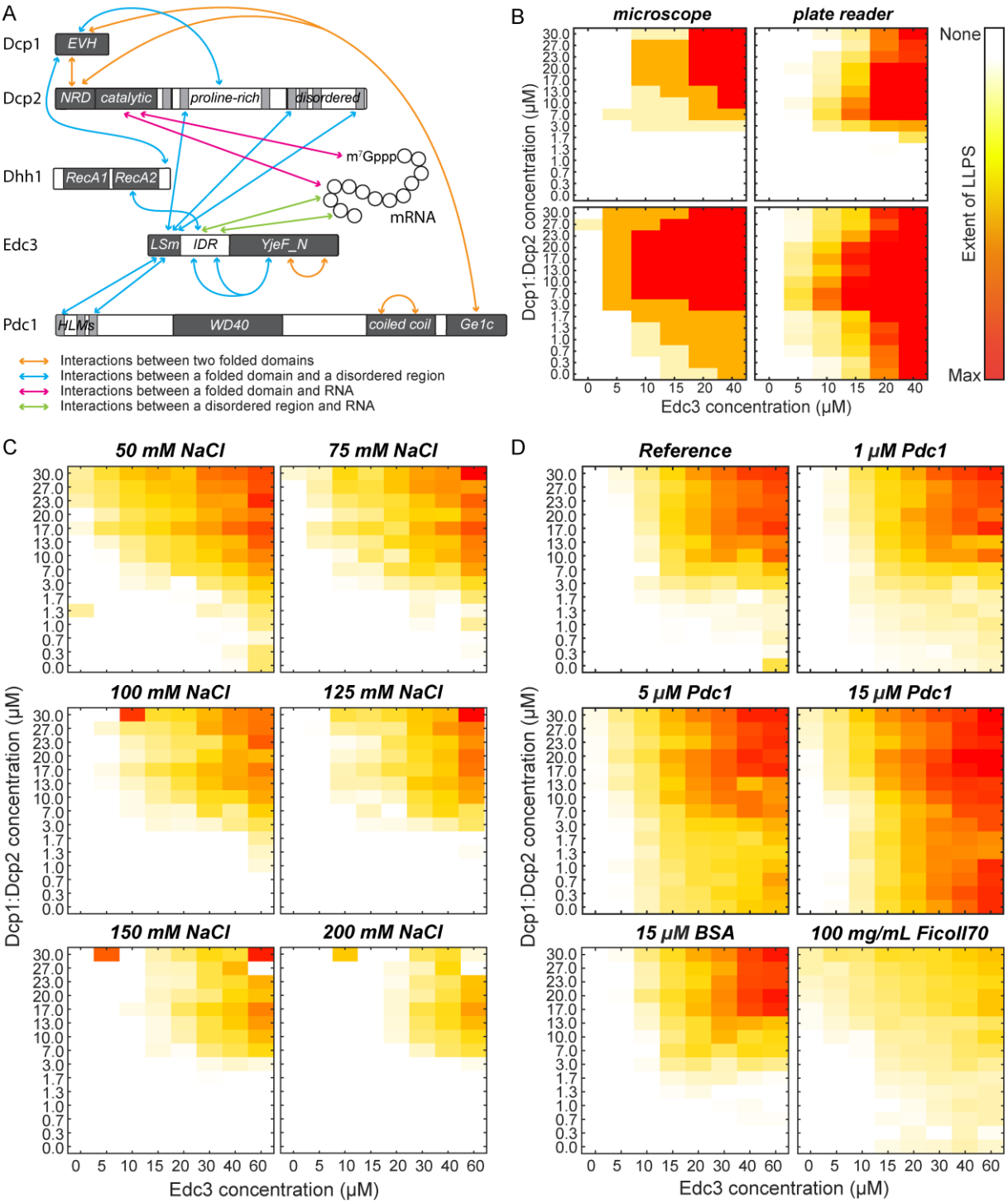

Figure 2.1: High throughput LLPS diagrams. (A) Schematic diagram of the proteins used in this study. Lines indicate the intra- and intermolecular interactions that are important for the LLPS process. **(B)** Comparison of LLPS diagrams for the Dcp1:Dcp2 (y-axis) and Edc3 (x-axis) protein that are constructed based on the visual inspection of microscopy images (left) or direct and quantitative turbidity measurements (right). Both methods provide highly similar phase diagrams, underlining that our high throughput approach provides rapid and reliable data. The degree of phase separation is indicated in a white (no LLPS), yellow, orange, red (high degree of LLPS) color scale. See also **Figure 2.2**. **(C)** The effect of salt on the LLPS process. Salt inhibits LLPS of Dcp1:Dcp2:Edc3. Note the few outliers in the phase diagrams (at e.g. 150 mM salt, 5 μM Edc3 and 30 μM Dcp1:Dcp2) are likely due to the presence of small air bubbles in these conditions. **(D)** Addition of Pdc1 significantly enhances the LLPS of Dcp1:Dcp2 and Edc3. The effect is due to specific interactions between Pdc1 and Dcp1:Dcp2:Edc3 and is not due to indirect effects of e.g. molecular crowding, as BSA (bottom left) has no effect on the LLPS process. A very high amount of the crowding agent Ficoll70 (bottom right) only has a minor effect on the Dcp1:Dcp2:Edc3 phase separation process

Previous studies have shown that salt can either inhibit[120,124,151] or enhance[268,269] the LLPS process, indicating the formation of cellular foci does not always follow the same molecular principles. Here, we used phase separation diagrams with 15x8 conditions and prepared six grids with salt concentrations between 50 and 200 mM (720 unique conditions; **Figure 2.1 C**). These data clearly show that an increase in the salt concentration results in a gradual shift of the phase separation boundary towards higher protein concentrations. From these data we conclude that electrostatic interactions play a role in the clustering of Dcp1:Dcp2 and Edc3. In addition, these data reveal that the increase in the number of *in vivo* cytosolic processing bodies upon increased extra-cellular salt concentrations[259] is not a direct effect.

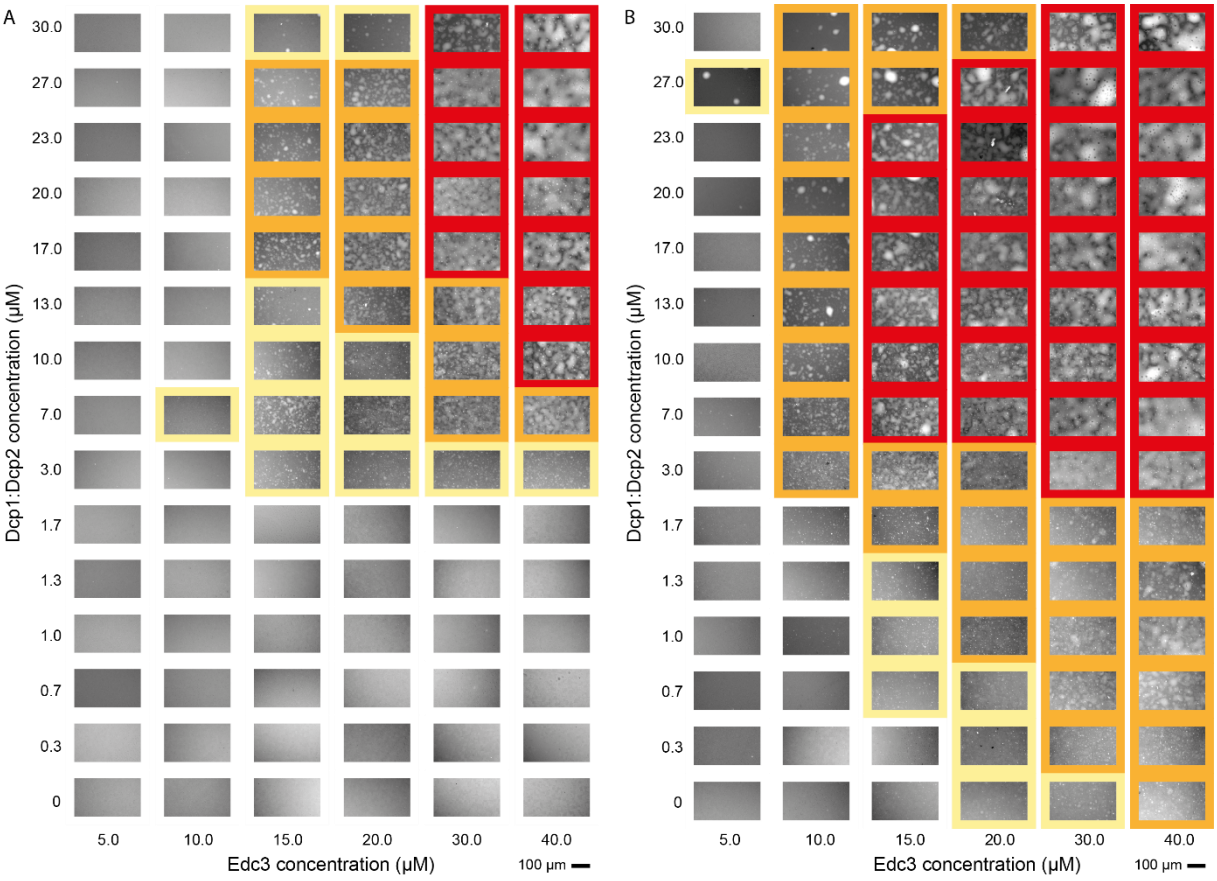

Figure 2.2: Interpretation of LLPS diagram based on fluorescence microscopy images. Phase separation diagram of Dcp1:Dcp2 and Edc3 in the absence **(A)** and presence **(B)** of 15 µM Pdc1. The degree of phase separation was visually classified as strong (red), intermediate (orange) or weak (yellow). See also **Figure 2.1**.

2.3.2 *Pdc1 stimulates LLPS*

In vitro, LLPS has only been addressed for systems with a limited number (two to three) of components. In a cellular setting, LLPS foci, including processing bodies, contain tens of proteins and nucleic acids. Based on our novel high throughput approach, we set out to assess the effect the addition of proteins has on the phase diagram boundaries.

First, we complemented the Dcp1:Dcp2:Edc3 *in vitro* system[103] with the processing body protein Pdc1[261] (**Figure 2.1 A**) and observe that the presence of Pdc1 reduces the amounts of Dcp1:Dcp2 and Edc3 that are required to induce phase separations (**Figure 2.1 D**). In the extreme case, the absence of Dcp1:Dcp2 can be compensated for by the presence of Pdc1, clearly illustrating the redundancy of the interactions that lead to phase separation and processing body formation. Importantly, the addition of the unrelated protein BSA does not influence the phase separation boundary, indicating that the Pdc1 effect is not due to molecular crowding (**Figure 2.1 D**). In agreement with that, also the addition of a crowding agent has no significant effect on the LLPS boundaries (**Figure 2.1 D**), although crowding agents are reported to increase LLPS in other systems[123,151]. As mentioned above in the salt-dependent LLPS experiments, there thus appears to be no general rules that can describe the mechanisms underlying the spontaneous clustering of proteins.

2.3.3 *RNA strongly stimulates LLPS*

In a second set of experiments, we investigated the effect of RNA on the phase separation diagram. To that end, we complemented Dcp1:Dcp2:Edc3 grids with 5 µM RNA of increasing length (**Figure 2.3 A**). Interestingly, we observe that a small RNA with 15 nucleotides has no effect on the phase diagrams. This is most likely due to the inability of this RNA to cross-link between different proteins. Longer RNAs, especially RNA with 30 or more nucleotides, have a strong impact on the phase diagrams and significantly shift those towards lower Edc3 and Dcp1:Dcp2 concentrations. Remarkably, we observe that Edc3 alone as well as Dcp1:Dcp2 alone can undergo efficient phase separation in the presence of RNA (e.g. **Figure 2.3 A**, bottom middle panel at zero Edc3 or Dcp1:Dcp2 concentrations). As for Pdc1, this underlines the redundancy in the phase separation systems. Noteworthy, the addition of DNA does not have an influence on the LLPS process, indicating that the Dcp1:Dcp2:Edc3 system is selective for RNA.

In the experiments above, we kept the molar RNA concentration constant and increased the length of the oligonucleotide. To directly address how the length of the RNA influences the phase separation process, we performed complementary experiments where we kept the number of nucleotides constant. In those experiments, we added a higher molar amount of a short RNA and a lower molar amount of the long RNAs. These experiments show that RNA substrates of 15 nucleotides or less have no or only a weak effect on the phase diagrams, whereas substrates of 30 or more nucleotides have a similarly strong effect (**Figure 2.3 B**). This shows that three 30mer RNAs have the same effect as one 100mer RNA, which provides strong evidence that the individual protein:RNA interactions are independent. In addition, our data indicate that the RNA length that is required to efficiently link into the interaction network is around 30 nucleotides.

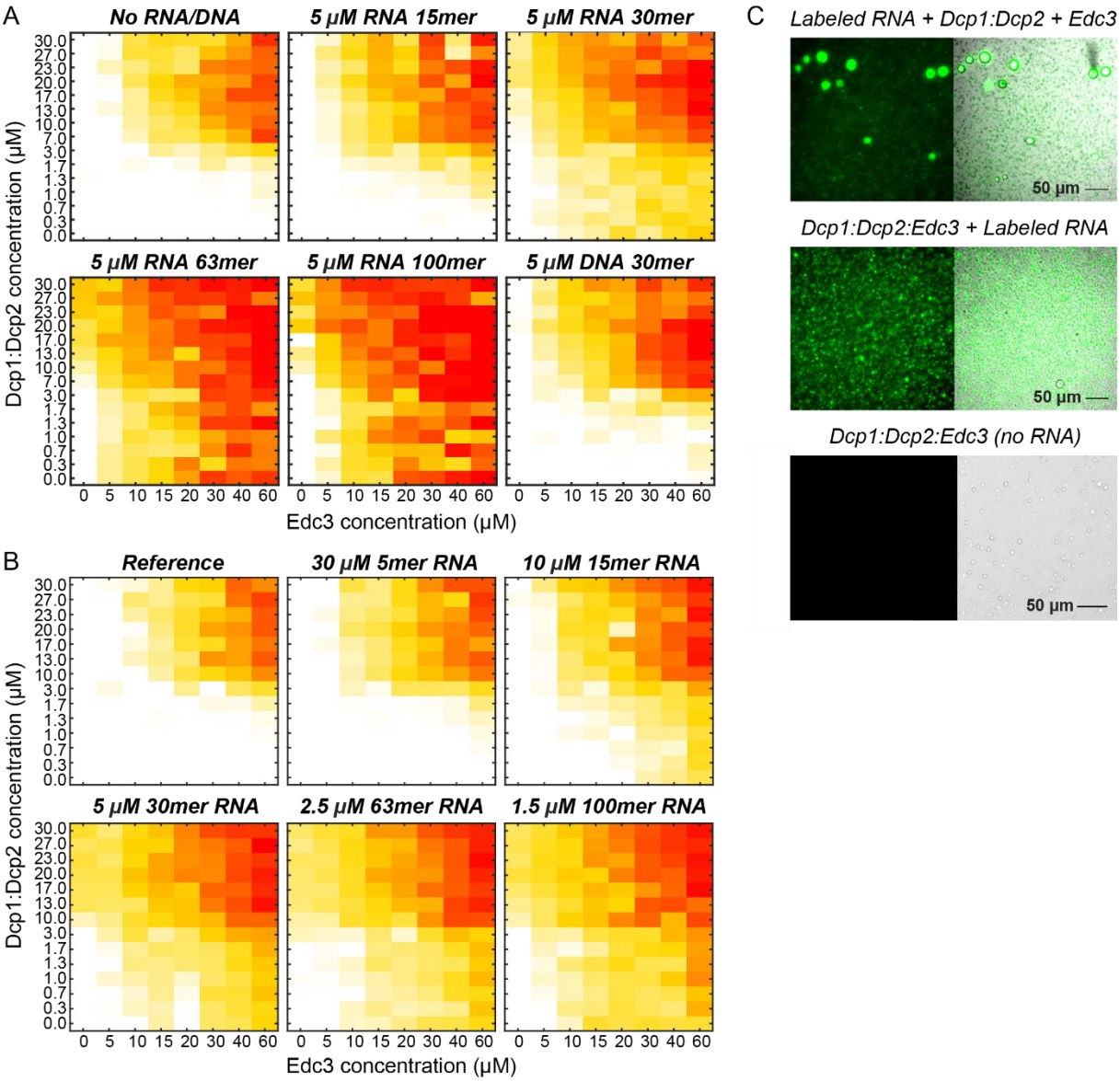

Figure 2.3: RNA strongly enhances LLPS. (A) Phase diagrams of Dcp1:Dcp2:Edc3 supplemented with 5 µM RNA of different length. RNA of 30 bases or more significantly moves the phase separation boundaries to lower concentrations, whereas DNA has no effect on the phase diagram. **(B)** Phase diagrams of Dcp1:Dcp2:Edc3 supplemented with RNA, however, as opposed to panel A, the total amount of nucleotides is kept constant. The phase diagrams show that one 100mer RNA has the same effect as three 30mer RNAs and indicate that the RNA binding events with Dcp1:Dcp2:Edc3 are fully independent. Shorter RNAs have no influence on the phase separation diagrams, indicating that the minimal length of an RNA that can efficiently be incorporated into the interaction network is around 30 bases. **(C)** Fluorescence microscopy images of Dcp1:Dcp2:Edc3 that is complemented with fluorescently labeled RNA. The RNA is highly enriched in the droplet phase. Addition of Dcp1:Dcp2 and Edc3 (no preformed Dcp1:Dcp2:Edc3 foci) to RNA (top) results in larger foci as when the RNA is added to Dcp1:Dcp2:Edc3 foci (middle), indicating that RNA is a highly efficient nucleation factor. No fluorescent signal is detected in the droplet phase in the absence of RNA (bottom), demonstrating that the fluorescent signal observed in the top and middle panels is due to fluorescently labelled RNA. The experimental conditions and the brightness and contrast are the same for all images. The size/ number of LLPS droplets in the bottom images is less than in the top/middle images due to the absence of RNA.

To experimentally validate that the RNA directly engages into the protein:protein interaction network we coupled the transcript to a fluorescence dye. Based on fluorescence microscopy images, we can unambiguously observe a strong enrichment of the labeled RNA in the droplet phase (**Figure 2.3 C**). Importantly, we observe that the size and the number of the liquid-like droplets depend on the order in which the components are mixed.

Addition of Edc3 and Dcp1:Dcp2 to the RNA resulted in larger and fewer droplets (**Figure 2.3 C**, top), whereas addition of RNA to preformed Edc3:Dcp1:Dcp2 droplets resulted in more and smaller droplets (**Figure 2.3 C**, middle). This indicates that RNA plays an important role in the nucleation dynamics of the LLPS process and that it can act as a strong nucleation factor[270] around which fast growing processing bodies are formed.

2.3.4 *RNA is protected against degradation by LLPS*

Inside the droplet phase the RNA is engaged in a tight network of protein:RNA interactions. This can provide advantages as it can potentially protect the transcript from exo- and endonucleolytic degradation. To test this hypothesis experimentally, we prepared an RNA substrate with 30 nucleotides that contains a single uridine at position 10. This site is a target for the enzyme RNase A that can cleave the substrate into 10mer and 20mer products (**Figure 2.4**). We then complemented the system with combinations of Dcp1:Dcp2 and Edc3. At the concentrations that we used, addition of Dcp1:Dcp2 resulted in weak phase separations (**Figure 2.3 A**, **B**) and a small degree of protection of the RNA (less cleavage). Addition of Edc3 resulted in intermediate phase separations and intermediate protection of the substrate. This protection is not due to direct interactions between the RNA and Edc3 as addition of the individual domains of Edc3 does not result in LLPS and does not protect the RNA from RNase A cleavage. Addition of both Dcp1:Dcp2 and Edc3 results in significant phase separations as well as in a significant protection of the RNA substrate. From these data we conclude that the LLPS process is able to protect the RNA substrate from enzymatic cleavages and degradation.

2.3.5 *LLPS reduces the catalytic activity of Dcp2*

One of the central components of processing bodies is the decapping complex Dcp1:Dcp2. To address how the activity of this enzyme complex is influenced by the inclusion into processing bodies we performed mRNA decapping experiments under multiple turnover conditions in the presence and absence of *in vitro* processing bodies. In these experiments we kept the concentration of the Dcp1:Dcp2 decapping complex constant and supplemented this with full-length Edc3 (to induce phase separation) or with exactly the same concentration of the isolated Edc3 domains (LSm, IDR and YjeF_N; **Figure 2.1 A**) that cannot induce LLPS. This ensures that the protein content in the different experiments is identical and thereby prevents indirect effects. Interestingly, we find that a small capped RNA with 21 nucleotides is decapped in the presence and absence of *in vitro* processing bodies with the same efficiency (**Figure 2.4 B**). This correlates with the limited effect that this short RNA has on LLPS (**Figure 2.3 B**).

In agreement with previous findings, we do, however, observe a modest stimulating effect of the Edc3 LSm domain on the decapping activity[101,104]. Interestingly, for a longer substrate with an RNA body of 100 nucleotides, which efficiently induces LLPS, we observe a significantly slower decapping activity under phase separation conditions (**Figure 2.4 C**). These *in vitro* results suggest that mRNA which is tightly embedded in *in vivo* processing bodies could be protected against decapping.

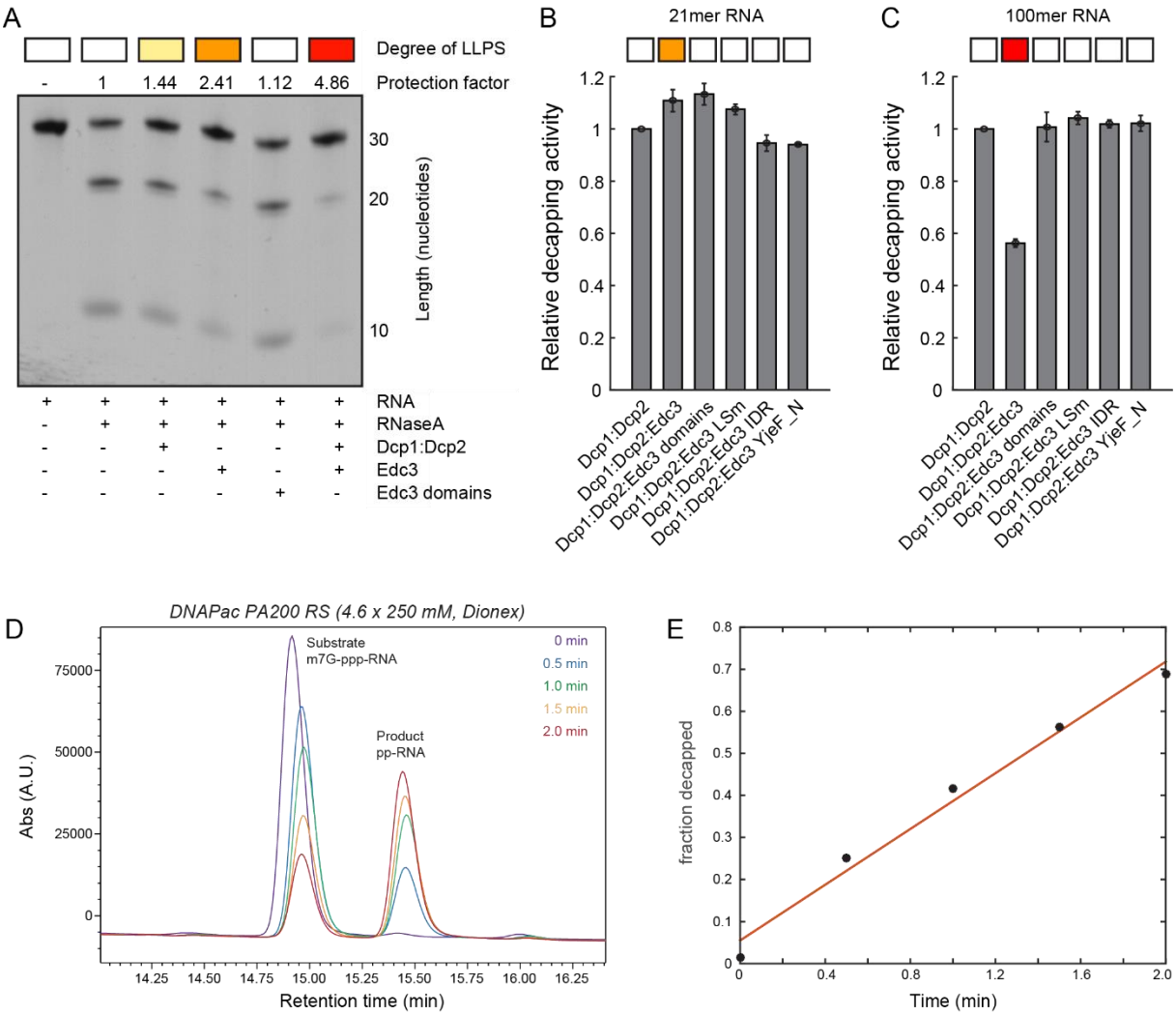

Figure 2.4: RNA is protected within LLPS foci. (A) RNase cleavage of a 30mer RNA (that contains a single RNaseA cleavage site at position 10) is reduced upon phase separation. This shows that the RNA within *in vitro* processing bodies is protected against RNase activity. This protection is not due to direct interaction between Edc3, Dcp1 or Dcp2 with RNA as addition of the same amount of the individual Edc3 domains (LSm, IDR and YjeF_N) does not result in RNase protection (and LLPS). The amount of LLPS is indicated on top, where the colour scheme of Figure 1B is used. **(B)** The Dcp2 activity on a short RNA of 20 nucleotides is independent of LLPS. **(C)** The Dcp2 activity on a longer RNA of 100 nucleotides is significantly reduced upon LLPS. This is not due to direct interactions, as the addition of the individual Edc3 domains has no influence on the decapping activity (see also A). The long RNA is efficiently embedded in the intermolecular interaction network that lead to LLPS (**Figure 2.3**), whereas the short RNA is not able to enhance phase separations of Dcp1:Dcp2:Edc3. It should be noted that the measured decapping activity is a weighted average of the activity of Dcp2 inside and outside the *in vitro* processing bodies, due to the exchange of components between the two phases. The activity of the decapping complex that is purely within the *in vitro* processing bodies is thus likely even lower than what we observe here. **(D+E)** Representative mRNA decapping experiment. 20 µM capped RNA substrate were supplemented with 1 µM Dcp1:Dcp2. At different time-points during the decapping reaction, a part of the sample was removed for analysis as described in the Methods section (2.2.7 and 2.2.8). **(D)** HPLC trace (DNAPac PA200 RS column (4.6 x 250 mM, Dionex) of an mRNA decapping experiment. The substrate (capped RNA) and the product (uncapped RNA) are indicated. **(E)** Time course of a decapping reaction, the best fit to the data is indicated. Note that the reaction is no longer linear after 2 minutes, due to product inhibition of the decapping enzyme.

2.3.6 *The IDR in Edc3 specifically interacts with RNA*

Edc3 contains an IDR region that links the N-terminal LSm domain and the C-terminal YjeF_N dimerization domain (**Figure 2.1 A**, **Figure 2.6 A**). To elucidate the importance of this IDR for LLPS, we deleted it from the Edc3 protein and prepared LLPS grids in the presence of Dcp1:Dcp2 (**Figure 2.5 A**). In the absence of RNA, we observe that deletion of the IDR moves the phase separation boundary slightly towards higher concentrations. More importantly, the addition of RNA to Edc3 lacking the IDR does not result in the strong shift of the phase separation boundary towards lower protein concentrations as observed for full-length Edc3 (**Figure 2.5 A**). This effect is even more prominent in RNA:Edc3 phase diagrams, where the removal of the IDR results in the almost complete loss of LLPS (**Figure 2.5 B**). This data indicates that the IDR in Edc3 is an important RNA interaction site.

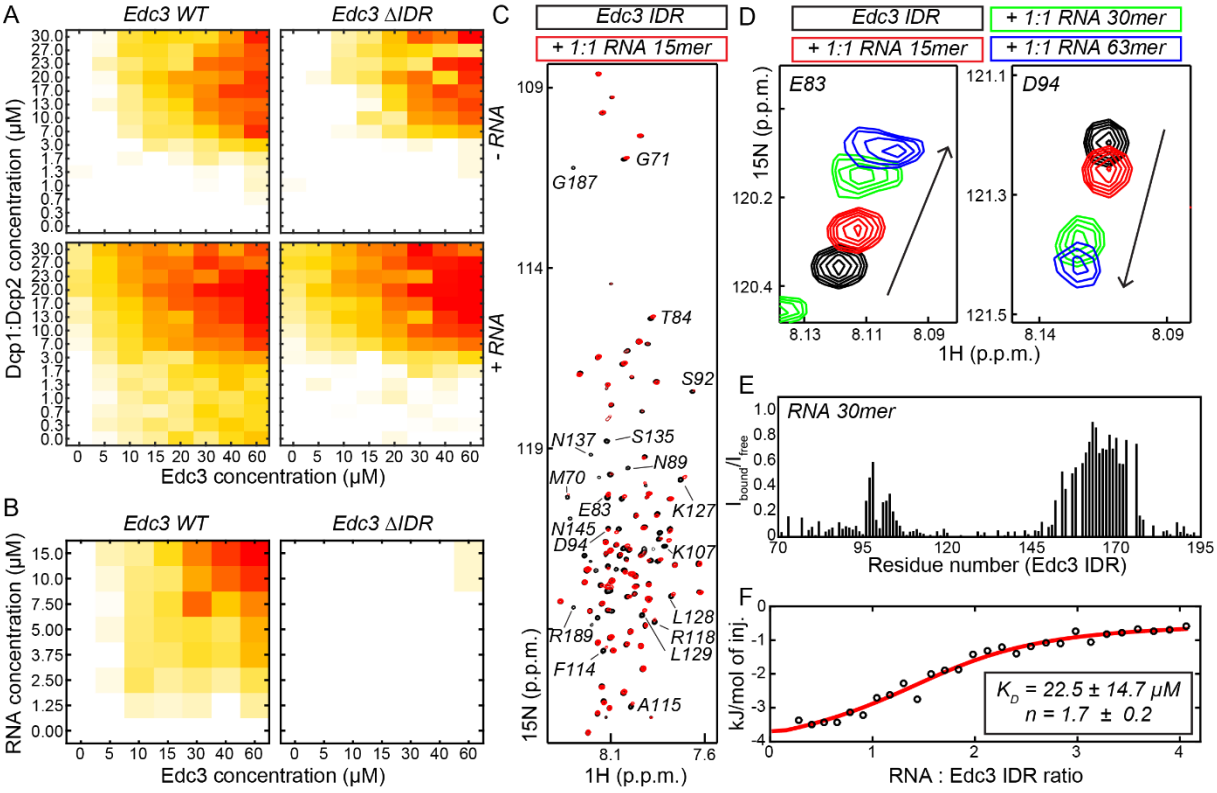

Figure 2.5: The Edc3 IDR specifically interacts with RNA. (A) Phase separation diagrams of Dcp1:Dcp2 and Edc3 in the absence (top) and presence (bottom) of a 30mer RNA. RNA significantly shifts the phase separation boundary for the wild-type (WT) Edc3 protein (left). Upon deletion of the IDR the effect of RNA on the phase diagrams is reduced (right). **(B)** Phase separation diagrams of Edc3 and a 30mer RNA. Edc3 and RNA are sufficient to induce LLPS (left). A version of Edc3 that lacks the IDR is no longer able to undergo phase separations in the presence of RNA only. **(C)** ^{1}H-^{15}N NMR spectra of the Edc3 IDR in the absence (black) and presence (red) of an equimolar amount of RNA of 15 nucleotides. Many resonances are significantly weaker or undergo CSPs, indicating a direct interaction between the RNA and the IDR. Several assignments are indicated. **(D)** The extent of the CSPs that are induced by the RNA correlate with the length of the RNA. This indicates that one RNA can interact with multiple IDRs. **(E)** Plot of the loss of intensity of the NMR signals in the Edc3 IDR upon addition of a 30mer RNA. Three regions in the Edc3 IDR interact with RNA: a region around residue 80, a region around residue 130 and a region around residue 180. **(F)** ITC binding experiments reveal a micromolar affinity between a 30mer RNA and the Edc3 IDR. The exact affinity cannot be extracted due to the unknown stoichiometry of the interaction.

To assess which residues in the Edc3 IDR interact with RNA, we used NMR spectroscopy. NMR spectra of the Edc3 IDR in the absence and presence of RNA show clear chemical shift perturbations (CSPs) that are restricted to a subset of the resonances (**Figure 2.5 C**). This indicates that a specific part of the IDR directly contacts the RNA. Interestingly, the extent of the observed CSPs correlates with the length of the added RNA (**Figure 2.5 D**), which shows that multiple IDRs can interact with a single RNA, in agreement with the phase separation diagrams of Edc3:Dcp1:Dcp2 and RNA (**Figure 2.3 A**, **B**). Based on the assignments of the Edc3 IDR resonances, there are three RNA interaction sites that are centered around residues 80, 130 and 180 (**Figure 2.5 E**). Notably, the region between residues 155 and 175, which is devoid of any positively charged residues, does not interact with RNA.

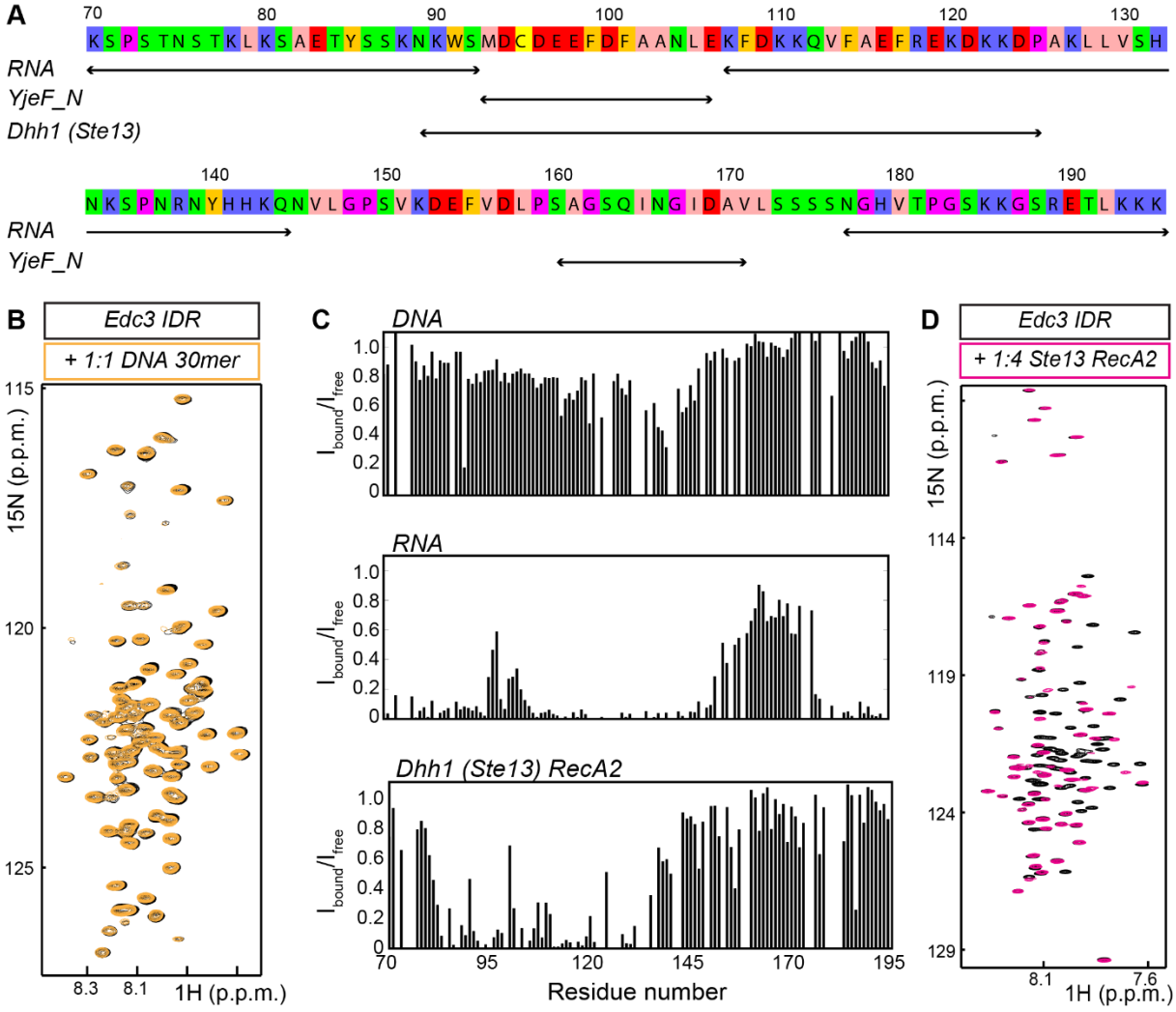

Figure 2.6: The Edc3 IDR does not interact with DNA, but with the second RecA domain of the helicase Dhh1 (Ste13). (A) Sequence of the Edc3 IDR. The residues are colored according to charge and hydrophobicity. The regions that interact with RNA (**Figure 2.5**), the Edc3 YjeF_N domain (**Figure 2.8**) and the helicase Ste13 (**Figure 2.7**) are indicated. **(B)** ^{1}H-^{15}N NMR spectra of the Edc3 IDR region in the absence (black) and presence (orange) of DNA. The lack of clear CSPs reveals that the Edc3 IDR does not interact with single-stranded DNA. **(C)** The intensity ratio of the NMR signals in the presence and absence of DNA is close to unity for all residues (top). As a comparison, the intensity ratio of the NMR resonances in the absence and presence of a 30mer RNA (middle) or the second RecA domain of Ste13 (bottom) is shown. The region in the Edc3 IDR that is affected by binding of Ste13 is around the conserved FDF motif. The binding sites of Ste13 and the Edc3 YjeF_N on the Edc3 IDR thus partially overlap. **(D)** ^{1}H-^{15}N NMR spectra of the Edc3 IDR region in the absence (black) and presence (hot pink) of the second RecA domain of Ste13.

ITC measurements reveal that the affinity between the RNA and the Edc3 IDR is in the micromolar range (**Figure 2.5 F**), although exact affinities cannot be extracted due to the unknown and dynamic stoichiometry of the Edc3 IDR:RNA interaction. The micromolar affinity is, however, in full agreement with other interactions that play a role in LLPS and allows for rapid readjustments within the interaction network[103,134]. Noteworthy, NMR titration experiments do not reveal an interaction of DNA with the Edc3 IDR (**Figure 2.6 B**, **C**). This is in agreement with our finding that DNA has no significant effect on the Edc3:Dcp1:Dcp2 phase diagram (**Figure 2.3 A**). Thus, the unstructured Edc3 IDR appears to be able to discriminate between RNA and DNA.

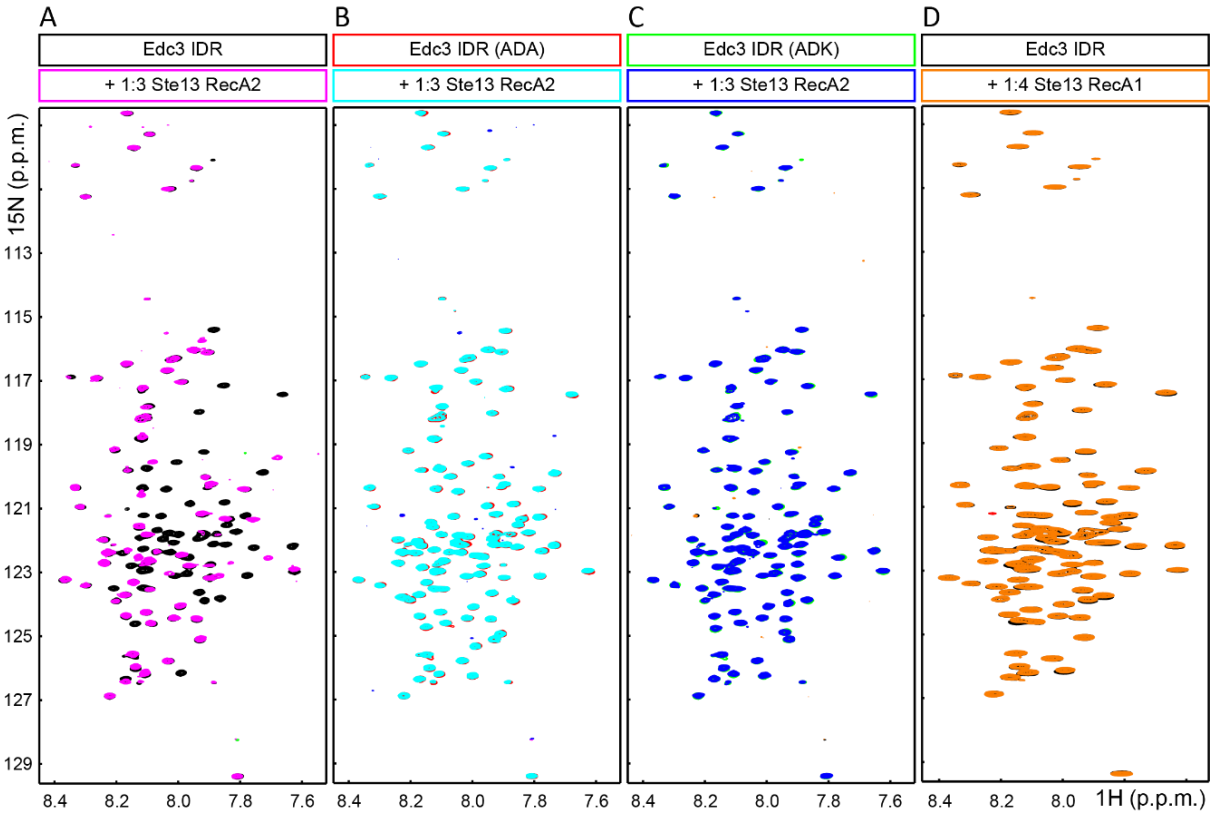

Figure 2.7: The Edc3 IDR interacts with Ste13 in a conserved manner. **(A)** ^{1}H-^{15}N NMR spectra of free (black) and bound (magenta) Edc3 IDR. Many resonances are significantly weaker or undergo CSPs in the presence of the second RecA2 domain of Ste13, indicating a direct interaction between the two proteins. **(B, C)** ^{1}H-^{15}N NMR spectra of mutant Edc3 IDR protein in the absence (B: red; C: green) and presence (B: cyan; C: blue) of the Ste13 RecA2 domain. The lack of CSPs indicate that either of the two mutants ADA **(B)** or ADK **(C)** abolishes binding of the Edc3 IDR to Ste13. **(D)** The first RecA1 domain does not interact with the Edc3 IDR as no CSPs are observed in the presence (orange) of Ste13 RecA1 compared to apo Edc3 IDR (black).

2.3.7 *Interactions of the Edc3 IDR and the RNA-helicase Dhh1 are conserved from yeast to human*

Major processing body components and decapping factors are conserved in all higher eukaryotes[271]. Additionally, orthologs of the Dhh1 helicase were found to interact with the Edc3 protein in a similar manner, where the second but not the first RecA-like domain of the helicase binds to conserved FDF and FDK peptide motifs in an intrinsically disordered region of Edc3[106,107,144].

We performed chemical shift perturbation experiment to test whether the interaction between Edc3 and the Dhh1 ortholog Ste13 is also conserved in *S. pombe*. We could identify residues in the Edc3 IDR that are perturbed upon binding of the second RecA2 domain of Ste13 (**Figure 2.6 D**). These residues span a region of around 40 amino acids (residues 90-130) and contain the conserved FDF and FDK motifs (**Figure 2.6 A**, **C**). Mutating either the FDF motif (to ADA) or the FDK motif (to ADK) completely abolishes Ste13 binding to Edc3. Notably, the Edc3 IDR does not interact with the RecA1 domain of Ste13. These results are in line with previous findings and argue for a conserved mode of interaction between Edc3 and Dhh1 in different yeast species and from lower to higher eukaryotes.

2.3.8 *Maturation of processing bodies*

Previous work has shown that cellular LLPS can undergo maturation and form more solid-like structures that, in specific cases, correlated with human disease[123]. Interestingly, we observe that fused Edc3:Dcp1:Dcp2 droplets also form a gel-like film over time that can be scratched from the bottom of the well of a 96-well plate with a tip of a pipette (**Figure 2.8**). In addition, as opposed to freshly formed LLPS droplets, this film does not rapidly dissolve upon dilution and is resistant to high salt concentrations (**Figure 2.8**). We therefore conclude that droplets that are formed from processing body components also undergo maturation. To unravel interactions contributing to the maturation effect, we focused on the Edc3 protein. Edc3 in isolation undergoes phase separations at high concentrations in a temperature- and salt-dependent manner, even in the absence of RNA and Dcp1:Dcp2 (**Figure 2.1 C** at zero Dcp1:Dcp2 concentration). Maturation occurs over time and to assess the underlying mechanisms of this process we measured both turbidity (as a measure of the phase separation) and the soluble protein concentration (as a measure of the amount of protein that engages in the insoluble mature phase) at different timepoints. To modulate the amount of LLPS we performed these measurements at ten salt concentrations between 20 and 125 mM.

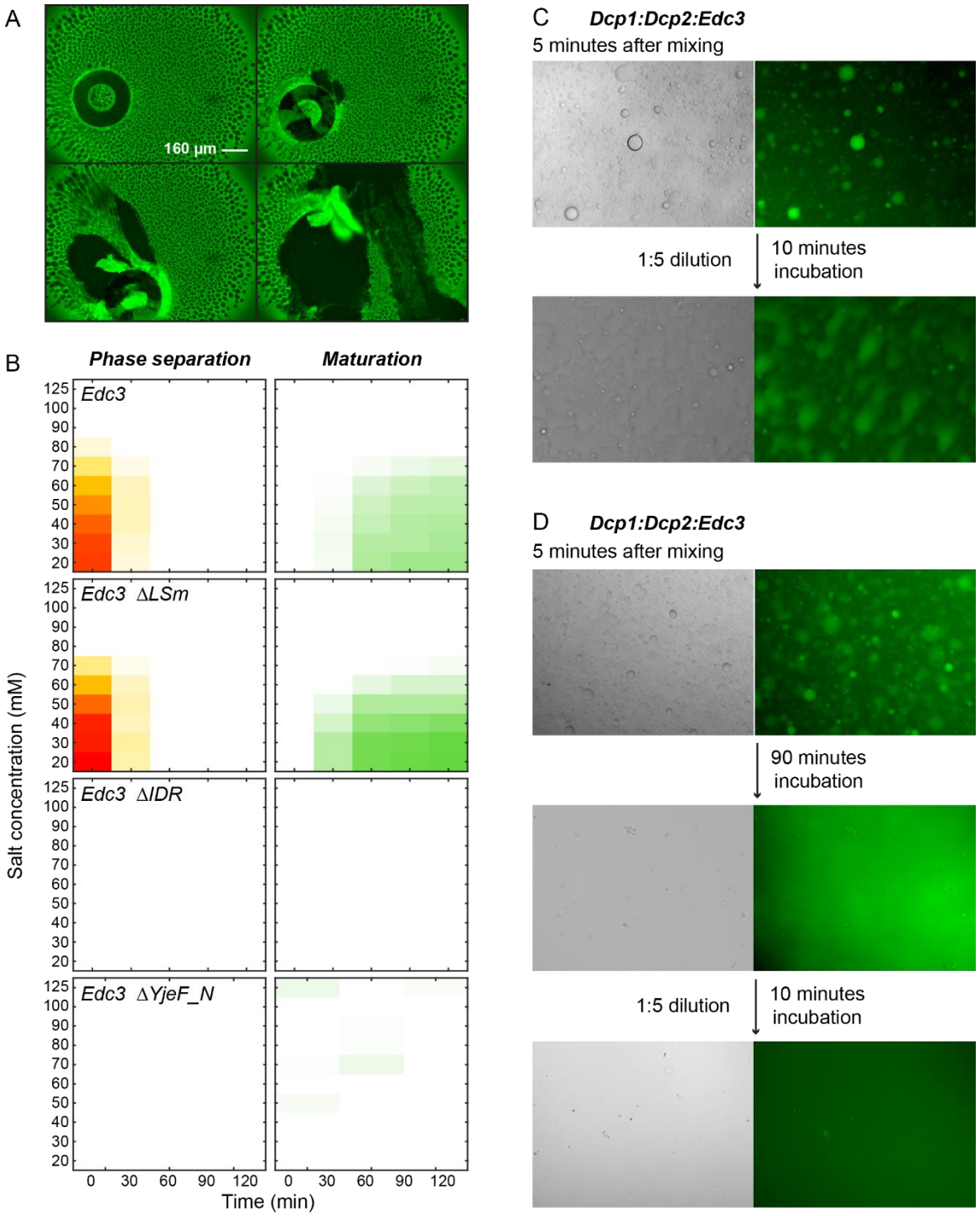

Figure 2.8: Maturation of *in vitro* processing bodies into a dilution-resistant phase. (A) Successive fluorescence microscopy images that were taken of Dcp1:Dcp2:Edc3 phase separated proteins after incubation for 12 hours. The proteins formed a gel-like film on the bottom of the well that could be scratched off using a pipette tip. **(B)** LLPS (left panels) and maturation (right panels) of the Edc3 protein at different salt concentrations (y-axis) were monitored over time (x-axis). Full-length Edc3 (top panel) undergoes phase separation at low salt concentrations. The formed foci (left) merge and subsequently form a gel-like film (right). Deletion of the IDR or the YjeF_N domain in Edc3 inhibited the phase separation and maturation processes, indicating that the interaction between these domains plays an important role in the maturation of Edc3-containing foci. **(C+D)** Dcp1:Dcp2 and fluorescently labeled Edc3 were mixed. **(C)** Top: Widefield and fluorescent images after 5 minutes incubation. Bottom: images of the same condition, after 1:5 dilution and 10 minutes incubation. The LLPS droplets largely disappeared. A small degree of matured droplets is visible in the fluorescent images. These have formed during the incubation times. **(D)** Liquid droplets matured for 90 minutes. During this time, they merge as a homogeneous layer on the bottom of the well. Subsequent dilution does not result in the disappearance of the homogeneous gel-like structure, indicating that the intermolecular interactions in the early **(C)** and late **(D)** *in vitro* processing bodies are significantly different.

At salt concentrations between 20 and 80 mM, the Edc3 wild-type protein undergoes LLPS and liquid-like droplets form (**Figure 2.8 B**, top left). Over time, these droplets disappear and in parallel a gel-like mature phase forms (**Figure 2.8 B**, top right). At salt concentrations above 80 mM, Edc3 does not undergo LLPS and no mature phase is formed. This indicates that the formation of the gel-like structure is a direct result of the formation of liquid-like droplets and that the LLPS is required for the aggregation process. Next, we repeated these experiments for Edc3 proteins lacking the LSm domain (**Figure 2.8 B**, second row), the IDR (third row) or the YjeF_N domain (bottom row) and observe that removal of either the IDR or the YjeF_N domain prevents the formation of mature droplets. Our data thus suggests that an interaction between the Edc3 IDR and YjeF_N domain underlies the maturation of *in vitro* processing bodies. To confirm an interaction between the Edc3 LSm and YjeF_N domains, we performed NMR titration experiments and identified two regions in the Edc3 IDR that directly and independently interact with the YjeF_N domain (**Figure 2.9 A**, **B**). Only upon deletion of both regions, the interaction between the Edc3 IDR and the Edc3 YjeF_N domain is abolished, while the ability of RNA binding persists (**Figure 2.9 C**), indicating that both interactions are not competing.

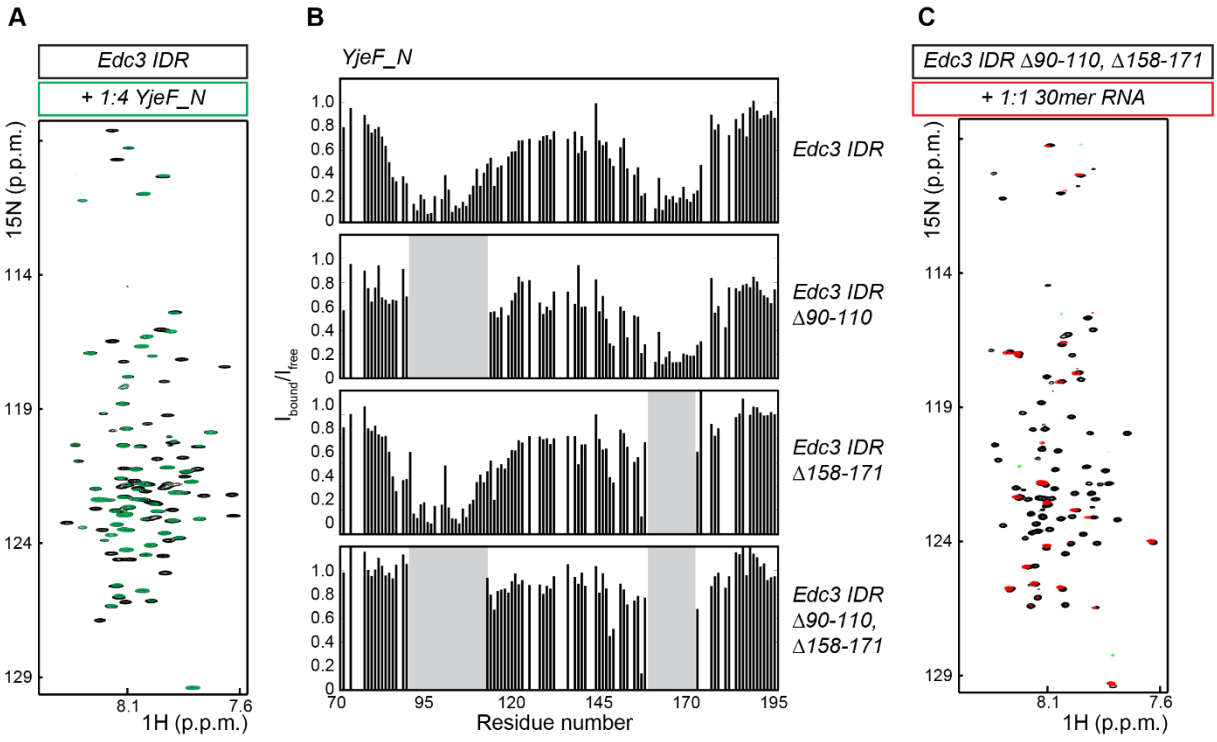

Figure 2.9: **Two regions in the Edc3 IDR interact independently with the Edc3 YjeF_N domain. (A)** ^{1}H-^{15}N NMR spectra of the Edc3 IDR region in the absence (black) and presence (green) of the Edc3 YjeF_N domain. **(B)** Intensity ratios of the ^{1}H-^{15}N resonances of the Edc3 IDR in the presence and absence of the Edc3 YjeF_N domain. Two regions in the Edc3 IDR are prominently influenced by this interaction: a region around residue 100 and a region around residue 165. Deletion of one of these regions does not influence the interaction with the other region (middle two panels, the deleted region is indicated in gray). Deletion of both regions (lower panel) abolishes the interaction between the Edc3 IDR and the Edc3 YjeF_N domain. **(C)** The Edc3 IDR double-deletion mutant, which is no longer able to interact with the YjeF_N domain, still binds RNA.

2.4 *DISCUSSION*

Processing bodies are ubiquitous in eukaryotic cells. Nevertheless, their function remains a matter of debate. Therefore, we here address the formation and subsequent maturation of processing bodies. In addition, we examine how the activity of the decapping enzyme is modulated by the incorporation into highly densely packed cellular foci.

The formation of cellular foci is governed by a redundant set of intermolecular interactions that initially result in the appearance of liquid-like droplet that spontaneously separate from the surrounding. *In vivo*, this LLPS process is counteracted by processes that can disintegrate these cellular foci into individual soluble proteins[139,163,164,257]. In an *in vitro* setting, however, liquid droplets continuously flow and fuse into larger assemblies that eventually settle as a homogeneous layer on the bottom of a test-tube[103]. To study the mechanisms behind LLPS processes in an *in vitro* setting, it is thus important to assess the formation of liquid droplets rapidly, before the proteins that are involved are irreversibly aggregating into gel-like structures. To that end, we here introduce an approach where hundreds of conditions can be screened rapidly, such that the irreversible aggregation of proteins can be neglected. Our approach is based on the use of small volume pipetting robotics and the assessment of the LLPS through turbidity measurements. These quantitative turbidity measurements on 5 µL volumes quantitatively correlate very well with the LLPS experiments on 80 µL volume samples using microscopic techniques (**Figure 2.1 B**). The small volume turbidity measurements, however, require significantly less time (factor of 100) and sample (factor of 16) and thus allow for the reliable assessment of LLPS screens that report only on the initial clustering process without interference from the maturation/aggregation process.

Based on the above methodology we assessed how incorporation of processing body proteins and RNA in *in vitro* assays influences the phase separation boundary. Pdc1 significantly reduces the required concentrations of Dcp1:Dcp2 and Edc3 to induce LLPS (**Figure 2.1 D**). Mechanistically this can be explained by the interactions between the Pdc1 HLMs and the Edc3 LSm domain that reduce the requirement of the Dcp2 HLMs to form indefinite interaction networks[103]. Importantly, RNA is also able to significantly reduce the concentrations that are required for the LLPS process of Dcp1:Dcp2 and Edc3 (**Figure 2.3 A**, **B**). To that end, the RNA needs to be longer than 20-30 bases. Shorter RNAs are most likely not able to bridge between different RNA binding sites in Dcp2 and Edc3 and can therefore not enhance the clustering process. RNA that is longer than 30 bases is efficient in influencing the LLPS process, where three 30mer RNAs have the same effect as one 100mer RNA has. This indicates that the multiple protein:RNA interactions in long RNAs are independent.

Cellular phase separation processes are modulated by interactions between folded protein domains (e.g. the interaction between Dcp1 and Dcp2, or the interaction between Dcp1:Dcp2 and the Ge1 domain in Pdc1[103]; **Figure 2.1 A**), interactions between disordered regions and folded protein domains (e.g. Edc3 LSm domain and the Dcp2 and Pdc1 HLMs[99,103]; **Figure 2.1 A**), interactions between folded protein domains and RNA (e.g. Dcp2 and RNA; **Figure 2.1 A**). Here, we add a fourth type interaction and show that the disordered region in Edc3 efficiently interacts with RNA (**Figure 2.1 A**, **Figure 2.5 C-F**). These IDR:RNA interactions are functionally important as removal of the IDR results in a reduction of the LLPS processes (**Figure 2.5 A**, **B**). We show that the collaborative sum of all these interaction types can move the LLPS boundary to significantly lower concentrations. Within a cellular setting the concentrations of the Dcp1, Dcp2, Pdc1 and Edc3 proteins are approximately between 20 and 200 nM[272]. Our phase diagrams now show clear LLPS close to these concentrations, despite the fact our used RNAs are significantly shorter than the average of cellular mRNAs and that many processing body components are still absent from our *in vitro* setup. Based on the additive effect that we describe here we confidently predict that addition of more P-body proteins will result in the formation of foci at cellular concentrations. In that light, it is important to note that the proteins Pat, Lsm1-7, Dhh1 (Ste13), Xrn1 (Exo2), Scd6 and Edc1 are all reported to interact with Dcp1, Dcp2, Edc3, Pdc1 and/or RNA[99,101] and that they can thus be efficiently embedded into the intermolecular interaction network that drives the LLPS process.

Maturation of cellular foci is implicated with medical disorders[273]. Due to the high local protein concentration in cellular foci the phase-separated proteins are likely prone to aggregation that can lead to the formation of amyloid like structures. Here, we observe that *in vitro* processing bodies also evolve into highly insoluble aggregates. This maturation is driven by interactions between the Edc3 IDR and the Edc3 YjeF_N domain (**Figure 2.8**). To what degree the maturation of processing bodies in an *in vivo* setting is correlated with medical disorders will be an exciting question for future studies.

Based on our data, the 120 amino acid long IDR of Edc3 plays multiple roles (**Figure 2.1 A**, **Figure 2.6 A**), in agreement with the previously reported importance of Edc3 in processing body formation[144]. First, it directly interacts with RNA and thereby enhances LLPS. Second, an FDF motif in this region can recruit the helicase Dhh1 (Ste13) through interactions with the second RecA domain. This interaction was previously shown for the human homologues[106] and is conserved in *S. pombe* (**Figure 2.6 C**, **D**). Finally, the Edc3 IDR directly interacts with the Edc3 YjeF_N dimerization domain, which is involved in the irreversible maturation process. Interestingly, the Dhh1 (Ste13), RNA and YjeF_N binding sites in the Edc3 IDR partially overlap (**Figure 2.6 A**, **C** and **Figure 2.9 B**). Whether the interaction of RNA with Edc3 influences the maturation process in processing bodies, as was recently shown for Whi3 droplets[122], remains to be determined.

Here, we define the intrinsically disordered region in Edc3 as a multivalent interaction platform for proteins and RNA, which underscores the importance of disordered regions in proteins for function[274].

To address the functional implication of processing bodies we assessed the mRNA decapping activity within these foci (**Figure 2.4 B**, **C**). Interestingly, the mRNA decapping activity is reduced within processing body-like structures for RNA substrates longer than 20 nucleotides. These longer RNAs are tightly incorporated within the network of interactions that drives processing body formation, which will restrict their local diffusion rates. The reduction in decapping activity within our *in vitro* processing bodies correlates well with the proposed role of processing bodies in temporary mRNA storage[114,259,262,275], where mRNA is temporarily taken out of the translational pool such that the cell can react efficiently to stress situations. In agreement with that, the RNA inside droplets is also protected against RNase activity (**Figure 2.4 A**).

In summary, our data reveal a highly intertwined network of intermolecular interactions that is responsible for the formation and maturation of processing bodies. Interestingly, this network of interactions is able to significantly reduce the mRNA decapping activity. Our studies thereby provide unique insights into the mechanisms of formation and the functional implication of processing bodies.

CHAPTER 3 *Atomic level insight into the maturated state of mRNA processing bodies by combining solid- and solution-state NMR spectroscopy*

This chapter was published in Nature Communications. Large parts of the text and figures are thus identical to the published manuscript. The project was supervised by Remco Sprangers and Marc Baldus. Experiments were designed by all authors. Stefan Schütz performed molecular cloning, protein labeling, expression and purification and prepared samples for NMR. Reinier Daman, Yanzhang Luo and Markus Weingarth acquired solid-state NMR data, while S.S. acquired solution-state NMR data. M.W. performed molecular dynamics simulations. All authors contributed to the analysis and interpretation of data. M.B. and R.S. wrote the manuscript and all authors commented on it.

3.1 *INTRODUCTION*

Increasing evidence shows that cells develop intra-cellular protein-rich organelles that grow and fuse, allowing certain molecules to become locally enriched while excluding others[118]. Such intracellular liquid–liquid phase separation (LLPS) processes seem to be a common approach to achieve spatial organization of cellular components into dynamic, membrane-less compartments. While physical models have helped to understand the formation of membrane-less protein compartments[119], studying such networks at atomic resolution has been challenging due to their intrinsic mobility and heterogeneous nature[276,277] and because genetic approaches in a cellular context are complicated by the highly redundant nature of the clustering process.

NMR spectroscopy has been shown to provide unique structural insights into heterogeneous and dynamical systems at atomic resolution[278–281]. Previously, solution-state NMR has been used to decipher molecular interactions among LLPS components under *in vitro* conditions using titration studies[103,269] and determined 3D structures of fast tumbling folded subdomains[282]. However, studying the structural organization of liquid-like droplets using solution-state NMR methods is complicated by the reduced diffusion of the liquid-liquid phase separated molecules[283]. Hence, direct insights into the structural and dynamical organization of liquid-like droplets has been limited[283]. Instead, solid-state NMR (ssNMR) has been employed to study the formation[158,284] and structural organization[158,285–288] of protein hydrogels and fibrillar assemblies that may be closely related to LLPS[276,289].

For example, we have previously used ssNMR to investigate the hydrogel state of the FG repeat domain of the nucleoporin Nsp1p related to the nuclear pore complex, revealing that transient amyloid-like β-sheet interactions among NTQS-rich protein regions are responsible for gelation and network formation[158,286]. Such on-pathway amyloid interactions were also recently proposed for hydrogel-forming peptides that contain GSY-amino acid rich stretches and, in isolation, form crystalline needles that shared essential features of amyloid fibrils[147]. Indeed, ssNMR revealed characteristic β-strand arrangements in fibrils formed by the low complexity domain of the FUS RNA-binding protein[148]. A relationship between LLPS and fibril-like protein states and their relevance for human disease has also been examined in stress granules, cytosolic membrane-less organelles composed of RNA-binding proteins (RBPs) and RNA[123], and, more recently, for huntingtin exon1 assemblies in mammalian cells and yeast[145]. While LLPS and fibrillization may represent two mechanistically distinct processes, models that connect these two processes have been developed[123,151,289]. For example, Molliex *et al.* have depicted the relationship between phase separation, fibrillization, and pathological inclusions by three steps with an intermediate maturation regime where two phases coexist in the granule state, which ultimately connects LLPS to disease[123]. However, the structural relationship between liquid-liquid phase separation and maturation has remained elusive.

In the following, we combined solid- and solution-state NMR experiments to directly study the LLPS and maturation of processing bodies (P-bodies) that are dynamic cytoplasmic ribonucleoprotein (RNP) granules containing proteins involved in translational repression and RNA degradation[89,113,252,290]. Their main constituents are mRNA and the mRNA decay machinery, including the Dcp1:Dcp2 mRNA decapping complex, the RNA helicase Dhh1, the Pat–LSm1-7 complex, the exonuclease Xrn1 and Edc3 (enhancer of decapping 3)[144,252]. The Edc3 protein is a central hub for processing body formation in yeast. The 50 kDa protein (**Figure 3.1**) comprises an N-terminal LSm domain that directly interacts with helical leucine-rich motifs (HLMs) in the P-body proteins Dcp2[101] and Pdc1[103]. C-terminal to the LSm domain, the Edc3 protein contains a 120 amino acid long intrinsically disordered region (IDR). This IDR is an interaction platform for RNA[252], the RNA helicase Dhh1[105] and the C-terminal Edc3 YjeF_N dimerization domain[252,291].

Previously, we showed the *in vitro* reconstitution of processing body-like assemblies using purified components[103,252]. These Edc3-containing *in vitro* foci undergo a maturation process, which over time results in the irreversible formation of an insoluble gel-like state. Based on solution-state NMR spectroscopy and phase separation diagrams, we identified key interactions in this sequential two-step process and found that interactions between the Edc3 LSm domain and HLMs in Dcp2 and interactions between the Edc3 IDR and RNA are required for LLPS[103,252].

Recently, we noticed that the isolated Edc3 protein is also able to undergo LLPS and we have shown that the IDR and the YjeF_N domain are key for this process[252]. Moreover, the ability of Edc3 to mediate phase transitions is significantly enhanced by RNA which can bridges between Edc3 protomers through interactions with the Edc3 IDR.

Solution-state NMR spectroscopy is ideally suited to study the proteins and protein:RNA complexes before they undergo LLPS, however, it is less optimal to study full-length processing body components after they have undergone phase separation. This is due to the limited diffusion within the high-density phase and due to the fact that these *in vitro* processing bodies mature into a more solid-like phase that does not spontaneously dissolve into a liquid phase upon dilution[252]. To study the structure, dynamics and RNA interactions of the Edc3 protein in the matured high-density protein state we here turned to solid-state (ss) NMR.

In the following, we conducted $^{13}C/^{15}N$- as well as ^{1}H-detected ssNMR experiments using a combination of scalar- and dipolar-based correlation methods that, as we have shown previously[158,286,292], provide a powerful tool to separate ssNMR protein signals of mobile and rigid protein regions. For ^{1}H-detected ssNMR experiments, we furthermore made use of proton dilution approaches[293,294] to increase spectral resolution under high-field/high-speed magic angle spinning (MAS) conditions. We applied our strategy to the full-length Edc3 protein as well as to the matured state of an Edc3 construct that lacks the N-terminal LSm domain (ΔLSm, **Figure 3.1 A**) representing the minimal Edc3 construct that can undergo LLPS[252]. In addition, we examined the C-terminal YjeF_N domain by both solid- and solution-state NMR and conducted ssNMR experiments of Edc3 samples after the addition of RNA, which significantly enhances the LLPS process of Edc3.

Our studies reveal that the intra- and intermolecular contacts that are present between the Edc3 protein and RNA before LLPS are also present in the matured granule state of processing bodies. These interactions are diverse and range from weak electrostatic to strong hydrophobic contacts and include folded protein domains as well as IDRs.

3.2 *MATERIALS AND METHODS*

3.2.1 *Protein expression and purification*

The genes for the full-length Edc3 protein, for Edc3 lacking the LSm domain (Edc3-ΔLSm) and for the Edc3 YjeF-N domain from *Schizosaccharomyces pombe* were cloned into modified pET vectors that carried an N-terminal TEV-cleavable His_6-tag.

The gene for the IDR of Edc3 was cloned into a modified pET vector that carried an N-terminal TEV-cleavable His_6-GST-tag. Point mutations and deletions were introduced using standard site-directed mutagenesis methods (**Table 3.1**).

Chemical competent *Escherichia coli* BL21 (DE3) Codon Plus RIL cells (Stratagene) were transformed with the appropriate plasmid and grown at 37 °C to an OD_{600} of 0.8 in lysogeny broth (LB) medium. Overexpression was induced with 0.5 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG). After overnight expression at 20 °C cells were pelleted by centrifugation and resuspended in buffer 1 (25 mM sodium phosphate, pH 7.4, 250 mM NaCl, 1 mM DTT) complemented with 10 mM imidazole, 5 mM $MgCl_2$, 0.1 % Triton X-100, lysozyme and 0.2 U/mL DNase I. Cells were lysed with ultrasound and insoluble cell debris were removed from the lysate by centrifugation. The supernatant was bound to equilibrated Ni-NTA resin, which was washed subsequently with buffer 1 complemented with 10 mM imidazole. Bound protein was eluted from the resin with buffer 1 complemented with 300 mM imidazole. Tobacco etch virus (TEV) protease was added to the elution fractions to cleave the purification tag from the target proteins.

The Edc3 IDR was dialyzed overnight at 4 °C into buffer 2 (25 mM HEPES, pH 8.0, 50 mM NaCl, 1 mM DTT) supplemented with 0.5 mM EDTA. The dialysate was applied to a cation exchange column (HiTrap 5mL SP FF, GE Healthcare) equilibrated with buffer 2. To separate the protein of interest from the cleaved GST-tag, a gradient from 0 to 35 % buffer 3 (25 mM HEPES, pH 8.0, 2 M NaCl, 1 mM DTT) was applied over a volume of 100 mL. Fractions containing the target protein were pooled and the buffer was exchanged to buffer 4 (25 mM HEPES, pH 7.3, 125 mM NaCl, 1 mM DTT) using centrifugal filters. Proteins carrying only an N-terminal His_6-tag were directly dialyzed into buffer 4 supplemented with 0.5 mM EDTA. Purification to homogeneity was achieved by size exclusion chromatography on HiLoad 16/600 Superdex 75 (Edc3 IDR, Edc3 YjeF_N) or Superdex 200 (Edc3-FL and Edc3-ΔLSm) columns (GE Healthcare) in buffer 4.

NMR-active proteins were obtained by overexpression in (50-100%) D_2O-based M9 minimal medium supplemented with 0.5 g/L $^{15}NH_4Cl$ as the sole nitrogen source and either 2 g/L $^{1}H^{13}C$-glucose for solid-state NMR samples or 4 g/L $^{1}H^{12}C$-glucose for solution-state NMR samples. For labeling of the Ile-δ_1, Met-ε, Val-γ_1/γ_2 and Leu-δ_1/δ_2 methyl groups within a U-[$^{2}H^{12}C$]-background, the medium was supplemented with α-ketobutyrate (3-$^{2}H_2$-4-$^{13}CH_3$; 60 mg/L), methionine (methyl-$^{13}CH_3$; 100 mg/L), α-ketoisovalerate (3-^{2}H-3-(methyl-$^{13}CH_3$)-4-$^{13}CH_3$; 100 mg/L) one hour prior to induction. For simultaneous labeling of Ala-β methyl groups, 2-^{2}H-3-^{13}C-L-alanine was added to the medium 20 minutes prior to induction.

Table 3.1: Edc3 constructs used in this study

Residue number / domain / mutations	Solubility / purification tag	Internal reference
1-454	N-His_6-TEV	#180
72-194 (IDR)	N-His_6-GST-TEV	#1293
195-454 (YjeF_N)	N-His_6-TEV	#196
YjeF_N (M-2G, L254I)	N-His_6-TEV	#1797
YjeF_N (M-2G, K258M)	N-His_6-TEV	#1798
YjeF_N (M-2G, L390M)	N-His_6-TEV	#1799
YjeF_N (M-2G, K392M)	N-His_6-TEV	#1800
YjeF_N (M-2G, W393M)	N-His_6-TEV	#1776
YjeF_N (M-2G, L445I)	N-His_6-TEV	#1801
YjeF_N (M-2G, Q447I)	N-His_6-TEV	#1802
YjeF_N (V253I)	N-His_6-TEV	#1786
YjeF_N (L370M)	N-His_6-TEV	#1787
YjeF_N (V416M)	N-His_6-TEV	#1788
72-454 (ΔLSm)	N-His_6-TEV	#1301
72-454, Δ90-110 (IDRΔ1 – YjeF_N)	N-His_6-TEV	#1515
72-454, Δ158-171 (IDRΔ2 – YjeF_N)	N-His_6-TEV	#1513
1-454	N-MBP-TEV	#259
195-454 (Yjef_N)	N-MBP-TEV	#258

3.2.2 *RNA in vitro transcription and purification*

RNA *in vitro* transcription and purification was carried out as described before[252]. *In vitro* transcription of a 30mer RNA was accomplished using in-house purified T7 polymerase[263] and a DNA primer as the template (**Table 3.2**). Transcribed RNA was purified under denaturing conditions using anion exchange chromatography, followed by isopropanol precipitation, desalting and vacuum concentration. The obtained pure dry RNA product was resuspended at concentrations required for subsequent experiments. The quality of the RNA was assessed using urea-polyacrylamide gel electrophoresis (PAGE) in 1× TBE (89 mM Tris, pH 8.0, 89 mM boric acid, 2 mM EDTA). RNA was visualized by methylene blue staining.

Table 3.2: RNA used in this study

RNA	Properties	Sequence	Internal reference
30U15mer	GA-only, but single U at position 15	GGAGGAGAGGAAGGUAAGGGAAGAAAGAAG	#10 primer

3.2.3 *Liquid-liquid phase separation assays*

LLPS maturation experiments were carried out as described previously[252]. In brief, Edc3-ΔLSm samples were prepared at 150 µM in 25 mM HEPES, pH 7.3 and different salt concentrations. At each time-point, the solution was mixed by pipetting and the turbidity (OD_{600}) was measured. 8 µL of the sample were collected, liquid droplets were dissolved by adding 2 µL 2.5 M NaCl and the protein concentration was determined by absorption measurements at 280 nm. Based on the dilution factor the concentration of soluble protein within the sample was obtained. The decrease in the protein concentration over time was used as a measure of the fraction of insoluble protein (matured state).

3.2.4 *Solution-state NMR experiments*

All solution-state NMR samples were prepared in buffer 4 and contained 5 % D_2O. NMR spectra were recorded at 298 K on Bruker AVIII-500 and AVIII-800 spectrometers equipped with room temperature and cryogenic probe-heads, respectively. NMR titration experiments were carried out with 0.05-0.1 mM ^{15}N- and ILVM methyl-labeled protein (Edc3 YjeF) and a 4- to 6-fold excess of unlabeled protein (Edc3 IDR). NMR spectra were processed using the NMRPipe/NMRDraw software suite[267]. Figures displaying NMR spectra and protein structures were prepared using NMRview (onemoonscientific.com) and Pymol (pymol.org), respectively.

3.2.5 *Solid-state NMR experiments*

Phase-separated Edc3 samples for solid-state NMR measurements were obtained by lowering the salt concentration of concentrated protein from 125 mM to 25 mM by dilution with buffer 5 (25 mM HEPES, pH 7.3, 1 mM DTT). For samples to contain RNA, the protein was supplemented with a 2.5-fold molar excess of 30mer RNA in buffer 5. Protein and protein:RNA samples were incubated for 30 minutes at room temperature and droplets were pelleted by centrifugation at 15.000 g for 15 minutes. The supernatant was discarded, and the gel was transferred to the solid-state NMR rotors.

Solid-state NMR experiments were conducted using 3.2 mm and 1.3 mm triple-resonance (^1H, ^{13}C, and ^{15}N) magic-angle-spinning (MAS) probe heads in static magnetic fields of 9.4, 16.4, 18.8 and 22.2 T, corresponding to proton resonance frequencies of 400, 700, 800 and 950 MHz, respectively. Scalar-based correlation experiments utilized HC INEPT[295] and 6 ms mixing C-C TOBSY[296] transfer steps. Samples were chilled to approximately 278 K sample temperature and spun at a MAS rate of 9 kHz. ^1H-detected experiments were performed at a MAS rate of 60 kHz and a sample temperature of approximately 310 K. Dipolar-based sequences were used with cross-polarization (CP) steps with an amplitude ramp of 80-100% on ^1H and 15 kHz PISSARRO[297] decoupling during detection periods.

For J-based experiments, decoupled HSQC[182] sequences were used with a 4 and 3.35 ms INEPT transfer time for NH and CH transfers, respectively. NHHC experiments[298] were performed on [^{12}C,^{15}N]/[^{13}C,^{14}N]-mixed labeled samples in the matured state. Dipolar-based CP steps were used for magnetization transfer from ^{13}C to ^{1}H and from ^{1}H to ^{15}N. Proton-mixing times were 0, 1 and 2 ms.

3.2.6 *Molecular Dynamics Simulations*

Molecular dynamics (MD) simulations to probe the interaction between the YjeF_N domain and the Edc3 IDR were run with the g54a7 force field[299] and the GROMOS simulation package version 4.6.3.[300]. A 3D model of the dimeric YjeF_N domain was obtained using homology modelling[301] based on the structure of the human protein (PDB:3D3K)[291]. A linear 14-meric peptide corresponding to the IDR stretch M93-E106 was built in Pymol (pymol.org). For the starting state, peptides were placed within approximately 1 nm distance from the dimeric YjeF_N domain in a cubic aqueous box. The system was first energy minimized, afterwards equilibrated in an NVT ensemble for 100 ps with position restraints, then further equilibrated in an NPT ensemble for 100 ps with position restraints, and eventually freely evolved for 100 ns. In total, six replicates of 100 ns duration each with randomized initial peptide positions were run. Contacts between IDR segment M93-E106 and the YjeF_N domain were quantified with the g-mindist tool of the GROMACS simulation package using a contact cut-off of 8 Å. Contacts were counted over the last 50 ns of the trajectory. Since the YjeF_N domain is symmetric, the contacts for both monomers were added together. The final analysis is the summation of the intermolecular IDR - YjeF_N domain contacts from three replicas.

3.3 *RESULTS*

3.3.1 *The LSm domain of Edc3 is mobile in the matured state*

We initially examined the global dynamics of the matured state of full-length Edc3 (Edc3-FL) using scalar-based ssNMR experiments[158,286,292]. To this end, we recorded HC INEPT[295] (**Figure 3.1 B**) and C-C TOBSY[296] (**Figure 3.1 C**) spectra that reveal regions of the protein that are highly mobile. These spectra indicate that at least one subdomain of the full-length Edc3 protein remains mobile, even in the matured state. As we previously showed that the Edc3 LSm domain is not essential for the LLPS process of isolated Edc3[252], we speculated that this domain might remain flexible in the matured processing body state. To test this hypothesis, we repeated the scalar-based ssNMR experiments on matured processing bodies that were prepared from Edc3 protein lacking the LSm domain (Edc3-ΔLSm).

In these spectra (**Figure 3.1 B** and **C**; blue), we observed a near complete reduction of the dynamics that were visible for the wild-type Edc3 protein. In particular, backbone correlations were strongly reduced in two-dimensional CC TOBSY spectra of the matured states of the Edc3-ΔLSm protein compared to Edc3-FL (**Figure 3.1 C**). Many of these disappearing correlations in **Figure 3.1 C** matched with previous solution-state NMR assignments of the isolated LSm domain and the corresponding residues are indicated in **Figure 3.1 D** in green. Deletion of the LSm domain removes two proline residues from the amino acid sequence. Yet, for Edc3-ΔLSm a total of four proline resonances disappeared (**Figure 3.1 E**), suggesting that the presence of the highly mobile LSm domain increases the flexibility of a part of the IDR. From these data, we conclude that the Edc3 LSm domain remains highly flexible in the matured state of the Edc3 protein. This notion also implies that the LSm domain is accessible for interactions with other processing body proteins and that it can thus engage in interactions with Dcp2, as we previously showed[101].

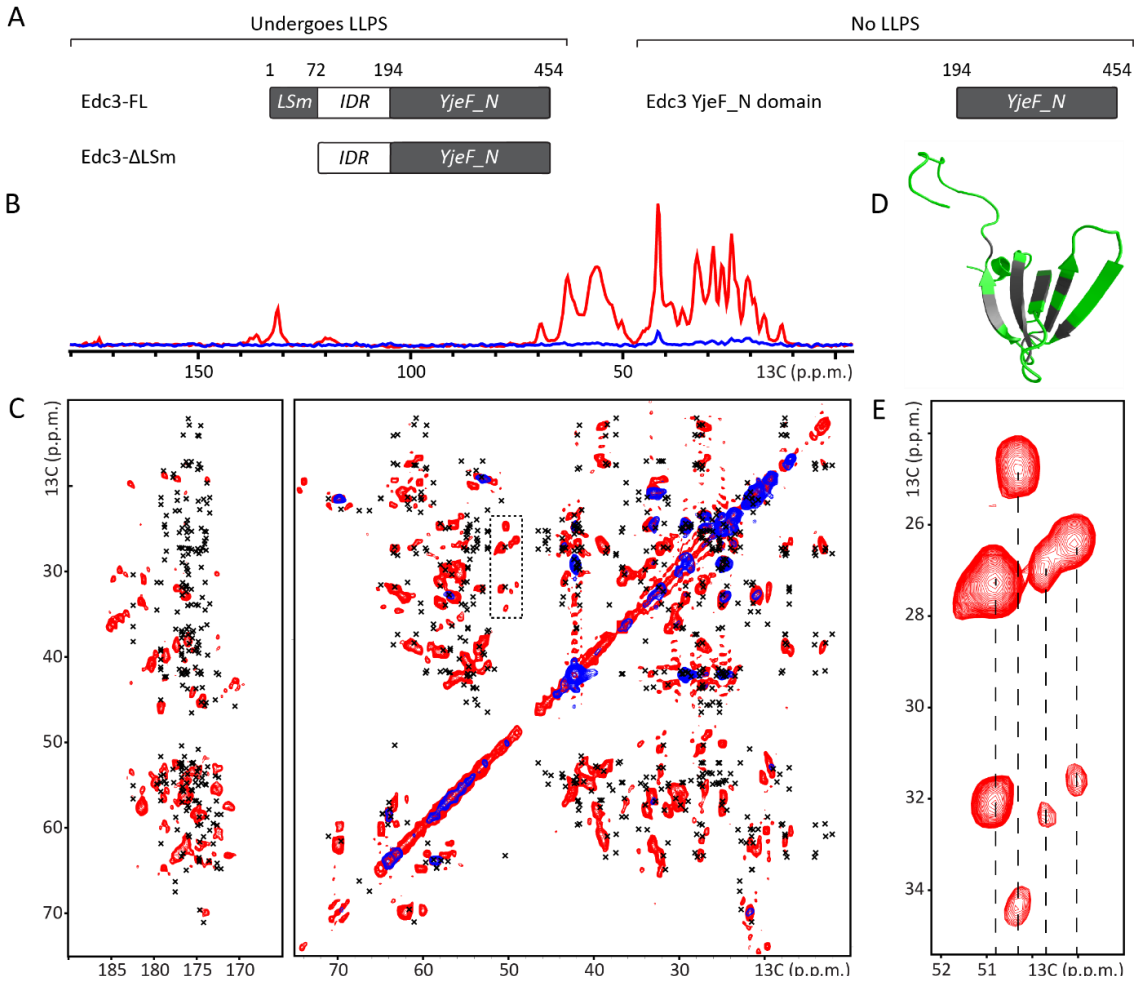

Figure 3.1: The LSm domain of Edc3 remains mobile after LLPS. (A) Overview of the Edc3 constructs used in this study. Residue numbers are indicated on top. **(B)** HC INEPT spectra highlighting dynamic regions in the full-length phase-separated Edc3 protein (Edc3-FL, red) and in the phase-separated Edc3 protein that lacks the N-terminal LSm domain (Edc3-ΔLSm, blue). **(C)** 2D CC TOBSY spectra of the phase separated Edc3-FL (red) and Edc3-ΔLSm (blue) proteins. The crosses indicate the solution chemical shifts of the LSm domain[101]. **(D)** The NMR structure of the LSm domain color-coded for residues where the solution state chemical shifts match with resonances in the CC TOBSY spectrum (green) and for residues where the solution state chemical shifts are not in a region where there are resonances observed in the CC TOBSY spectrum (gray). **(E)** Zoom on the proline region highlighted by the black dashed box in panel **C**, indicating four unique proline spin systems.

3.3.2 *The YjeF_N domain forms a rigid core in the matured state*

The C-terminal YjeF_N domain of Edc3 forms a strong dimer interface[291] (**Figure 3.2**) that does not dissociate over time and that is essential for the LLPS behavior of Edc3[252]. Solution-state NMR ^{1}H-^{15}N TROSY spectra of the isolated and soluble YjeF_N domain display a well-folded protein (**Figure 3.3 A**, green). To assess if the fold of the YjeF_N domain changes upon maturation of the phase-separated Edc3 protein, we recorded ^{1}H-detected dipolar NH ssNMR spectra on the matured state of the Edc3-ΔLSm protein (**Figure 3.3 A**, black). These spectra display Edc3 regions that are rigid on the time scale of our ssNMR experiment. As the Yjef_N domain comprised the majority of the Edc3-ΔLSm protein, we thus expected to observe mainly resonances from the dimerization domain.

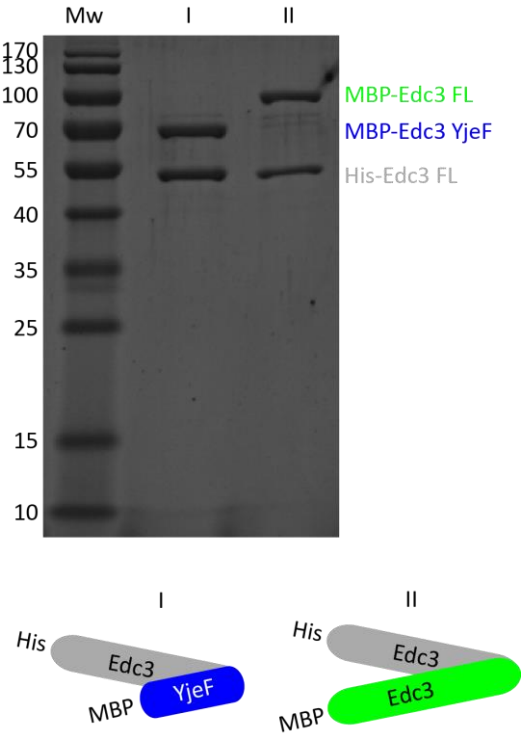

Figure 3.2: Edc3 forms a stable dimer in solution. His_6-tagged full-length Edc3 was co-expressed in *E. coli* with either the MBP-tagged Edc3 Yjef_N domain (I) or with MBP-tagged full-length Edc3 (II). The expressed proteins were purified in two steps. In step one, all His-tagged Edc3 complexes (dimers) were selected for using Ni-NTA chromatography. In step two, all MBP-tagged Edc3 complexes (dimers) were selected for using amylose affinity chromatography. Edc3 complexes that were purified this way contain at least one His-tag and at least one MBP-tag. The purified complexes were analyzed by SDS-PAGE and show a 1:1 ratio of a His-tagged and an MBP-tagged Edc3 monomer. This clearly shows that the Edc3 protein forms a stable dimer in solution through the Yjef_N domain. Mw: molecular weight marker in kDa.

Indeed, we found a remarkable overall agreement between the NH correlation spectrum of the isolated Yjef_N domain in solution and the matured form of the Edc3-ΔLSm construct (**Figure 3.3 A**, green versus black). The same agreement was observed for ^{1}H-^{13}C methyl TROSY[181] spectra recorded on the soluble YjeF_N domain as compared to a ^{1}H-detected dipolar CH ssNMR spectra recorded on the matured state of Edc3-ΔLSm (**Figure 3.3 B**, green versus black). Note that the solution-state NMR spectrum in **Figure 3.3 B** only comprises signals of methyl groups from Ala-β, Ile-δ_1, Met-ε, Val-γ and Leu-δ positions, as the other carbon and proton atoms are NMR-inactive.

Even in the absence of residue-specific assignments, this strong resemblance of solution- and solid-state NMR data suggests that the fold of the YjeF_N domain is largely conserved in the matured state of the Edc3 protein without major structural rearrangements. This notion rules out a scenario where the Yjef_N domain forms β-sheet-rich fibril-like structures after phase separation and maturation and is further supported by the strong correlation between ssNMR spectra of the matured phase separated Edc3-ΔLSm protein and of the lyophilized Yjef_N domain (**Figure 3.4**).

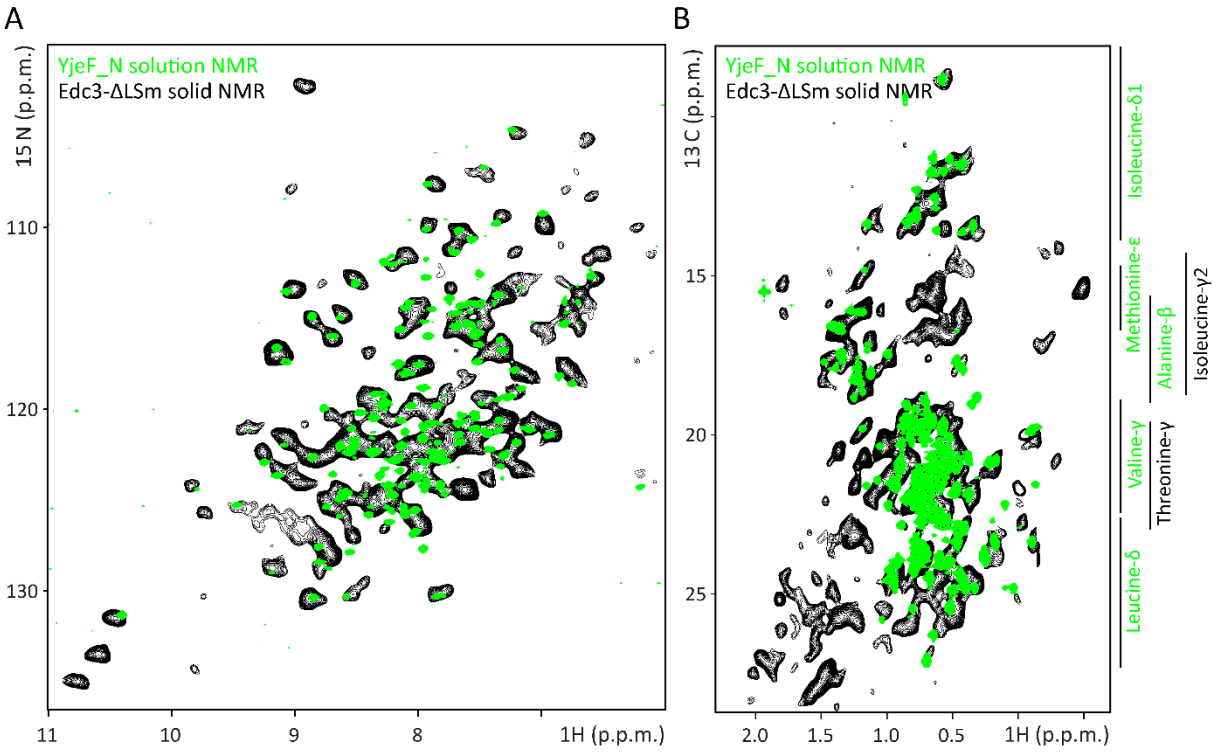

Figure 3.3: The fold of the YjeF_N dimerization domain is conserved after LLPS. Comparison of dipolar ssNMR NH spectra recorded on Edc3-ΔLSm (black) with (^{1}H-^{15}N or methyl ^{1}H-^{13}C) TROSY-based spectra recorded in solution on the YjeF_N domain (green). **(A)** ^{1}H-^{15}N region of the spectrum and **(B)** methyl region of the spectrum. The overall similarity between the spectra shows that the Edc3 Yjef_N domain is folded similarly in solution and in the matured state. As a reference, average ^{13}C chemical shifts for methyl-containing residues are indicated next to ^{1}H-^{13}C spectrum (note that Thr-γ, Ile- γ2 methyl groups and CH2 groups are not NMR-active in the solution-state sample).

Despite the strong correlation between the ssNMR spectrum of the Edc3 protein in the phase-separated state and the solution-state NMR spectrum of the isolated Yjef_N domain, there are some notable differences. This is mainly caused by the fact that the ssNMR spectrum was recorded on a protein construct that contains the IDR and the Yjef_N regions, whereas the solution-state NMR spectrum was obtained on a sample that only contains the Yjef_N domain. Note that it is not possible to record solution-state NMR spectra on an Edc3 construct that includes the IDR, as both Edc3-FL and Edc3-ΔLSm undergo LLPS and will transition into a matured state over time, which is incompatible with solution-state NMR methods. Due to the presence of the IDR in the sample that was used to record the ssNMR spectra, additional NMR resonances are expected and observed (**Figure 3.3**, compare black and green spectra). In addition, the IDR is known to interact with the Yjef_N domain in the phase separated protein (see below). This will result in chemical shift perturbations of a set of the Yjef_N resonances and thus in differences between the solution and ssNMR spectra. In addition, as the YjeF_N domain was expressed in D_2O for the solution-state NMR experiments, not all backbone NH protons might have back-exchanged during the purification process, which could result in absence of resonances in the solution-state ^{1}H-^{15}N NMR spectrum.

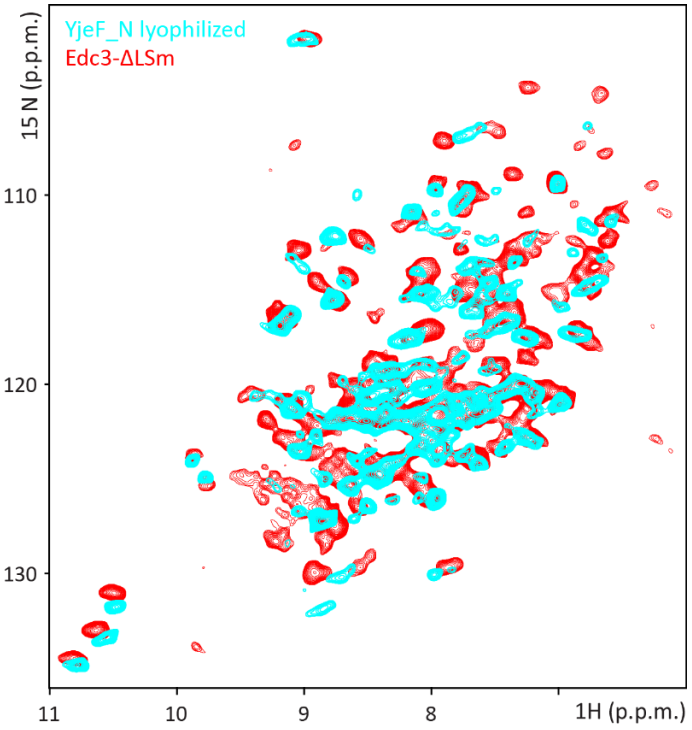

Figure 3.4: The fold of the YjeF_N domain in isolation is conserved in the matured state. Comparison of ^{1}H-detected dipolar NH spectra recorded on Edc3-ΔLSm in the matured state (red) and lyophilized YjeF_N domain (cyan) highlighting the conservation of the fold of the YjeF_N domain as seen by ssNMR. Differences between the two spectra can result from the IDR that interacts with the Yjef_N domain in the Edc3-ΔLSm construct.

Moreover, Ile-γ2 and Thr-γ methyl groups were not labeled for the solution-state NMR experiments and the corresponding resonances are thus not visible in the ^{1}H-^{13}C methyl TROSY spectrum. Finally, changes between the solution- and solid-state NMR spectra can result from differences in the measurement temperature (solid: 277 K, solution: 298 K) and sample preparation. Based on our experience on globular proteins, we however judge that these temperature effects only cause minor changes in the NMR spectra. In summary, based on a comparison of Yjef_N spectra of the Edc3 protein in solution and spectra in the matured state, we conclude that the structure of the Yjef_N domain is largely maintained after undergoing LLPS.

3.3.3 *Interactions between the IDR and the Yjef_N domain are important for phase separation of Edc3*

Previously, we have shown that the interaction between the Edc3 IDR and the Edc3 dimerization domain is important for the LLPS process of Edc3. In particular, we have identified two regions in the IDR that can interact with the Yjef_N domain[252]. These regions comprise two amino acid segments: K90-K111 (termed IDR1, **Figure 3.5 A**) and L159-L172 (termed IDR2, **Figure 3.5 A**). To assess the relative importance of these segments for the self-assembly process of Edc3, we performed LLPS assays with versions of the protein that lack either the first (IDR1) or the second (IDR2) disordered segment that interacts with the Yjef_N domain (**Figure 3.5 B**). We found that both segments in the IDR enhance LLPS.

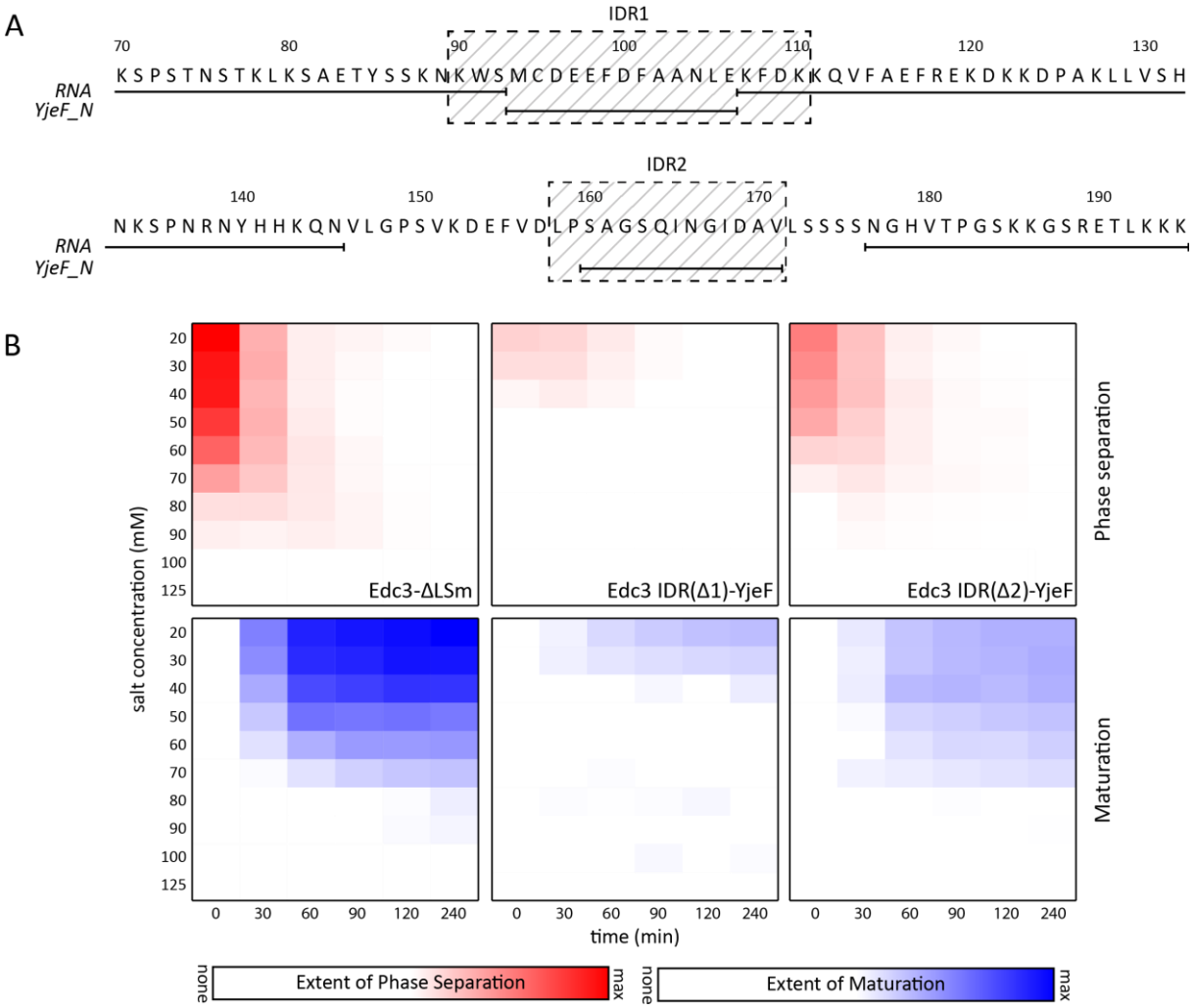

Figure 3.5: Two distinct regions in the IDR are important for phase separation. **(A)** Edc3 IDR sequence with highlighted residues that interact with the YjeF_N domain and RNA. **(B)** Assays that report on the LLPS (top panels) and maturation (bottom panels) of the Edc3 ΔLSm protein. The behavior of the protein was monitored over time (x-axis) at different salt concentrations (y-axis). The wild-type Edc3-ΔLSm construct exhibits a high degree of LLPS at low salt concentrations (top left panel). The droplets merge over time and form an insoluble gel-like matured phase (bottom left panel). Deletion of YjeF-interacting segments (middle and right panels) from the IDR results in a decrease of phase separation (top panels) and thus also of maturation (bottom panels). These data thus show that the IDR:YjeF_N interaction is important for Edc3 foci formation and maturation. IDR1 that contains residues 90-110 has a larger impact on LLPS than IDR2, that comprises residues 158-171.

The first segment has a stronger influence on LLPS than the second one, as removing IDR1 from the protein reduces LLPS to a larger degree than removing IDR2. These results underscore the degeneracy in the interaction networks that are responsible for cellular phase transitions and identify IDR1 and 2 as important interaction sites for LLPS. In support of this data, [^{13}C, ^{14}N] and [^{12}C, ^{15}N] mixed-labeled Edc3-ΔLSm samples were prepared in the matured state to probe intermolecular contacts between dimers. We observed intermolecular contacts between HN and HC atom pairs in NHHC experiments[298] verifying intermolecular contacts between the IDR and YjeF_N during LLPS (**Figure 3.6**).

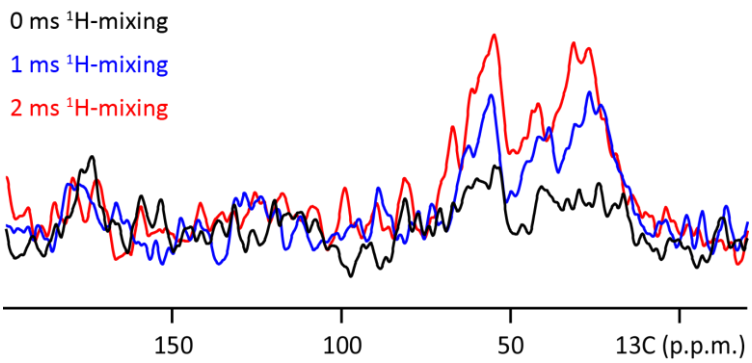

Figure 3.6: Edc3 dimers show intermolecular interactions during LLPS. NHHC spectra recorded on mixed ^{13}C-only labeled and ^{15}N-only labeled Edc3-ΔLSm with different proton mixing times shows the intermolecular interactions of Edc3-ΔLSm dimers in the matured state.

3.3.4 *Details of the IDR and Yjef_N domain interactions*

To address the interaction between the IDR and the Yjef_N domain with atomic detail, we turned to solution-state NMR techniques. In titration experiments we observed clear chemical shift perturbations in ^{1}H-^{15}N TROSY and ^{1}H-^{13}C methyl TROSY NMR spectra of the Yjef_N dimerization domain when the IDR is added in *trans* (**Figure 3.7 A** and **B**).

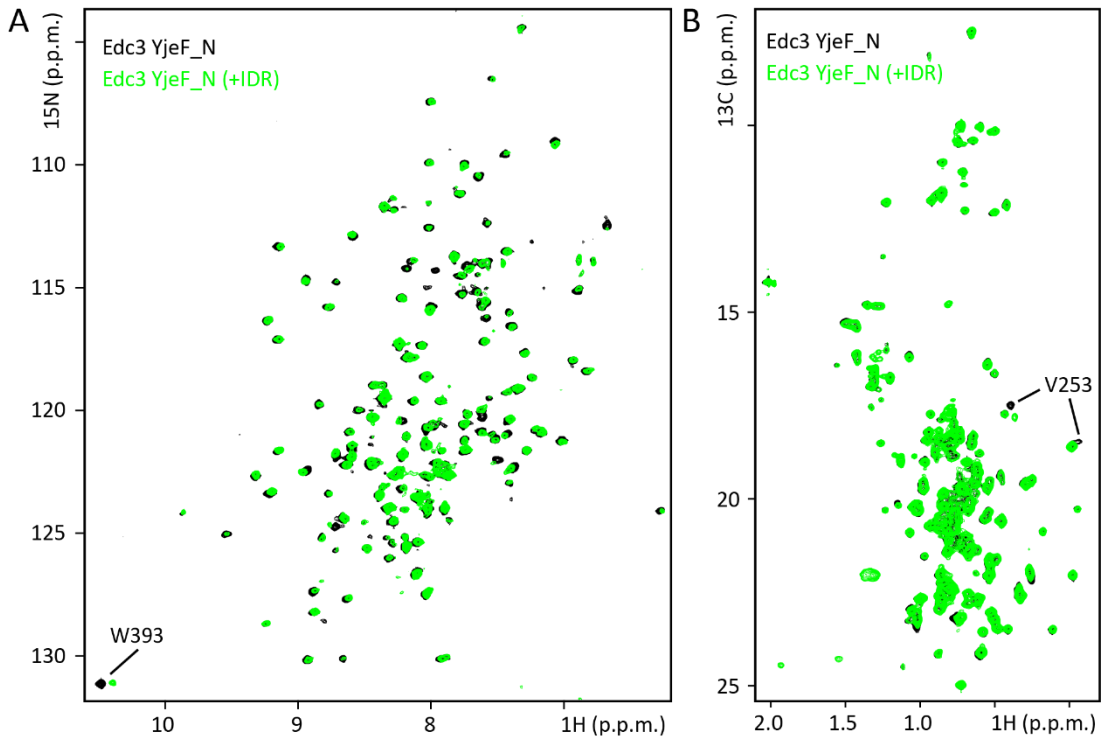

Figure 3.7: The Edc3 IDR interacts with the YjeF_N domain in solution. (A) Solution-state ^{1}H-^{15}N TROSY and **(B)** ^{1}H-^{13}C HMQC spectra of the YjeF_N domain in the absence (black) and presence (green) of the unlabeled IDR. The assignment of W393 and V253 are highlighted as these residues show a significant chemical shift perturbation upon interaction with the IDR.

To identify where the IDR interacts with the YjeF_N domain, we assigned several resonances in the ^{1}H-^{15}N and ^{1}H-^{13}C spectra through a mutational approach (**Figure 3.8 A**). In particular, we were able to identify that W393 (based on ^{1}H-^{15}N data) and V253 (based on methyl TROSY data) in the dimerization domain are both part of the interface with the Edc3 IDR.

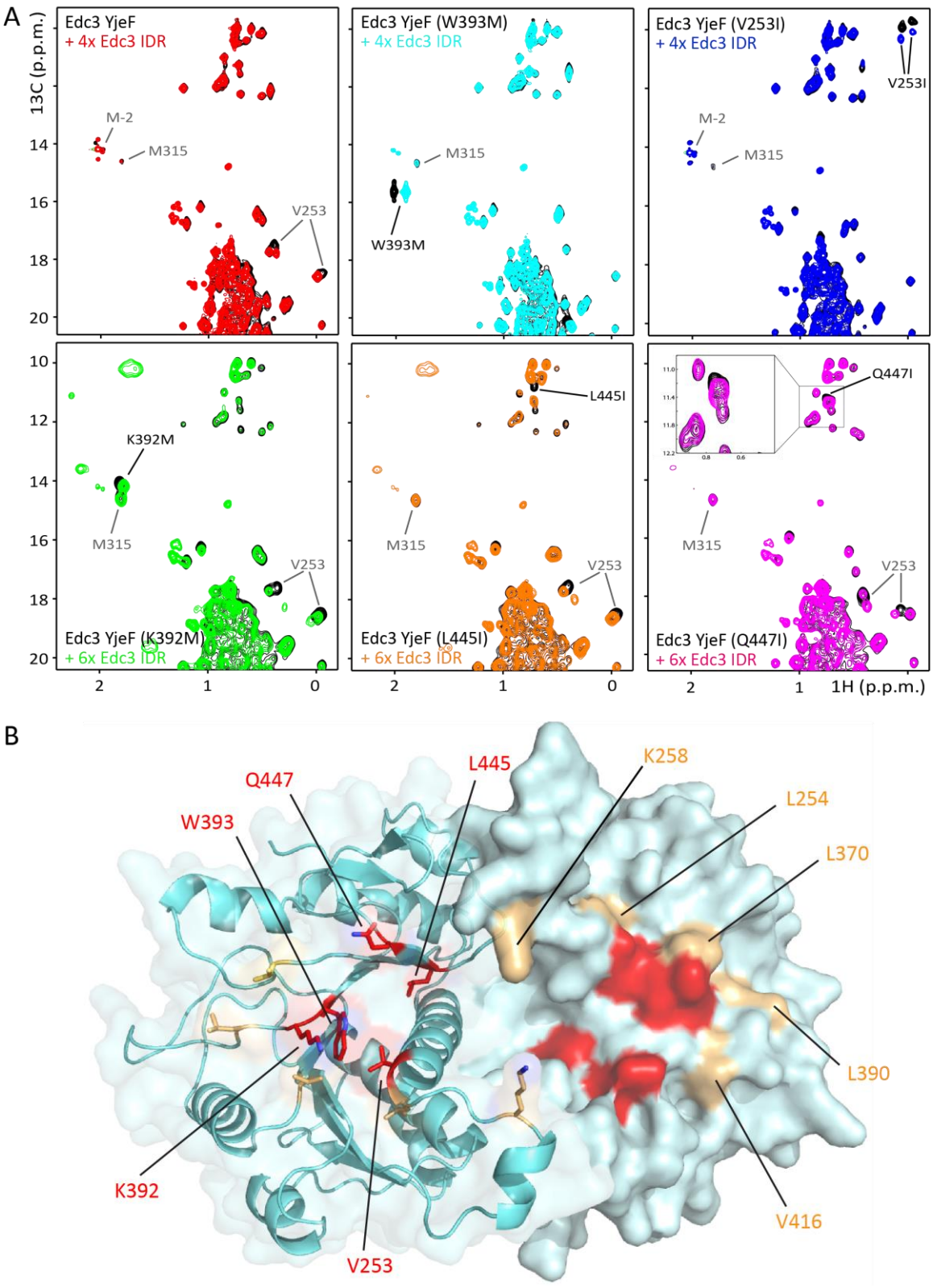

Figure 3.8: The IDR interacts directly with the YjeF_N domain. (A) The binding site of the Edc3 IDR on the Edc3 Yjef_N domain was mapped using methyl TROSY NMR spectra. Specific residues (W393M (cyan), V253I (blue), K392M (green), L445I (orange) and Q447I (magenta)) on the Yjef_N surface were mutated into methyl group containing amino acids to add NMR-active probes. These mutant proteins were subsequently used in titration experiments were the NMR-inactive Edc3 IDR was added in a 4- to 6-fold molar excess. Chemical shift perturbations of the introduced methyl groups indicate if the corresponding residue is part of the Edc3 IDR binding site on the Edc3 Yjef_N domain. Mapping of binding sites in this manner has proven to be highly sensitive and accurate[214,243]. **(B)** The identified IDR binding site is highlighted on the Yjef_N homology model. Residues of the YjeF_N domain that interact with the IDR are colored red while non-interacting residues are colored yellow. M315 is located at the back of the YjeF_N dimer in the represented orientation, outside of the binding site of the IDR.

As W393 and V253 are close in space on the surface of the Yjef_N domain, they are part of the same binding pocket. To define additional residues that are part of this pocket, we took advantage of a slightly modified version of the methionine scanning approach[214,243]. In this method, methyl group-containing residues were introduced in the spatial vicinity to W393 and V253. These methyl groups were subsequently used as probes for the interaction of the YjeF_N domain with the IDR. Based on this strategy, we were able to identify K392, L445 and Q447 as part of the IDR binding pocket (**Figure 3.8 A**). The introduced methyl groups did not alter the interaction mode between the Yjef_N domain and the IDR as the naturally occurring methyl groups of V253 underwent the same CSPs in the WT protein and the mutants. At the same time, we showed that L254, K258, L370, L390 and V416 are located outside the IDR binding interface (**Figure 3.8 B**), as these reporter methyl resonances are insensitive to the IDR, while the methyl groups of V253 remain effected by the IDR:Yjef_N interaction. In short, we find that the Edc3 IDR interacts with the dimerization domain at two specific pockets on the surface, one on each protomer.

To shed further light on how the IDR interacts with the YjeF_N domain, we used atomistic molecular dynamics (MD) simulations (**Figure 3.9 A**). To this end, we used a homology model of the YjeF_N domain and performed six simulations of the complex formation between the YjeF_N domain and a short peptide that corresponds to IDR segment M93-E106 (IDR1). For each MD simulation, two 'IDR peptides' were randomly placed in the aqueous phase above the YjeF_N domain, shortly minimized and equilibrated with position restraints, and the system was freely evolved for 100 ns.

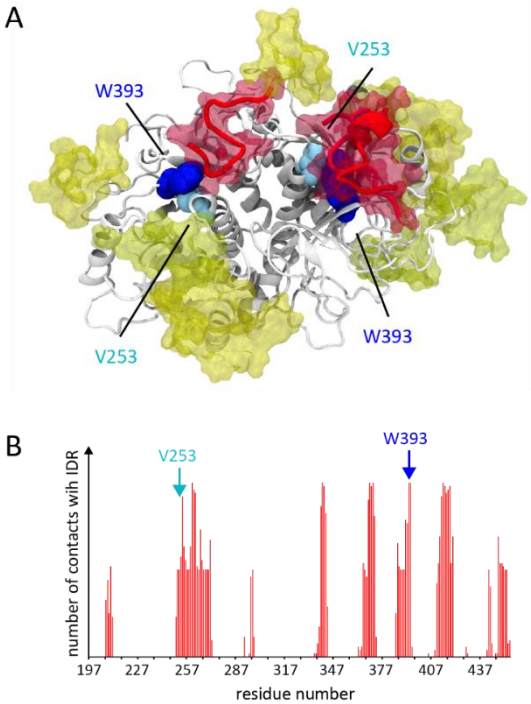

Figure 3.9: Molecular dynamics (MD) simulations confirm a direct interaction between the Edc3 IDR and the YjeF_N domain.

(A) MD simulations on the complex formation between the YjeF_N dimer (in white) and peptides corresponding to the first YjeF_N-interacting segment of the IDR M93-E106 (IDR 1; red). In the MD simulation, for three out of six replicas, the IDR peptides directly interacted with the residues V253 (cyan) and W393 (blue), that were identified as YjeF_N-interacting hotspots in our solution-state NMR data. The figure shows the endpoints of the simulations after 100 ns.

(B) The analysis of molecular dynamics simulations results in the identification of the binding epitope that was also determined based on solution-state NMR experiments. For the simulations where the IDR bound in agreement with the solution-state data (in red in A), the contacts between the IDR and the Yjef_N domain were summed up over the last 50 ns of the MD simulations.

Interestingly, in three of the six runs, we observed that the peptides (in red, **Figure 3.9 A**) bound the YjeF_N domain in direct proximity of V253 and W393 (in cyan and blue), corroborating the solution-state NMR data. Peptides that bound in proximity of these residues also remained stably associated with the YjeF_N domain (**Figure 3.9 B**).

3.3.5 *RNA interacts with the IDR and thereby increases rigidity*

Previously, we observed that RNA significantly enhances the LLPS tendency of the Edc3 protein as three distinct regions on the IDR can directly interact with RNA[252] (**Figure 3.5 A**). To obtain insights into the molecular mechanism by which RNA is embedded into the matured state of phase-separated Edc3, we conducted ssNMR experiments on Edc3-ΔLSm protein in the presence of RNA. Interestingly, based on scalar-based CH and NH spectra we observed that the incorporation of RNA into the protein-rich Edc3 phase results in a significant loss in ssNMR signals, consistent with a reduction of the internal dynamics in the matured phase (**Figure 3.10 A** and **B**; compare blue and green spectra). In particular, we observe that many resonances that are typical for unstructured backbone and side chain regions disappear upon interaction with RNA (**Figure 3.10 B**, green). These data thus directly report on the interaction between RNA and the IDR which results in a loss of mobility of the IDR (**Figure 3.10 B**).

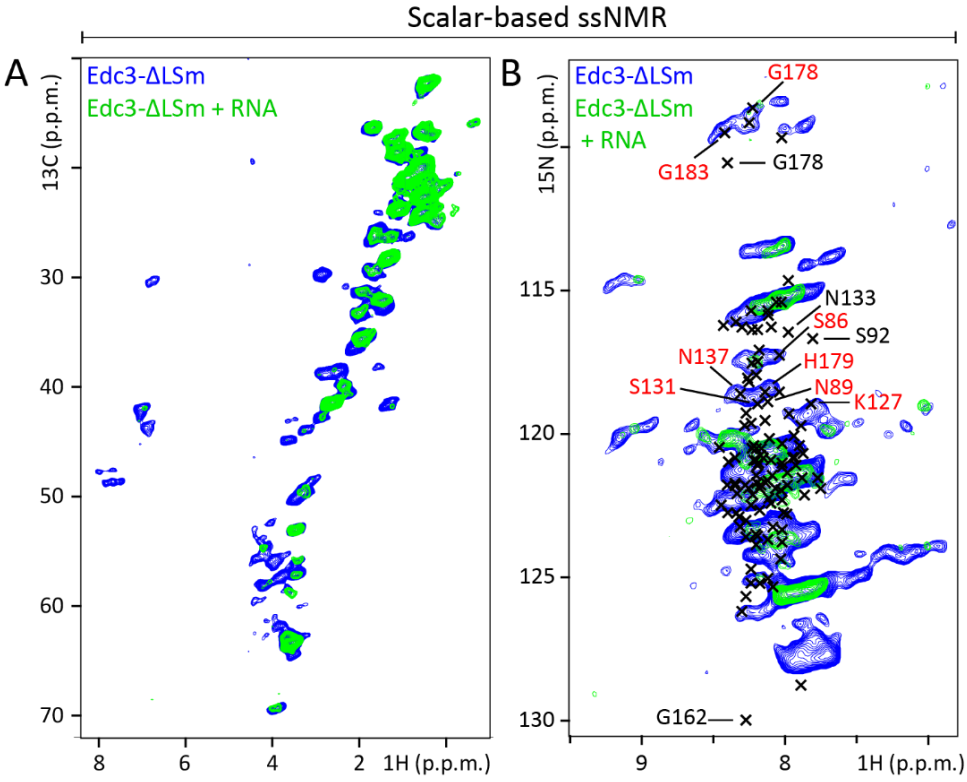

Figure 3.10: RNA interacts with the IDR and results in reduced dynamics in the matured state. (A, B) Scalar-based CH and NH spectra that detect molecular elements with fast nanosecond motions recorded on Edc3-ΔLSm in the absence (blue) and presence (green) of RNA, respectively. In panel **B**, the solution-state ^{1}H-^{15}N assignments of the IDR are indicated with black crosses. Putative interaction sites with RNA (red labels) or the YjeF_N domain (black labels) are indicated.

Using assignments obtained from solution-state NMR[252] (black crosses), the regions where signals disappear upon RNA incorporation (red labels) were identified to include S86, K127, S131, N137 and G183. This confirms our previous solution-state NMR analysis on the putative IDR:RNA interaction site (**Figure 3.5 A**), as these residues were identified to directly interact with RNA. Furthermore, many resonances of the IDR that are known to interact with the Yjef_N domain are not observed in these ssNMR spectra, as these amino acids are rigid both in absence and in presence of the RNA.

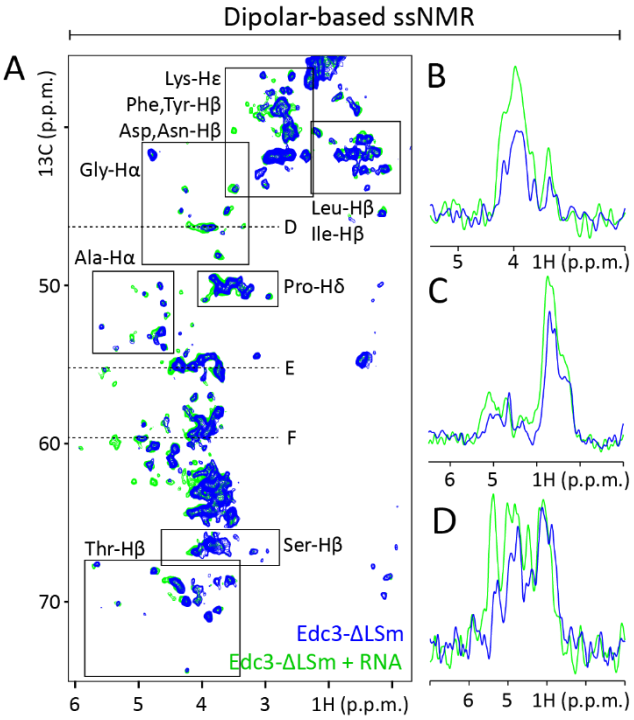

Figure 3.11: Dipolar-based ssNMR experiments report on the rigidification of the Edc3 protein in the presence of RNA. (A) Dipolar CH spectra that detect rigid, immobile entities were recorded on Edc3-ΔLSm in the absence (blue) and presence (green) of RNA. Several boxes indicate predicted chemical shifts for specific amino acids. **(B-D)**, 1D slices of highlighted regions in **A**.

Table 3.3: Amino acid distribution and occurrence for the Edc3 YjeF_N domain and IDR

	YjeF	*IDR*
Ala	26	7
Arg	6	3
Asn	14	9
Asp	10	9
Cys	3	1
Gln	16	3
Glu	8	8
Gly	18	6
His	7	4
Ile	18	2
Leu	33	8
Lys	10	20
Met	1	1
Phe	9	6
Pro	11	6
Ser	19	18
Thr	21	5
Trp	3	1
Tyr	3	2
Val	23	7

To independently confirm the rigidification of the IDR upon interaction with RNA, we recorded dipolar CH experiments on the matured state of the Edc3-ΔLSm protein in the absence and presence of RNA. In those experiments, we observe that several peaks appear upon addition of RNA, confirming the hypothesis that residues of the IDR are stabilized by the RNA (**Figure 3.11**, green). The corresponding correlations (for example at 5.39 ppm ^{1}H and 59.9 ^{13}C ppm) are consistent with lysine side chain and serine/threonine backbone ^{1}H-^{13}Cα signals that are prominently present in the IDR segments that interact with RNA in solution (**Table 3.4**). In conclusion, our solid-state NMR data suggests that RNA is recruited to the dynamic and unstructured IDR in Edc3, which results in decreased internal mobility.

3.4 *DISCUSSION*

LLPS is increasingly recognized as a cellular process that allows for the organization of cellular compartments across different spatial and temporal scales. A prominent example of such systems are RNP granules that contain proteins involved in translational repression and mRNA degradation[252].

Currently, the mechanisms by which LLPS modulates cellular function is not understood in sufficient detail. One potential consequence of LLPS is that the catalytic activity of the embedded enzymes is effected, potentially through changes in the local concentrations or accessibility of substrates[252]. To shed light on these aspects, it is essential to gain atomic-level insight into the structural organization of LLPS proteins. Here, we demonstrate the power of combining solid- and solution-state NMR to study the process of phase separation events on the structural and dynamical level and at atomic resolution.

Our results on the central processing body protein Edc3 reveal that the different domains exhibit different levels of structural organization and dynamics after LLPS and a second transition into a matured phase. Firstly, we find that the N-terminal LSm domain remains largely dynamic in the matured state, which leaves this domain accessible for contacts with additional processing body proteins including the Dcp1:Dcp2 mRNA decapping complex. Secondly, we find that the C-terminal YjeF_N domain largely retains its structure when the Edc3 protein undergoes a transition from the soluble to the matured phase separated state. This is in strong contrast to other proteins that have been observed to form insoluble amyloidic fibril structures after phase separation[148]. The minor structural changes in the fold of the Edc3 protein reflects the fact that proteins that reside in cellular processing bodies can rapidly dissociate from these foci.

Within the phase-separated state, the Edc3 Yjef_N domain interacts with the IDR of Edc3. This transient interaction is likely intermolecular as it increases the network of interactions leading to LLPS. Finally, we observe that the Edc3 IDR interacts with RNA in the matured state. These interactions involve distinct segments of the IDR and result in enhanced phase separation as the RNA extends and tightens the intermolecular interaction network.

Based on our data, we can propose a model for the formation of the matured state of the Edc3 protein (**Figure 3.12**). In this model, the dimeric soluble Edc3 protein engages in a number of intermolecular interactions that involve RNA. These contacts result in the formation of an infinite interaction network that drives the transition of the Edc3 protein from the dimeric soluble form into a dense protein state that then further matures into an insoluble state.

Interestingly, our data clearly show that the interactions between Edc3 and RNA that take place between the soluble components are, to a large degree, preserved in the matured phase-separated state. The structural differences upon LLPS are thus small, which facilitates the reversibility of the process and necessitates only small cellular perturbations to induce phase separations (**Figure 3.12**).

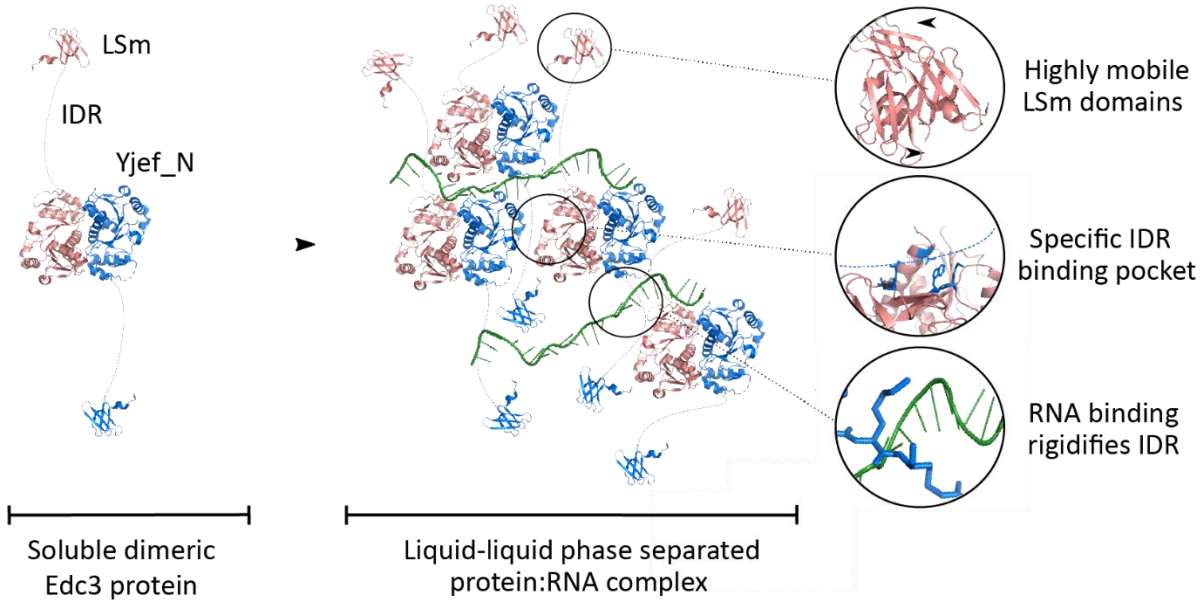

Figure 3.12: Model of Edc3- and RNA-mediated processing body formation. The soluble dimeric Edc3 protein (blue and red) contains an N-terminal LSm domain that is linked through an IDR with the C-terminal dimeric Yjef_N domain (left). Intermolecular interactions between RNA (green), the IDR and the Yjef_N domains (solution-state NMR) result in the establishment of an infinite interaction network that leads to liquid-liquid phase separation of the RNP. The LLPS proteins mature over time to form a gel-like state. Within this matured phase of processing bodies, the LSm domains remain highly flexible (solid-state NMR) and can thus interact with other processing body proteins (top circle). Based on solution-state and solid-state NMR data, we have shown that the Edc3 YjeF_N dimerization domain remains dimeric and interacts specifically via a surface patch with an Edc3 IDR of another Edc3 dimer (middle circle). The RNA interacts with the Edc3 IDR, which results in a rigidification of the unstructured region (bottom circle).

Our work paves the way for future studies that address the exact structural changes that take place upon LLPS and those that result in the formation of the irreversible assemblies. Based on our data, these latter changes are small, but they will, nevertheless, have a significant effect on cellular homeostasis, as the maturation of cellular condensed protein phases needs to be prevented. We envision that our approach to study the structure and dynamics of LLPS complexes will be able to provide important insights for the current context and will also be applicable for other LLPS systems where irreversible assemblies are closely linked to fibril formation and disease.

CHAPTER 4 *Deciphering the contributions of molecular interactions that lead to liquid-liquid phase separation of the conserved DEAD-box protein Dhh1*

Unpublished results. Remco Sprangers conceived the project. Experiments were designed by R.S. and Stefan Schütz. S.S. performed all experiments. R.S. and S.S. analyzed and interpreted data. S.S. prepared figures and drafted the manuscript, which has been revised by R.S.

4.1 *INTRODUCTION*

DEAD-box helicases[302] belong to a large family of non-processive ATP-dependent RNA helicases that are conserved from bacteria to humans[303]. *S. cerevisiae* Dhh1 belongs to the Dhh1/DDX6 sub-family of DEAD-box helicases[304] and is involved in translational repression, processing body (P-body) formation and mRNA decay[75,88,95,305–307]. Orthologs of Dhh1 are found in all higher eukaryotes[271] such as *S. pombe* (Ste13)[308], *Caenorhabditis elegans* (CGH-1)[309], *Drosophila melanogaster* (Me13B)[310], *Xenopus leavis* (Xp54)[311] and mammals (RCK/p54 or DDX6)[312–314].

Dhh1 contains two conserved central RecA-like[315] domains that are connected via a short flexible linker sequence. The two RecA-like domains form the so-called helicase core, which is flanked by disordered N- and C-terminal extensions (NTE and CTE, respectively). The sequence of these extensions are not conserved among the members of the Dhh1/DDX6 sub-family of DEAD-box helicases[303]. In *S. cerevisiae* Dhh1, the protein discussed in this work, the NTE comprises about 30 residues and is enriched in asparagine. The CTE contains about 80 residues and is highly enriched in proline and glutamine residues[146], reminiscent of a prion-like domain[316,317].

The crystal structure of the apo Dhh1 helicase core reveals an open but compact conformation that is rigidified by contacts between the two RecA-like domains (**Figure 4.1 A** and **B**)[318]. It has been shown that RNA binding to Dhh1 requires domain interactions, as mutations of the interdomain contacts result in weaker RNA binding[318]. In the absence of ATP, RNA binding does not induce conformational changes in Dhh1 as has been concluded from tryptic digestion experiments[318]. Contrarily, ATP alone is sufficient to induce a conformational change and further compaction is achieved in the presence of both ATP and RNA[318]. Based on structures of other DEAD-box helicases, it has been suggested that the interaction of Dhh1 with RNA and ATP leads to the formation of a closed state of the enzyme (**Figure 4.1 C**)[318].

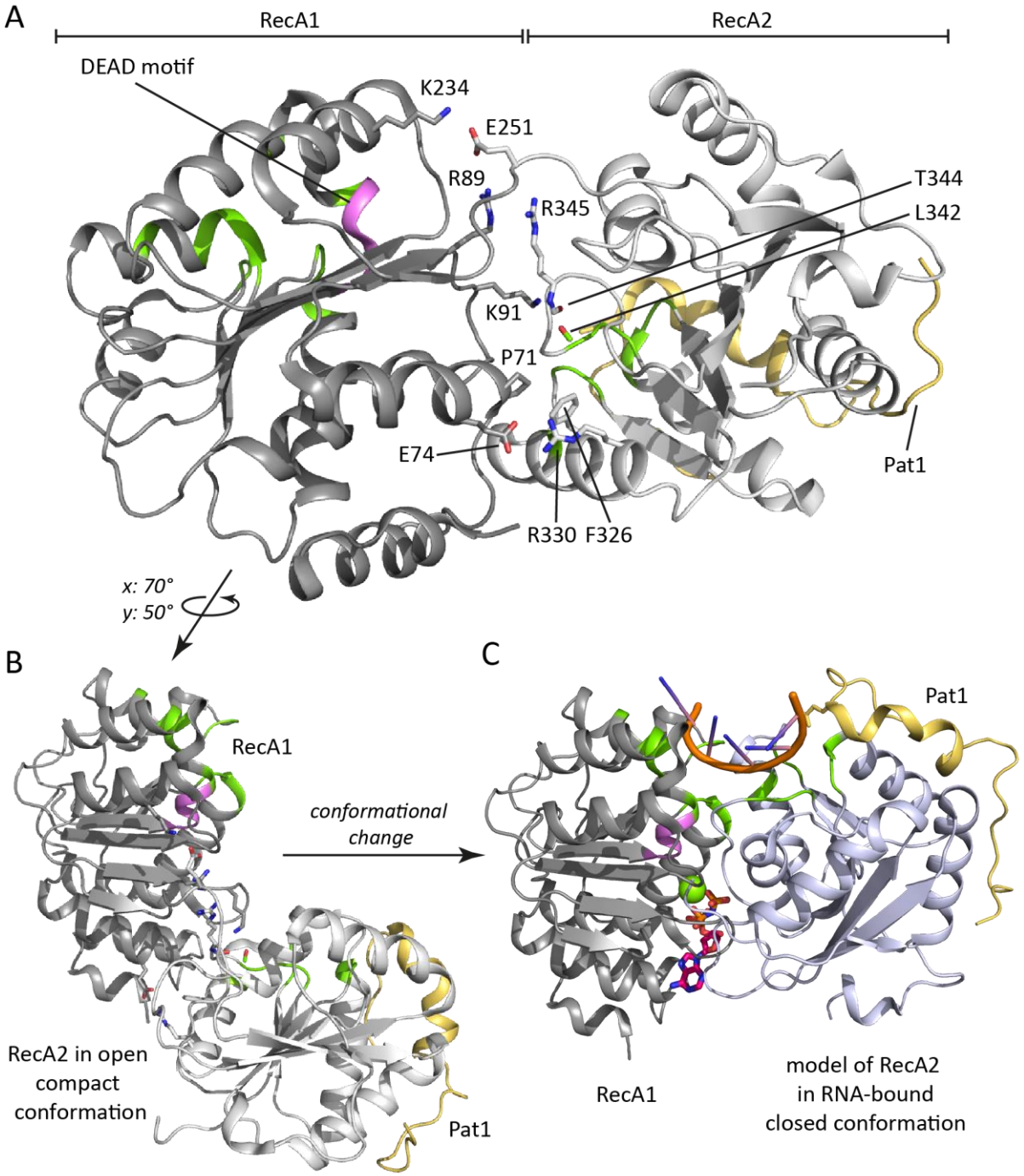

Figure 4.1: Crystal structure of the *S. cerevisiae* Dhh1 helicase core. (A) In the apo form (pdb: 1s2m), Dhh1 adopts an open but compact conformation with several contacts between the RecA1 (dark grey) and RecA2 (light gray) domains. These contacts interfere with the formation of a continuous RNA-binding surface (green). Binding of the Pat1 protein (yellow) to the Dhh1 RecA2 domain does not interfere with the interdomain contacts, but with crystal packing. The position of Pat1 was obtained from a superposition of apo Dhh1 (pdb: 1s2m) with Dhh1 in complex with Pat1 (5-79) (pdb: 4brw). Residues involved in the interdomain contacts are indicated and depicted in stick representation. The DEAD-box motif is highlighted in magenta. **(B)** Same as **(A)** but rotated as indicated. **(C)** Model of Dhh1 in its RNA- and ATP-bound form. The model was obtained by a superposition of the RecA1 domain from apo Dhh1 and of the RecA2 domain from the Dhh1:Pat1 complex with the structure of *D. melanogaster* Vasa (pdb: 2db3) in complex with an oligo-U RNA (orange), magnesium (green sphere) and AMPPNP (hot pink sticks). The orientation of the Dhh1 RecA1 domain is as in **(B)**. To highlight the domain reorientation in the closed conformation, the RecA2 domain is in light blue compared to light gray in apo Dhh1 (**A** and **B**). In the ATP- and RNA-bound form, Dhh1 residues involved in RNA-binding (as inferred from the Vasa:RNA:ATP complex structure) form a continuous interaction surface, while the interdomain contacts are disrupted. The DEAD-box motif is in proximity to the Mg^{2+}-ion that coordinates the β- and γ-phosphate groups of ATP. Note that Pat1 clashes with the RNA in the closed conformation, which provides a rationale for the mutually exclusive interactions of Pat1, Edc3 and RNA with the Dhh1 RecA2 domain.

In the closed state, the ATPase activity and the affinity for single-stranded RNA is increased in other DEAD-box helicases such as *D. melanogaster* Vasa[319–322]. Consistently, the low intrinsic ATPase activity that is associated with the rigid open conformation of Dhh1 increases about 10-fold in the presence of RNA and about a further 2.5-fold when the interdomain contacts are disrupted[323]. On the other hand, the RNA-binding affinity was found to be independent of ATP[323]. Notably, helicase activity has never been demonstrated for the purified Dhh1 enzyme[164,318,323].

In vivo, members of the Dhh1/DDX6 subfamily of DEAD-box helicases have a general function in mRNA storage and translational control[311,324–328]. The role of Dhh1 in translational repression likely results from its interaction with the CCR4-Not deadenylation complex[329,330] and by interfering with the assembly of the 48S preinitiation complex[88]. In addition, Dhh1 can interfere with translation elongation by binding to sub-optimal codons, which can trigger the degradation of the transcript[74].

Dhh1 interacts with a large number of proteins that are involved in mRNA turnover. First, Dhh1 binds to the CCR4-Not deadenylation complex via a direct interaction with the Not1 MIF4G domain, that also serves as a scaffold to tether the Pop2/Caf1 and Ccr4 deadenylases[331–334]. Second, Dhh1 is recruited to the mRNA degradation machinery through interactions between the Dhh1 C-terminal extension and the main decapping activator Dcp1[99,330]. Third, Dhh1 interacts directly with Dcp2, without influencing the activity of the mRNA decapping enzyme[95]. And fourth, the RecA2 domain of Dhh1 can recognize FDF, FDK and FG motifs that are present in the decapping factors Edc3, Pat1 and Scd6[105–108]. These motifs interact with specific, hydrophobic binding sites on the surface of the helicase (**Figure 4.2**). Notably, binding of Pat1 or Edc3 to Dhh1 was found to interfere with RNA binding due to overlapping binding sites on the surface of the RecA2 domain[105].

Dhh1 is a major component of cytosolic processing bodies[113,259] and recruited to these foci through interactions with the Edc3 protein[144]. The Dhh1 protein is then able to recruit the deadenylases Pop2/Caf1 and Ccr4[335]. ATP- and RNA-binding is required for Dhh1 localization to P-bodies, as the disruption of either binding site significantly reduces P-body formation[164]. Notably, the ATPase activity of Dhh1 is required for the disassembly of *in vivo* P-bodies, as mutations in the active site result in a permanent presence of P-bodies[164].

The mechanisms that result in the formation of processing bodies have been extensively studied *in vitro*, based on bottom-up experiments that can detect the propensity of protein complexes to undergo spontaneous liquid-liquid demixing[103,252]. It was found that the *in vitro* phase separation of Dhh1 depends on the presence of ATP and single-stranded RNA[164]. Consistent with this notion, RNase treatment resulted in the dissolution of *in vitro* Dhh1 droplets[164].

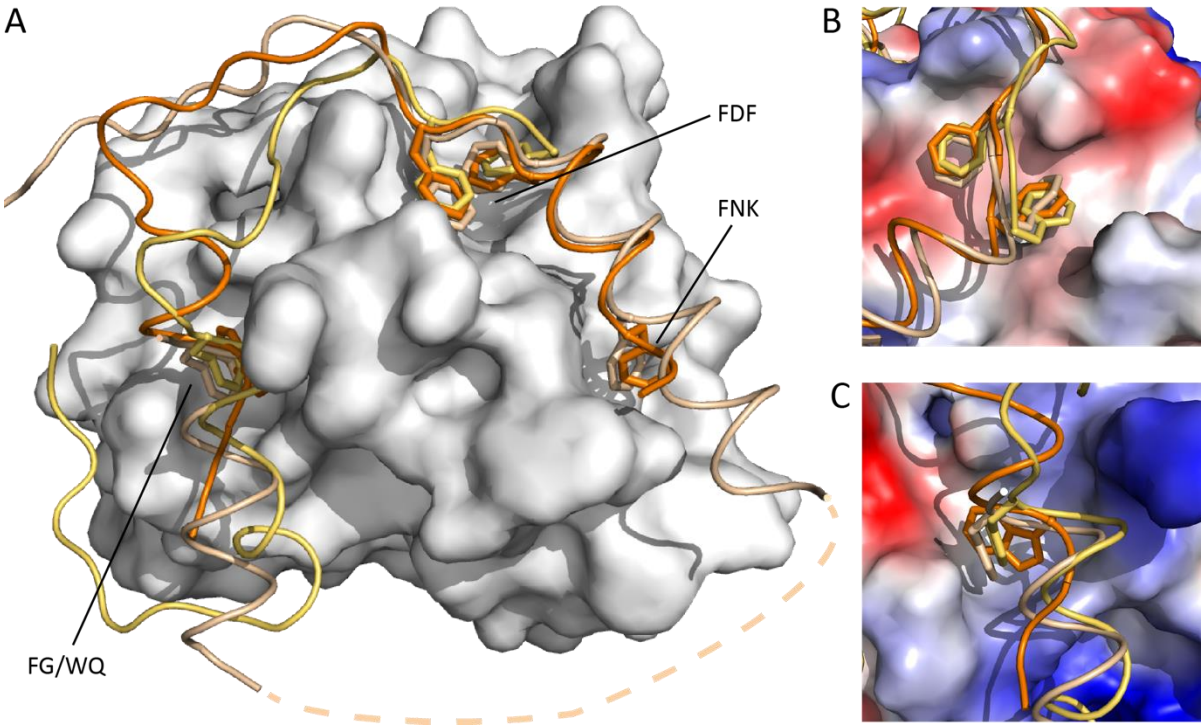

Figure 4.2: Decapping factors interact with the Dhh1 RecA2 domain. (A) Surface representation of the Dhh1 RecA2 domain (pdb: 4brw, light gray) in complex with peptides of Edc3 (orange), Pat1 (yellow) and LSm14/Scd6 (beige). For the model, the RecA2 domains of the Dhh1:Pat1 (pdb: 4brw) and DDX6:LSm14 (pdb: 6f9s) complexes were superposed on the RecA2 domain of the Dhh1:Edc3 complex (pdb: 4bru). In the DDX6:LSm14 crystal structure, LSm14 residues 209-254 (dashed line) are lacking electron density. This linker sequence connects the FG and FNK motifs. The superposition reveals a conserved binding mode of several mRNA decay factors and P-body components to the Dhh1 protein. **(B, C)** Zoom on the FDF **(B)** and FG/WQ motif **(C)** that bind to hydrophobic pockets on the RecA2 domain. Coloring of the peptides is as in **A**. The surface of Dhh1 is colored according to its electrostatic surface potential. Surface potentials were calculated in pymol 2.3 (Schrödinger, LLC; https://pymol.org/2/) and color-coded from red (negatively charged) to blue (positively charged).

Several intermolecular interactions modulate the Dhh1 LLPS process. On the one hand, the Not1 MIF4G domain stimulates the disassembly of Dhh1 clusters *in vitro* and triggers disassembly of P-bodies *in vivo*[164]. In agreement with that, mutations in the Not1:Dhh1 interface lead to a constitutive assembly of cellular P-bodies[164]. On the other hand, the decapping factor Pat1 enhances *in vitro* Dhh1 phase separation in the presence of ATP and RNA and promotes *in vivo* P-body formation through a direct interaction with the Dhh1 helicase core[152].

Here, we study the liquid-liquid phase separation process of Dhh1 at an atomic level. Based on a bottom-up approach with recombinant proteins and well-defined RNA, we quantify how specific intermolecular interactions contribute to the LLPS propensity of Dhh1. In summary, we establish that the isolated Dhh1 helicase core also undergoes liquid-liquid phase separation in a manner that is independent of ATP or RNA. Based on a large number of mutations in Dhh1 and on methyl TROSY NMR spectroscopy we identified residues in Dhh1 that are important for the phase separation process. In addition, we show that the LLPS of the Dhh1 core is enhanced by ATP, RNA and the N- and C-terminal extension of Dhh1. Our studies thus shed light on molecular principles that result in LLPS of proteins that contain both folded domains and intrinsically disordered regions.

4.2 *MATERIALS AND METHODS*

4.2.1 *Protein expression and purification*

The genes for full-length or truncated constructs of Dhh1 from *Saccharomyces cerevisiae* and *Chaetomium thermophilum* were cloned into modified pET vectors that carry an N-terminal TEV-cleavable His_6-tag. The genes for the first and second RecA-like domain of Dhh1 (RecA1 and RecA2, respectively), the N-terminus of Pat1 (Pat_N) and the intrinsically disordered region (IDR) of Edc3 from *S. cerevisiae* were cloned into modified pET vectors that carried an N-terminal TEV-cleavable His_6-GST-tag. Point mutations, deletions and insertions were introduced using standard site-directed mutagenesis or enzymatic assembly methods[336,337] (**Table 4.1**).

Table 4.1: Protein constructs used in this study

Protein / Construct	**Residues / Mutations**	**Solubility / Purification tag**	**Internal reference**
S.c. Dhh1	1-506	N-His_6-TEV	#1872, #1906
S.c. Dhh1 (Dhh1ΔN)	31-506	N-His_6-TEV	#1873
S.c. Dhh1 (Dhh1ΔC)	1-425	N-His_6-TEV	#1883
S.c. Dhh1 (ΔNC / core)	31-425	N-His_6-TEV	#47, #1907
S.c. Dhh1 (RecA1)	31-251	N-His_6-GST-TEV	#1696
S.c. Dhh1 (RecA2)	250-425	N-His_6-GST-TEV	#1697
S.c. Dhh1 (stretch1)	31-425 (core), 316-KMPQAARNR-324	N-His_6-TEV	#1874
S.c. Dhh1 (stretch1+2)	31-425 (core), 316-KMPQAARNR-324, 328-DFRNGVCRN-336	N-His_6-TEV	#1912
S.c. Dhh1 (stretch1+3)	31-425 (core), 316-KMPQAARNR-324, 407-HPIPAEIPKNLYVYENP-423	N-His_6-TEV	#1914
S.c. Dhh1 (stretch3)	31-425 (core), 407-HPIPAEIPKNLYVYENP-423	N-His_6-TEV	#1936, #1957
S.c. Dhh1 (AAtoHP)	31-425 (core), A407H, A408P	N-His_6-TEV	#1958
S.c. Dhh1 (DDtoPP)	31-425 (core), D414P, D423P	N-His_6-TEV	#1958
S.c. Edc3_idr	86-125	N-His_6-GST-TEV	#1887
S.c. Pat_N	5-79	N-His_6-GST-TEV	#1905
C.t. Dhh1	27-421 (core)	N-His_6-TEV	#1903
S.c. RecA1 – *C.t.* RecA2 (ScA1-CtA2)	31-250 (*S.c.*) fused to 251- 421 (*C.t.*)	N-His_6-TEV	#1911
C.t. RecA1 – *S.c.* RecA2 (CtA1-ScA2)	17-250 (*C.t.*) fused to 251-425 (*S.c.*)	N-His_6-TEV	#1803

Chemical competent *E. coli* BL21 (DE3) Codon Plus RIL (Stratagene) cells were transformed with the appropriate plasmid and grown in LB (lysogeny broth) medium at 37 °C. At an OD_{600} of 0.8, protein overexpression was induced with 0.5 mM IPTG (Isopropyl β-D-1-thiogalactopyranoside) at 20 °C. Following overnight expression, cells were pelleted by centrifugation and lysed by ultrasound in buffer A (25 mM sodium phosphate, pH 7.4, 500 mM NaCl, 1 mM DTT) complemented with 10 mM imidazole, lysozyme, 0.1 % Triton X-100, 5 mM $MgCl_2$ and 0.2 U/mL DNase. Insoluble debris were removed from the cell lysate by centrifugation and the supernatant was applied to Ni^{2+}-nitrilotriacetic acid (NTA) resin equilibrated in buffer A. The resin was washed with 15 column volumes (CV) buffer A that was complemented with 10 mM imidazole. The protein bound to the resin was eluted with buffer A complemented with 300 mM imidazole. 1 mg TEV protease was added to the eluted proteins to cleave the purification tag from the target protein.

Subsequently, proteins were dialyzed overnight at 20 °C into buffer B (20 mM MES, pH 6.0, 125 mM NaCl, 1 mM DTT) supplemented with 0.5 mM EDTA. Solutions of dialyzed Dhh1 constructs that contain the helicase core (RecA1-RecA2) were adjusted to 250 mM NaCl and supplemented with 4 % (v/v) 3 M sodium acetate, pH 5.3. Proteins were concentrated using centrifugal filters. Purification to homogeneity was achieved by size exclusion chromatography (SEC) on a HiLoad 16/600 Superdex 75 column (GE Healthcare) in buffer C (20 mM MES, pH 6.0, 100 mM NaCl, 1 mM DTT).

After SEC, pooled fractions of Pat_N (Pat1, residues 5-79) or Edc3 IDR (residues 86-125) still contained trace amounts of the His_6-GST expression tag as assessed by SDS-PAGE. The impurities were removed by reverse affinity chromatography. In brief, the solution was applied to Ni^{2+}-NTA resin equilibrated with buffer C that was supplemented with 10 mM imidazole. The resin was washed with two CV buffer C complemented with 10 mM imidazole. The flow-through and wash fractions, that contained only the target protein, were pooled and concentrated for a second size exclusion chromatography step, after which pure and homogenous target protein was obtained.

4.2.2 *Complex reconstitution*

Dhh1:Pat_N and Dhh1:Edc3_idr complexes were reconstituted following published protocols[105]. In brief, purified components were mixed in a 1:1.5 molar ratio, whereby the peptide was in excess over the Dhh1 protein. After incubation for 1 hour at 30 °C, unbound protein was separated from the complex on a HiLoad 26/600 Superdex 75 column (GE Healthcare) in buffer C. Fractions containing the complex were pooled and concentrated using centrifugal filters.

4.2.3 *Protein labeling for NMR spectroscopy*

Labeling of proteins with NMR-active nuclei was achieved by over-expression of the gene in M9 minimal medium. The growth medium was supplemented with 1 g/L $^{14}NH_4Cl$ as the sole nitrogen source and 2 g/L $^{2}H^{12}C$-glucose, if the medium was based on 100 % D_2O, or 4g/L $^{1}H^{12}C$-glucose, if the deuteration level was less. Labeling of the Ala-β, Ile-δ_1, Met-ε, Leu-δ_1/δ_2 and Val-γ_1/γ_2 methyl groups (ILVMA-labeling) within a *U*-[^{2}H,^{12}C]-background was achieved by supplementing the medium with α-ketobutyrate (3-$^{2}H_2$-4-^{13}C; 60 mg/L), L-methionine (*methyl*-$^{13}CH_3$; 100 mg/L) and α-ketoisovalerate (3-^{2}H-3-(*methyl*-$^{13}CH_3$)-4-^{13}C; 100 mg/L) one hour prior to induction. L-alanine (2-^{2}H-3-^{13}C, 100 mg/L) was added 20 minutes prior to induction. For samples to be labeled in Val but not Leu methyl groups, 100 mg/L unlabeled α-ketoisocaproate is added in combination with methyl-labeled α-ketoisovalerate. For labeling of Ala-β methyl groups without simultaneous labeling of valine and leucine, the medium is supplemented with perdeuterated α-ketoisovalerate instead of its methyl-labeled form (**Figure 1.4**).

4.2.4 *NMR spectroscopy*

All NMR samples were prepared in buffer C and contained 5 % D_2O in the case of IM-, IA-, IV-, IMA- and IMV-labeled samples, or 100 % D_2O for ILVMA-labeled samples. The latter were used to record 3D HMQC-NOESY-HMQC spectra with short and long mixing times (40-60 ms and 300 ms, respectively) for automated methyl group assignments using MAGIC[240]. NMR spectra were recorded at 298 K on a Bruker AVIII-800 spectrometer equipped with a cryogenic probe-head. NMR titration experiments were carried out with 40 µM IM-labeled Dhh1 RecA2 and a 4-fold excess of unlabeled Dhh1 RecA1. NMR spectra were processed using the NMRPipe/NMRDraw software suite[267]. Figures displaying NMR spectra were prepared using NMRview (onemoonscientific.com) or Sparky[338].

4.2.5 *RNA in vitro transcription and purification*

In vitro transcription and purification of a 30mer RNA (**Table 4.2**) was carried out as described before[252]. A DNA primer served as the template and transcription was accomplished using in-house purified T7 RNA polymerase[263]. In brief, transcribed RNA was purified under denaturing conditions using anion exchange chromatography, followed by isopropanol precipitation and desalting. Solvent evaporation resulted in the pure dry RNA product, which was resuspended at a concentration required for phase separation experiments. The quality and integrity of the RNA was controlled using urea-polyacrylamide gel electrophoresis in 1x TBE (89 mM Tris, pH 8.0, 89 mM boric acid, 2 mM EDTA). RNA bands were visualized by methylene blue staining.

Table 4.2: 30mer RNA used in this study

RNA	Properties	Sequence	Internal reference
30U15mer	GA-only, but single U at position 15	GGAGGAGAGGAAGGUAAGGGAAGAAAGAAG	#10 primer

4.2.6 *Liquid-liquid phase separation experiments*

For LLPS experiments, proteins were prepared as 3x concentrated stock solutions in buffer C that contains 125 mM NaCl. 3.3 µL of protein were diluted into prepared mixtures (6.6 µL) of 20 mM MES/HEPES and NaCl that resemble a grid of conditions with the following final parameters: pH 6.0, 6.5, 7.0, 7.5 and 35 mM, 50 mM, 75 mM, 100 mM NaCl. For each condition, three technical replicates were prepared. Immediately after dilution, the absorption at 600 nm (OD_{600}) as a quantitative estimate of phase separation was measured on a Nanodrop spectrophotometer. Mean values for each technical replicate are from three measurements. Phase separation experiments were carried out at 20 °C unless stated otherwise.

To investigate the temperature-dependence of the Dhh1 LLPS process, three 50 µL samples at 100 µM protein concentration were prepared from a 3-fold concentrated stock solution at a final pH of 6.65 and a final salt concentration of 75 mM. The samples were kept for 2 minutes at a defined temperature (4, 20 or 37 °C) before the OD_{600} was measured on a Nanodrop spectrophotometer. Afterwards, the samples were transferred to another temperature. The steps were as follows: room temperature, 4 °C, 20 °C, 4 °C, 20 °C, 37 °C, 20 °C, 4 °C, 37 °C, 4 °C, 20 °C.

4.3 *RESULTS*

4.3.1 *Dhh1 undergoes phase separation in the absence of RNA*

In previous work, we established that a synergistic network of protein-protein and protein-RNA interactions promotes the formation of *in vitro* processing bodies[252]. We found that weak but multivalent binding events between decapping factors form the basis of theoretically infinite interaction networks which manifest in a liquid-liquid phase separation process *in vitro* and eventually processing body formation *in vivo*[103,252]. Considering the finding that Edc3 can undergo LLPS independent of any other binding partner, we hypothesized that *S. cerevisiae* Dhh1 may also possess a potential intrinsic phase separation capability.

To test our hypothesis, we first focused on the helicase core of Dhh1, which consists of the two RecA-like domains RecA1 and RecA2 but lacks the unstructured N- and C-terminal extensions (**Figure 4.3 A**). We assessed LLPS of the Dhh1 helicase core at different conditions (pH 6.0-7.4 and 35-100 mM NaCl) with turbidity measurements that accurately report on the degree of LLPS[134,252]. We found that the Dhh1 core phase separated preferably at high pH and low salt concentrations and that this LLPS behavior is most pronounced at high protein concentrations (**Figure 4.3 B**). The observed LLPS process is dependent on the intact helicase core as a mixture of the two separately purified RecA-like domains (100 µM each) is not able to undergo phase separation (**Figure 4.3 C**).

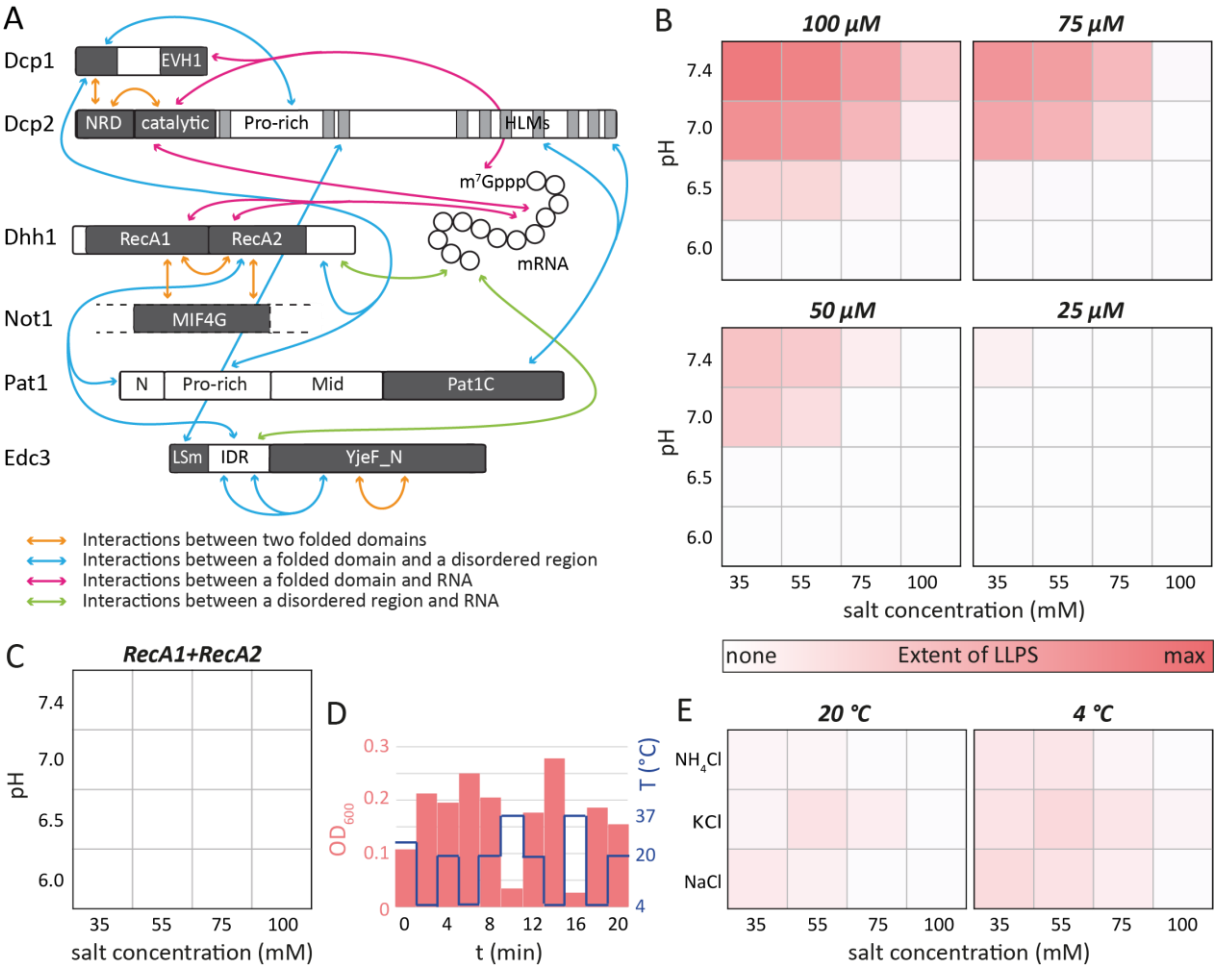

Figure 4.3: Phase separation of the Dhh1 helicase core. (A) Schematic representation of *S. cerevisiae* proteins, RNA and interactions important for LLPS *in vivo*. The DEAD-box RNA helicase Dhh1 is tightly embedded in an interaction network formed by mRNA degradation factors and RNA. Different interaction modes (colored arrows) complement each other in the formation of *in vitro* liquid droplets and *in vivo* processing bodies. **(B)** Phase diagrams of the *S. cerevisiae* Dhh1 helicase core protein at different concentrations. The protein undergoes LLPS in a concentration-dependent manner at high pH and low salt concentrations. White color indicates no phase separation, while red indicates a high degree of LLPS as assessed by turbidity measurements at 600 nm. The color scheme is the same for all grids shown in this chapter. **(C)** LLPS of Dhh1 is dependent on the intact helicase core. A mixture of individually expressed and purified RecA1 and RecA2 domains does not phase separate. An interaction between the two domains in *trans* is not sufficient for LLPS. **(D)** Phase separation of Dhh1 is temperature-dependent. Low temperatures promote phase separation, while at elevated temperatures liquid droplets dissolve. **(E)** Dhh1 LLPS does not depend on the type of salt present in the buffer. Experiments were carried out with 75 µM protein at pH 6.5.

Depending on the amino acid composition, the type of interactions and thermodynamic processes that govern phase separation, proteins can undergo liquid-liquid phase separation at low or high temperatures[142,339]. We found that Dhh1 undergoes phase separation at low temperatures (**Figure 4.3 D**). Enhanced phase separation at low temperatures is associated with polar residues that govern the LLPS process[142,339]. Moreover, the temperature-dependent phase separation of Dhh1 is reversible (**Figure 4.3 D**): at high temperatures, preformed liquid droplets of Dhh1 rapidly dissolve, while a subsequent decrease in temperature results in droplet re-formation. The observed reversibility is a hallmark of LLPS processes[103,120,137,142,154,273].

To test the influence of different salts on LLPS of the Dhh1 core, we performed phase separation experiments in the presence of NaCl, KCl and NH_4Cl (**Figure 4.3 E**). We did not observe any strong salt dependency and thus used NaCl in all further experiments. Independent of the salt type, phase separation of the Dhh1 core is enhanced at lower temperatures (**Figure 4.3 E** right panel).

4.3.2 *Residues at the C-terminus of the RecA2 domain are crucial for LLPS*

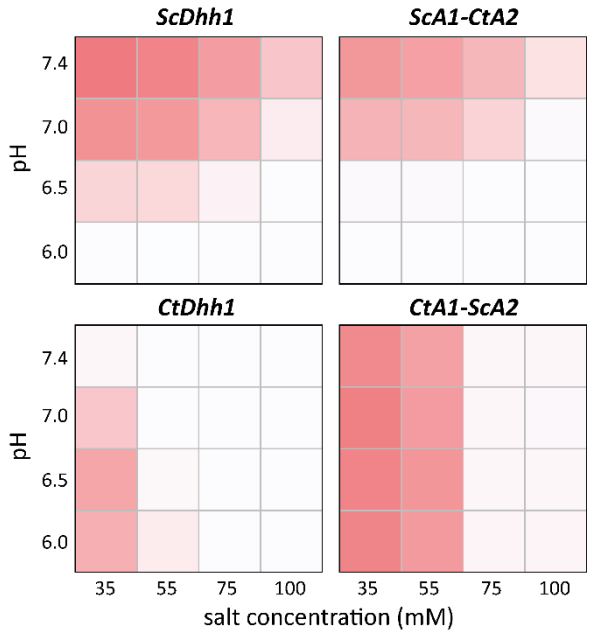

Figure 4.4: Dhh1 phase separation patterns are species-dependent. Phase diagrams for Dhh1 helicase core proteins from *S. cerevisiae* (top left), *C. thermophilum* (bottom left) and chimeric constructs (right) at 100 µM. *S.c.* and *C.t.* Dhh1 proteins exhibit different LLPS patterns. The *S.c.* RecA2 appears to promote phase separation as its loss (from wild-type ScDhh1 to chimeric ScA1-CtA2) results in reduced phase separation. Conversely, gain of the *S.c.* RecA2 domain (from CtDhh1 to chimeric CtA1-ScA2) enhances phase separation.

We next asked which parts of the Dhh1 helicase core mediate phase separation. To our surprise, we found that the *C. thermophilum* and *S. cerevisiae* Dhh1 helicase core proteins (CtDhh1 and ScDhh1, respectively) show different phase separation patterns (**Figure 4.4** left), although they share over 75 % sequence identity (**Figure 4.5 A**). As for ScDhh1, CtDhh1 shows a high degree of phase separation at low salt concentrations. However, LLPS of CtDhh1 is strongest at low pH, while for ScDhh1 LLPS is strongest at high pH (**Figure 4.4**, compare top left and bottom left).

We used the *C. thermophilum* homolog of the *S. cerevisiae* Dhh1 protein as a toolbox to determine crucial residues for phase separation of the *S.c.* Dhh1 helicase core.

First, we designed two chimeric proteins with swapped domains. The first chimera consists of the RecA1 domain of *S.c.* Dhh1 fused to the RecA2 domain of *C.t.* Dhh1 (ScA1-CtA2) and the second chimera consists of the RecA1 domain of *C.t.* Dhh1 fused to the RecA2 domain of *S.c.* Dhh1 (CtA1-ScA2). We then assessed the phase separation behavior of the two chimeric proteins to determine if they exhibit intermediate phase separation patterns or if the LLPS patterns are dominated by only one of the two RecA-like domains.

The phase separation pattern of the ScA1-CtA2 chimera is reminiscent of the pattern observed for the *S.c.* Dhh1 helicase core (ScDhh1) although the overall degree of phase separation is reduced relative to the wild-type protein (**Figure 4.4**, compare top left and right). In contrast, for the CtA1-ScA2 chimera we observe a LLPS pattern which resembles that of CtDhh1 but shows enhanced phase separation compared to the latter (**Figure 4.4**, compare bottom left and right). These results indicate that the *S.c.* RecA2 domain is a driving force of Dhh1 phase separation: gain of the *S.c.* RecA2 domain (from CtDhh1 to CtA1-ScA2; **Figure 4.4** bottom) enhances LLPS, while its loss (from ScDhh1 to ScA1-CtA2; **Figure 4.4** top) reduces phase separation.

We hypothesized that stepwise mutations at positions in the RecA2 domain, which are different in *S. cerevisiae* and *C. thermophilum*, would result in a transition of the LLPS pattern from ScDhh1 towards ScA1-CtA2 (**Figure 4.4** top left and right). To that end, we identified three sequence stretches in the RecA2 domain which are least conserved between *S. cerevisiae* and *C. thermophilum* (**Figure 4.5 A**). These stretches (1-3) are solvent-exposed and, for stretch 1 and 2, alter the electrostatic surface potential of Dhh1 (**Figure 4.5 B**, top and bottom left). Changes in the presence and distribution of surface-exposed charged residues might influence the LLPS behavior of the Dhh1 protein, as electrostatic interactions are one possibility of proteins to undergo intracellular phase transitions[119].

Consequently, we expressed and purified Dhh1 helicase core proteins which harbor mutations in the three stretches. In that process, the wild-type *S. cerevisiae* Dhh1 sequence was changed towards the *C. thermophilum* protein, and LLPS behavior was tested based on pH/salt-grids (**Figure 4.6**). Changing the sequence in stretch 1 has no effect on Dhh1 LLPS, as this mutant (**Figure 4.6**, top right) shows phase separation comparable to wild-type ScDhh1 (**Figure 4.6**, compare top left and right), likewise, changing both stretch 1 and 2 has no strong effect on the Dhh1 LLPS behavior (compare left top and bottom). Contrarily, in the stretch 1+3 and stretch 3 mutants phase separation is completely abolished (**Figure 4.6**, bottom middle and right). Hence, we conclude that residues in stretch3 are important for Dhh1 LLPS.

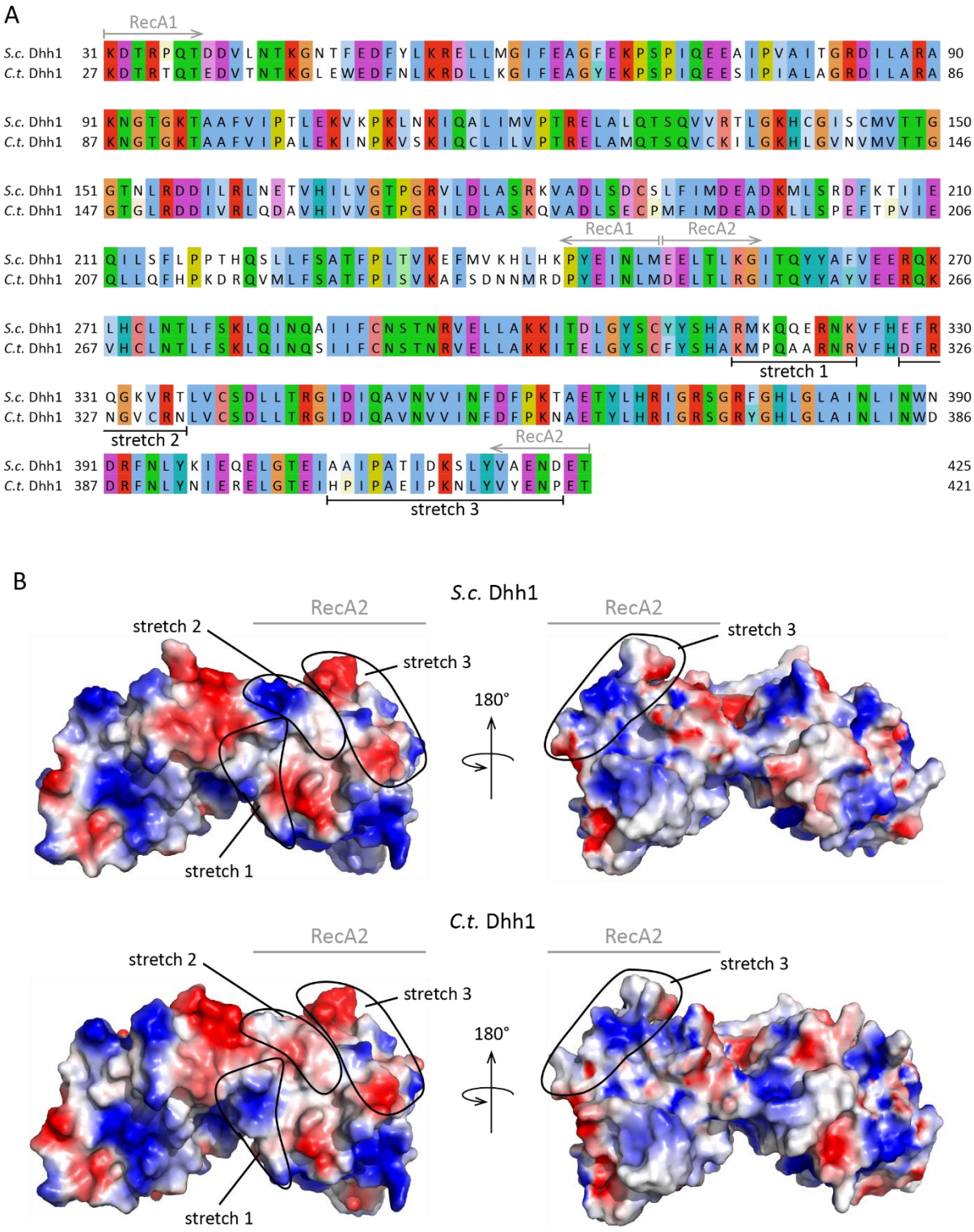

Figure 4.5: Dhh1 proteins from *S. cerevisiae* and *C. thermophilum* are highly similar. (A) Sequence alignment of the Dhh1 helicase core regions of *S. cerevisiae* and *C. thermophilum*. Domain boundaries are indicated in gray. Three sequence stretches in the RecA2 domain (stretch 1-3, black) exhibit low conservation between the two yeast species. **(B)** Electrostatic surface potentials of *S.c.* Dhh1 (top) and *C.t.* Dhh1 (bottom). The three stretches are solvent-exposed, and their location is indicated. The sequence variations in these patches between *S. cerevisiae* and *C. thermophilum* results in an altered surface potential (compare especially patch 1 and 2 of *S.c.* and *C.t.* Dhh1). The orientation is the same for *S.c.* and *C.t.* Dhh1 and the location of the RecA2 domain is indicated. Surface potentials were calculated in pymol 2.3 (Schrödinger, LLC; https://pymol.org/2/) and color-coded from red (negatively charged) to blue (positively charged). The structure of *C.t.* Dhh1 is a homology model obtained with SWISS-MODEL (https://swissmodel.expasy.org/) and based on the crystal structure of *S.c.* Dhh1 (pdb: 1s2m).

Two sets of sequence variations in stretch 3 (residues 407-423) between *S. cerevisiae* and the *C. thermophilum* caught our attention. First, in the crystal structure of *S.c.* Dhh1 A407 and A408 form a small β-strand which augments the central β-sheet of the RecA2 domain[318]. In *C. thermophilum*, these two alanine residues are substituted by a His-Pro dipeptide (**Figure 4.5 A**). As proline is generally considered to break secondary structures[340,341], the A408P substitution might thus prohibit β-sheet augmentation, which could be the reason for the different LLPS behavior of *S. cerevisiae* and *C. thermophilum* Dhh1. And second, the aspartate residues D414 and D423, which are in a loop in *S.c.* Dhh1, are substituted by proline residues in *C.t.* Dhh1 (**Figure 4.5 A**). Since LLPS can be mediated by charge-charge interactions[119], net loss of two negative charges due to the double Asp-to-Pro substitution might result in reduced phase separation.

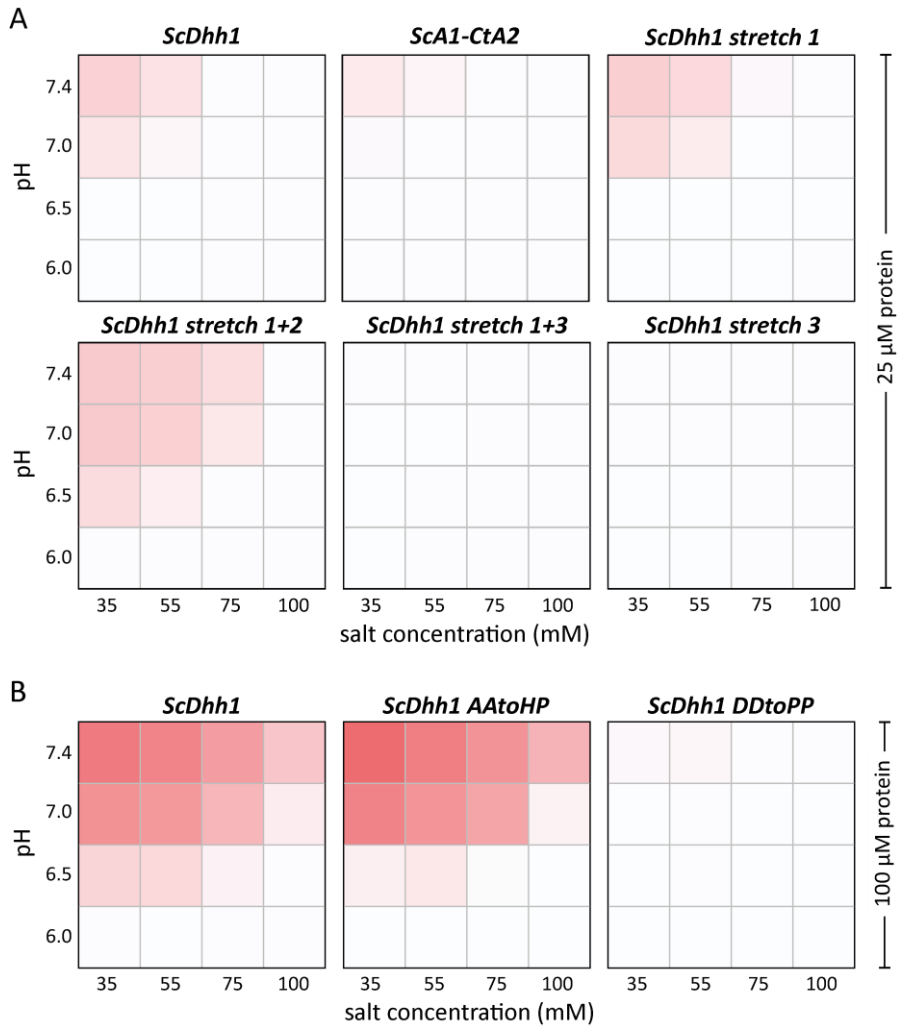

Figure 4.6: Mutations in the C-terminus of the RecA2 domain abolish LLPS of Dhh1. (A) Phase diagrams of Dhh1 helicase core proteins at 25 µM. Mutations in two regions (stretch 1 and stretch 2; top right and bottom left, respectively), which exhibit sequence variations between *S.c.* and *C.t.* (see **Figure 4.5**), do not reduce phase separation compared to wild-type ScDhh1 (top left). Mutations in the very C-terminus of the RecA2 domain (stretch 3) from *S.c.* to *C.t.* sequence result in a complete loss of LLPS (bottom middle and right). The chimeric ScA1-CtA2 protein still exhibits a small degree of phase separation under these conditions (top middle). **(B)** Phase diagrams of mutant Dhh1 helicase core proteins at 100 µM. The phase diagram of wild-type Dhh1 is shown as a reference (left). Within stretch 3, double mutation of an Ala-Ala dipeptide, which is engaged in augmentation of a β-sheet in the RecA2 domain, has no influence on LLPS of ScDhh1 (middle). Stretch 3 also contains two Asp residues, whose double mutation to Pro is sufficient to abolish phase separation completely (right).

To test which of these two features are important for the LLPS of Dhh1, we created the double mutants ScDhh1 A407H, A408P ("AAtoHP") and ScDhh1 D414P, D423P ("DDtoPP") and tested their phase separation capability (**Figure 4.6 B**). The AAtoHP double mutant (bottom left) shows phase separation to the same extent as the wild-type ScDhh1 protein (top left), indicating that the differences in the β-sheet architecture are not important for the enhancement of LLPS. On the other hand, the DDtoPP double mutant is sufficient to abolish phase separation (bottom middle). This indicates that charge-charge interactions involving in the RecA2 domain are responsible for the phase separation of Dhh1. Importantly, these results demonstrate that few (in this case two) amino acid substitutions can have a drastic impact on the phase separation behavior of a protein.

4.3.3 *The unstructured extensions enhance LLPS of the Dhh1 helicase core*

As we have previously established that the IDR in *S. pombe* Edc3 is important for phase separation of this protein, we now speculate that the disordered N- and C-terminal extensions (NTE and CTE, respectively) of *S. cerevisiae* Dhh1 might enhance phase separation of the helicase core. To test that, we prepared Dhh1 constructs that either include or lack the NTE, CTE or both extensions and evaluated their LLPS behavior (**Figure 4.7 A**). The Dhh1 helicase core (ΔNC) shows only limited LLPS at 25 µM protein concentration (bottom right). For a Dhh1 construct comprising the NTE and the helicase core but lacking the longer CTE (Dhh1 ΔC), phase separation is slightly enhanced relative to the Dhh1 core (**Figure 4.7 A,** compare bottom left and right). A stronger enhancement of LLPS relative to the Dhh1 core was observed for a construct comprising the helicase core and the CTE (Dhh1 ΔN; **Figure 4.7 A,** compare top and bottom right). The highest degree of phase separation was observed for full-length Dhh1 (**Figure 4.7 A,** top left), where both extensions are present. Our results indicate that both the N- and the C-terminal extensions provide additional interactions between the Dhh1 proteins, which results in a shift of the phase separation boundary towards higher salt concentrations and lower pH. The contribution of the CTE to LLPS of full-length Dhh1 is higher than that of the N-terminal extension, probably due to the pronounced prion-like nature of the CTE. These data are consistent with previous studies that also showed *in vitro* phase separation for full-length Dhh1[164].

4.3.1 *ATP and RNA strongly enhance LLPS of full-length Dhh1*

The DEAD-box helicase Dhh1 is able to interact with nucleotides, including ATP[323]. Interestingly, in the presence of a 100-fold excess of ATP, phase separation of the full length Dhh1 protein is strongly enhanced (**Figure 4.7 B**, compare top left and right). It should be noted that in our assays the turbidity is measured immediately after induction of LLPS.

Given the low intrinsic ATPase activity of Dhh1 in the absence of RNA[323], the observed increase of LLPS in the presence of ATP is likely not due to ATP hydrolysis or consequential effects such as ADP binding.

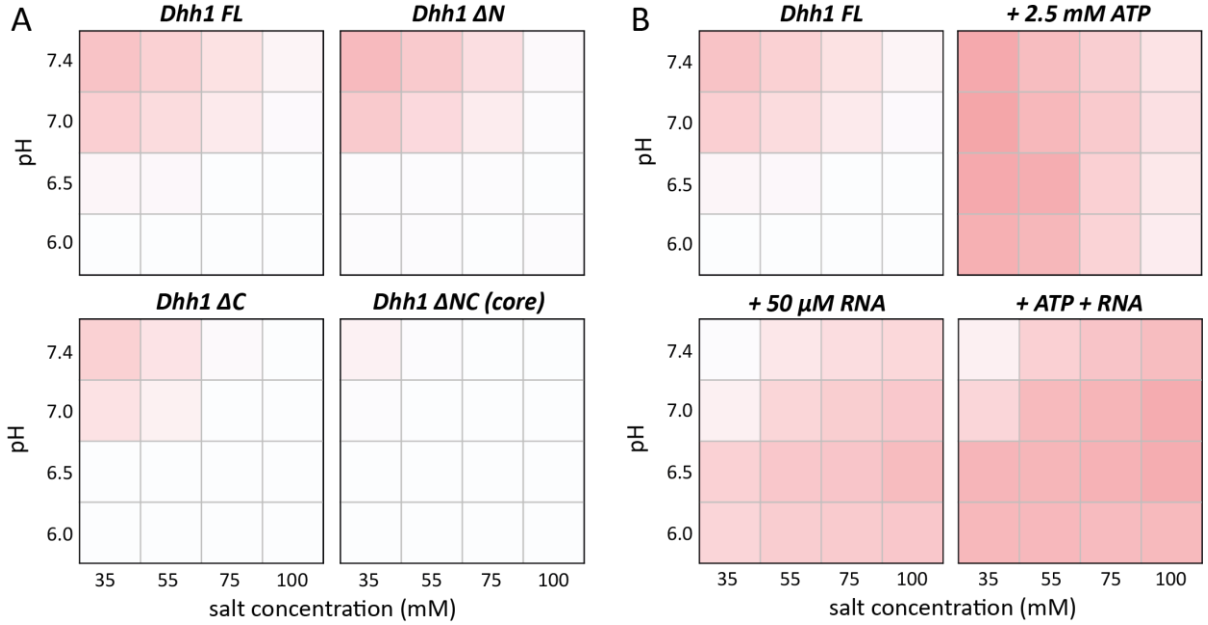

Figure 4.7: LLPS of the Dhh1 core is enhanced by the unstructured tails, ATP and RNA. (A) Phase separation diagrams for full-length (FL) Dhh1 as well as truncation constructs lacking either the N- (ΔN) or the C-terminus (ΔC) or both (ΔNC) at 25 µM protein concentration. The disordered tails enhance phase separation of the Dhh1 helicase core (ΔNC) indicated by the shift of the phase separation boundary towards lower pH and higher salt concentrations. **(B)** Phase diagram of 25 µM FL Dhh1 (top left) and in presence of a 100-fold excess of ATP (top right), a 2-fold excess of 30mer RNA (bottom left) or both ATP and RNA (bottom right). ATP enhances phase separation of Dhh1, while presence of RNA inverts the phase separation pattern.

RNA is a key factor for the formation of cytoplasmic granules, such as processing bodies, and as such RNA has been shown to enhance *in vitro* LLPS[150,151,252,342]. In previous reports, poly(U) RNA has been shown to enhance the LLPS tendency of Dhh1[152,164]. Poly(U) RNA is a mixture of RNAs with undefined lengths up to a few thousand nucleotides. This inhomogeneity poses a difficulty when investigating the contributions from different interactions on the phase separation process of Dhh1. Therefore, we here make use of a well-defined *in vitro* transcribed RNA: an unstructured 30mer with a sequence composed of adenosine and guanosine nucleotides and a single uridine at position 15 (30mer GA-RNA, see **Table 4.2**). Based on crystal structures of other DEAD-box RNA helicases in complex with RNA[319–322,343–346], Dhh1 likely binds 6-10 nucleotides in a cleft that is formed in the closed conformation of the helicase core. The length of the RNA we used here was chosen such that at least two Dhh1 molecules can bind simultaneously to one RNA molecule, and thereby bridging of two Dhh1 proteins via one RNA molecules is possible.

In our LLPS assays, we added a two-fold molar excess of the 30mer GA-RNA to full-length Dhh1. Unexpectedly, we observed an “inversion” of the phase separation pattern relative to the full-length Dhh1 protein (**Figure 4.7 B**, compare left top and bottom).

For full-length Dhh1, we observe LLPS at high pH and low salt concentrations. In the presence of RNA, we observe strong phase separation at lower pH (6.5) and high salt concentrations, where the full-length protein displays no phase separation. The addition of ATP to the Dhh1:RNA complex lowers the LLPS boundary, without further changing the salt and pH preferences (**Figure 4.7 B**, compare bottom left and right). These results indicate that the Dhh1 protein exploits different phase separation mechanisms in the absence and presence of RNA.

4.3.2 *Binding of Edc3 or Pat1 disrupts phase separation of the Dhh1 core*

Dhh1 is tightly embedded in the mRNA degradation network (**Figure 4.3 A**) and we aimed to assess the effect of decapping factors on Dhh1 phase separation. Here, we focused on the Dhh1 helicase core as the known interaction partners Pat1 and Edc3 bind in a mutually exclusive manner to the Dhh1 RecA2 domain[105]. The two conserved Dhh1-interaction motifs FG/WQ and FDF reside in the disordered Pat1 N-terminus and in the IDR of Edc3 (**Figure 4.3 A**). The Edc3 IDR additionally harbors a Dhh1-interacting FNK motif, which is not present in Pat1.

To test the effect of Pat1 and Edc3 on Dhh1 LLPS, we reconstituted complexes of the Dhh1 helicase core with peptides of either Edc3 (residues 86-125; Edc3_idr) or Pat1 (residues 5-79; Pat_N) (**Figure 4.8**). Remarkably, the presence of Pat1 or Edc3 almost completely abolishes the phase separation propensity of Dhh1 (**Figure 4.9 A**, compare middle and right panels with left panel).

Interestingly, in the presence of Pat1, a strong phase separation at low pH and low salt is reproducibly observed. It remains unclear if this is due to an altered phase separation behavior of the Dhh1:Pat_N complex or if Pat_N dissociates from the complex with Dhh1 under the aforementioned conditions and starts to undergo homotypic phase separation.

In summary, we have shown that the Dhh1 helicase core is able to undergo LLPS. We have identified two aspartate residues on the RecA2 domain that are important for LLPS. The phase separation behavior of Dhh1 is enhanced by the N- and C-terminal extensions, as well as by ATP and RNA. The interaction of Pat1 and Edc3 on the other hand diminish *in vitro* LLPS of Dhh1 in the absence of ATP and RNA. Structurally, the mechanism of this behavior is not clear, as Pat1 and Edc3 bind in a region that is remote from the important aspartate residues.

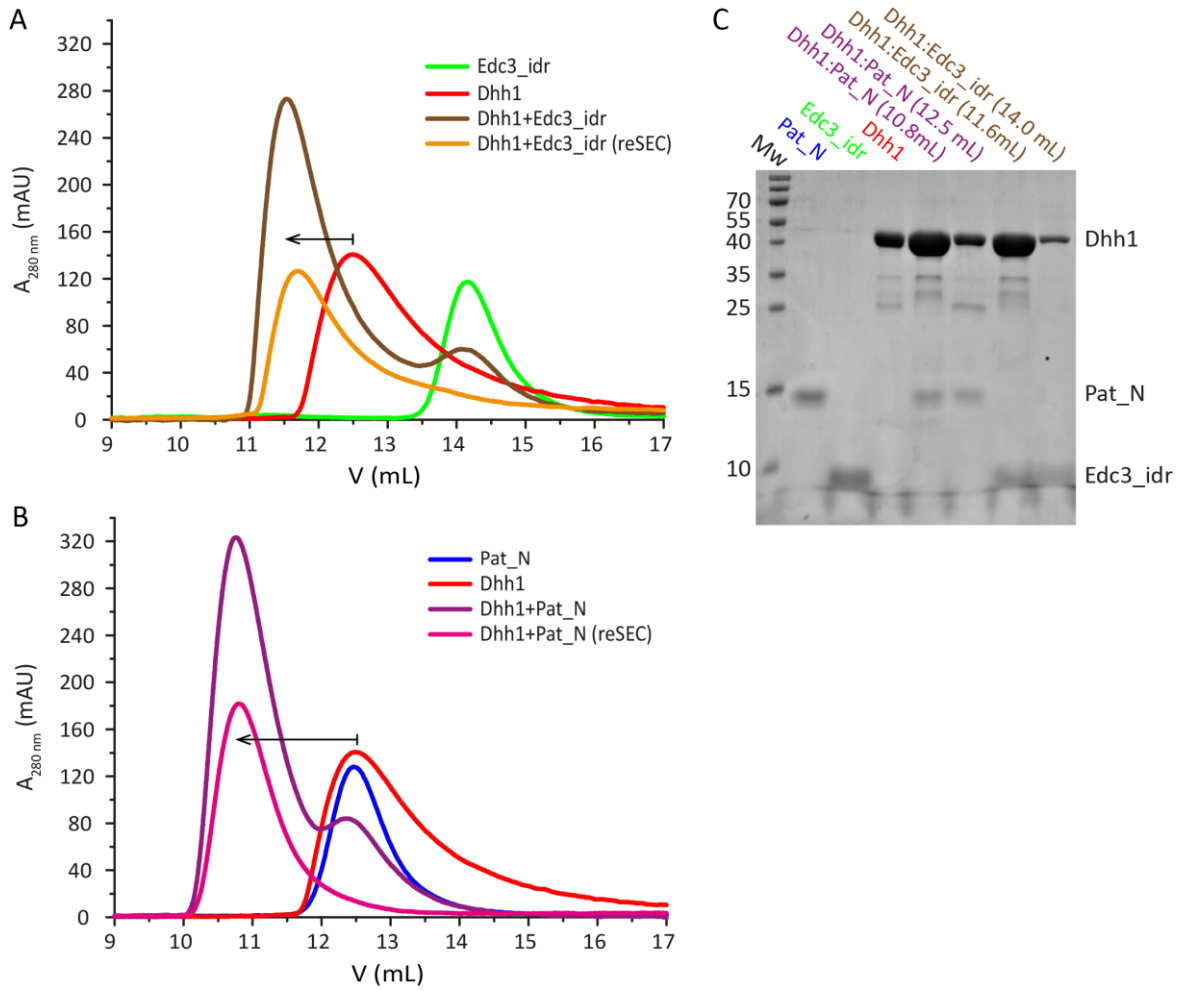

Figure 4.8: The Dhh1 helicase core forms stable complexes with Pat1 and Edc3 peptides. (A, B) Analytical gel filtration chromatograms of Edc3_idr (green), Pat_N (blue), Dhh1 helicase core (red) and complexes thereof, Dhh1:Edc3_idr (brown) and Dhh1:Pat_N (violet). Peak fractions of the complexes (V_E of 11.6 and 10.8 mL for Dhh1:Edc3_idr and Dhh1:Pat_N, respectively) were concentrated and subjected again to analytical gel filtration (orange and magenta chromatograms). The Dhh1:Edc3_idr and Dhh1:Pat_N complexes are stable as the re-run complexes do not show significant amounts of the free peptides, which were used in excess for complex reconstitution. Black arrows indicate the shift in elution volume from free Dhh1 to Dhh1:Edc3_idr or Dhh1:Pat_N, respectively, indicative for the formation of stable complexes. **(C)** SDS-PAGE analysis of peak fractions from analytical gel filtration. The reconstituted complexes contain the Dhh1 helicase core protein and the Pat_N or Edc3_idr peptides. Note that the peptides do not run true-to-size and stain poorly.

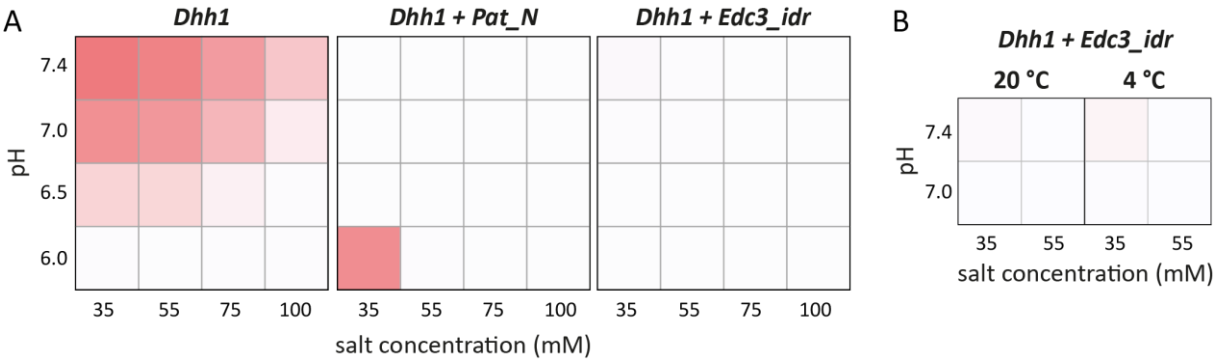

Figure 4.9: Binding of Pat1 and Edc3 peptides affects LLPS of the Dhh1 helicase core. (A) Phase diagrams of the Dhh1 helicase core (left), the Dhh1:Pat_N (middle) and the Dhh1:Edc3_idr (right) complex at 100 µM. Binding of Pat_N changes the phase separation pattern of the Dhh1 helicase core protein towards low pH and low salt, while binding of the Edc3 peptide almost completely abolishes phase separation. **(B)** LLPS experiments for the Dhh1:Edc3_idr complex at 4 °C reveal that there is residual phase separation at high pH and low salt. This indicates that binding of the Edc3_idr peptide significantly reduces phase separation but does not change the LLPS pattern observed for free Dhh1 protein.

4.3.3 *In solution, the two RecA domains of Dhh1 tumble independently*

To gain structural insights into the molecular interactions within the phase-separated Dhh1 protein, we turned to NMR. Recently, we demonstrated that solid-state NMR is suitable to study a matured, gel-like phase of the Edc3 protein from *S. pombe* (Damman R, Schütz S, *et al.*, under revision; see CHAPTER 3). Here, we aim at applying solution-state NMR to study liquid droplets of the folded Dhh1 helicase core. Solution NMR studies on phase-separated proteins are still rare and mostly involve intrinsically disordered proteins only[283,347,348]. We applied methyl TROSY techniques, as these are well suited to study high-molecular weight assemblies in solution[184].

^{1}H,^{13}C-HMQC spectra of the ILVMA-labeled 45 kDa Dhh1 helicase core protein under non-phase separating conditions (**Figure 4.10**) are of very high quality. This spectrum (black) overlays very well with the spectra of the isolated RecA1 and RecA2 domains (red and green, respectively). This indicates that both RecA domains tumble independently in the full-length protein. The previous finding that both domains do not interact in the absence of RNA and ATP is confirmed by the lack of chemical shift perturbations when the NMR-inactive RecA1 domain is added to the NMR-active RecA2 domain (**Figure 4.11**). Remarkably, in the crystal structure of apo Dhh1 extensive interdomain contacts were associated with a compact open conformation. Our results indicate that these contacts are released or absent in solution.

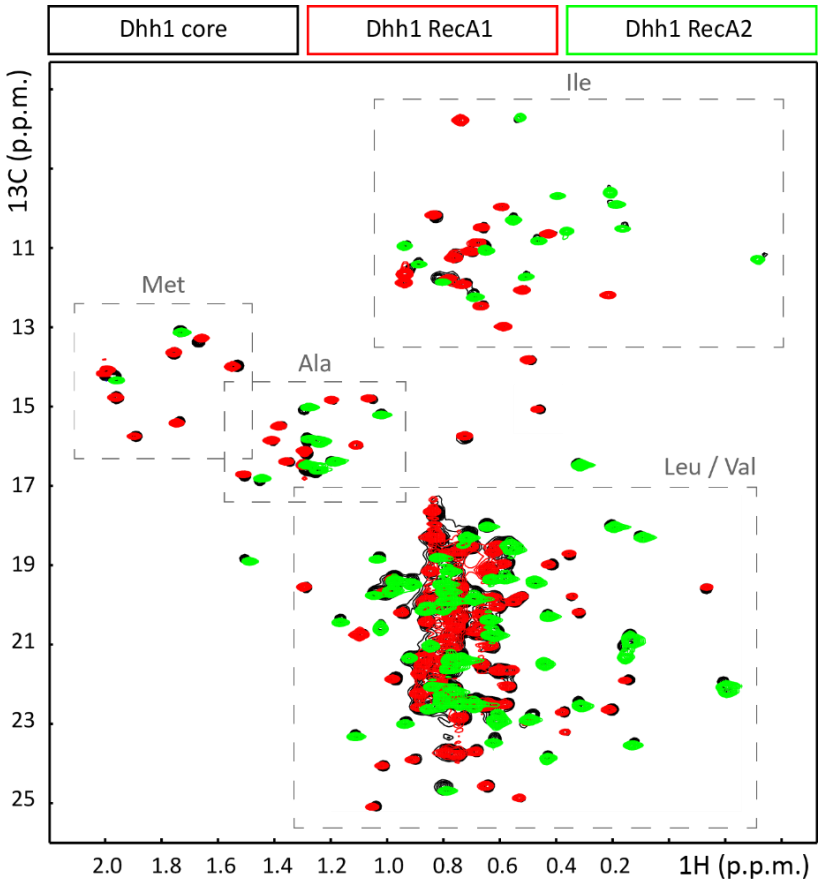

Figure 4.10: The two RecA-like domains of Dhh1 tumble independently in solution. Methyl TROSY spectra of ILVMA-labeled Dhh1 (black), Dhh1 RecA1 (red) and Dhh1 RecA2 (green). The black spectrum is, in good approximation, the sum of the red and the green spectrum, indicating that both RecA-like domains in the Dhh1 protein have limited contacts in solution. Resonance assignments by residue type can be approximated from the peak positions in the spectrum[187] as indicated by the dashed rectangles. These boundaries are not exclusive, overlap of the respective regions and outliers can occur. However, absolute residue-type specific assignments are required to facilitate automated assignment procedures[240].

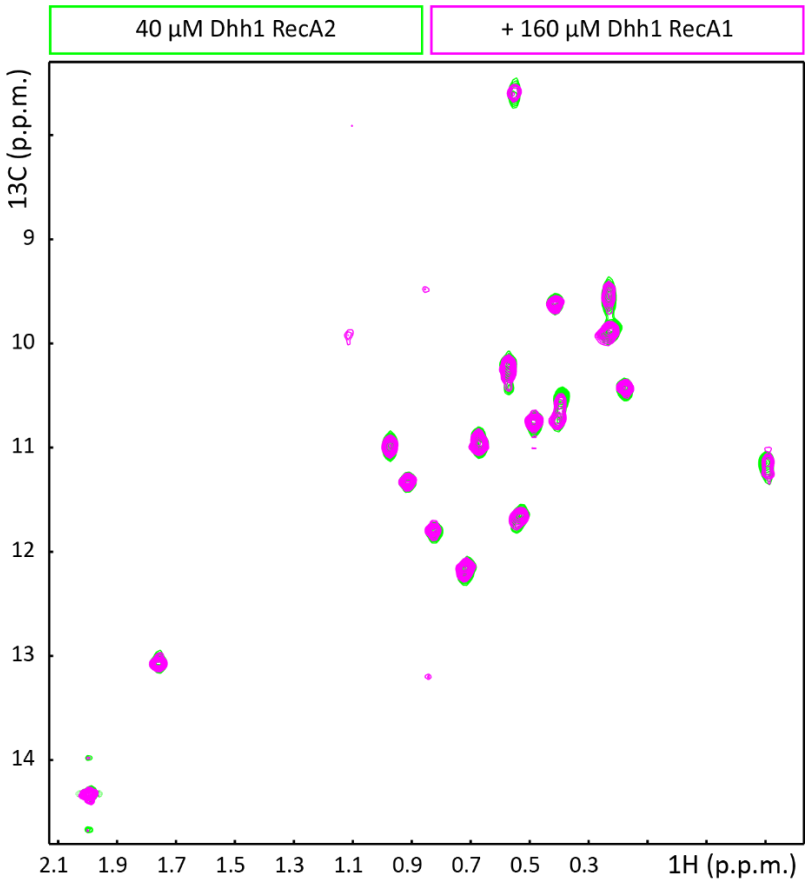

Figure 4.11: The RecA-like domains of Dhh1 do not interact in solution. Methyl TROSY spectra of the IM-labeled Dhh1 RecA2 domain in the absence (green) and presence (magenta) of a four-fold excess of unlabeled RecA1. The absence of clear CSPs indicates that both domains do not interact in *trans* under these conditions.

4.3.4 *Assignment of the Dhh1 methyl groups*

To assign the methyl resonances in the Dhh1 protein we took advantage of the divide-and-conquer approach (see 1.4.4). Traditional assignment strategies based on backbone experiments turned out to be insensitive and the individual domains were thus assigned based on the MAGIC approach[240]. In this approach, methyl resonances are assigned based on an existing structural model and experimental NOEs[240]. To facilitate the assignment, we assigned the residue type of each methyl resonance in the Dhh1 RecA domains by preparing a set of protein samples that are methyl-labeled only in IMA or IMV residues (for RecA1) and IA or IV (for RecA2) (**Figure 4.12**). From these spectra, we conclude that all the expected Ile (16), Met (8) and Ala (14) as well as 26 of 28 Val resonances and most of the Leu methyl groups (47 of 54) can be detected in the RecA1 domain (**Figure 4.12 A**). For the RecA2 domain, the expected 15 Ile, 2 Met, 12 Ala and 18 Val methyl groups as well as at least 35 out of 38 expected Leu resonances are visible (**Figure 4.12 B**).

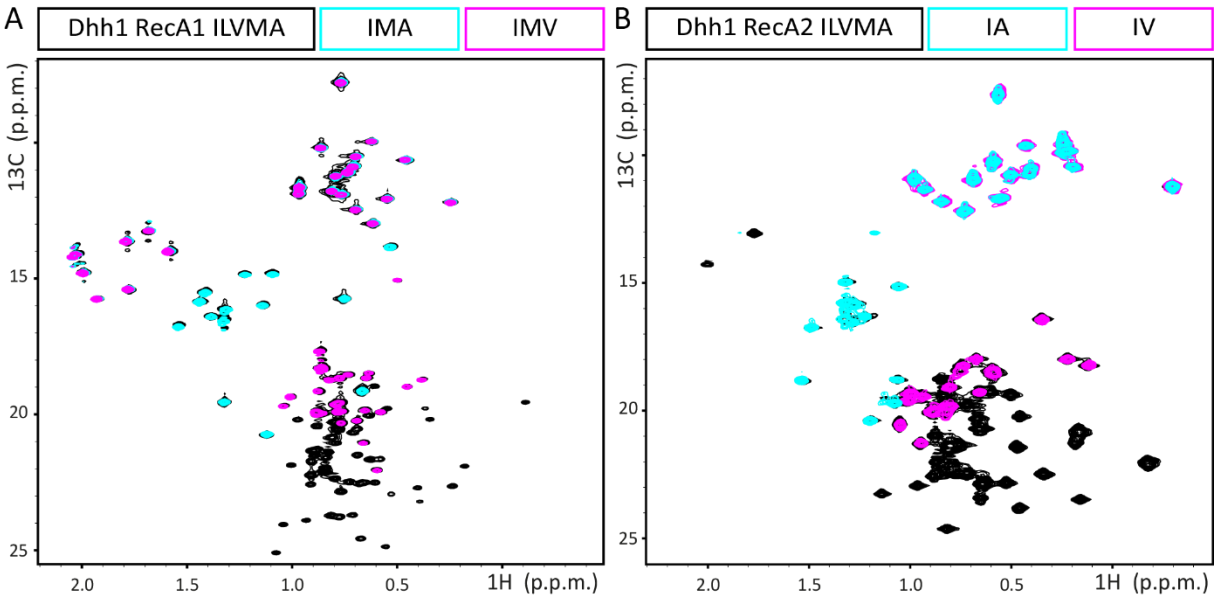

Figure 4.12: Residue type-specific assignment of the Dhh1 RecA-like domains. (A) Three methyl-labeled samples were prepared for the residue type-specific resonance assignment of the Dhh1 RecA1 domain: ILVMA (black), IMA (cyan) and IMV (magenta). Peaks colored in black only correspond to Leu methyl groups. Resonances in cyan and black correspond to Ala methyl groups, while overlays of magenta and black peaks correspond to Val methyl groups. Peaks appearing in black, cyan and magenta correspond to Ile and Met. Resonances of both residue types are distinguished by their location in the spectrum. Occurrence of Ile resonances in the typical Met region and *vice versa* is scarce. **(B)** Analogously to the RecA1 domain, ILVMA- (black), IA- (cyan) and IV-labeled samples (magenta) were prepared to assign methyl groups of the RecA2 domain in a residue type-specific manner. Resonances in cyan and black correspond to Ala residues and resonances in black and magenta resemble Val methyl groups. Overlays of cyan, magenta and black peaks correspond to Ile residues. Peaks appearing only in black correspond to either Met or Leu. Resonances of both residue types are easily distinguished by their location in the spectrum. Overlap of regions where Met and Leu resonances appear is not observed.

The MAGIC-based methyl group assignment of the RecA1 domain was based on the crystal structure of the Dhh1 helicase core (pdb: 1s2m) and on HMQC-NOESY-HMQC data that was recorded with NOE mixing times of 40 and 300 ms. The two strongest methionine resonances in the RecA1 methyl TROSY spectrum (**Figure 4.13 A**, dashed arrows) correspond to highly flexible residues. We thus assigned these to the N-terminal methionine M29, which originates from the cloning site, and to the C-terminal methionine M250, which is part of the short linker region between the two RecA-like domains. Accordingly, the strongest alanine resonance was assigned to the N-terminal alanine residue A28 (**Figure 4.13 A**, open arrow), which originates from the cloning site as well. During the automated assignment process these assignments were fixed. The MAGIC algorithm subsequently assigned all isoleucine (16), methionine (8) and alanine (14) resonances. For three out of the 26 visible valine resonances and eight out of the 47 visible leucine methyl groups an assignment was not possible.

For the methyl resonance assignment of the Dhh1 RecA2 domain, HMQC-NOESY-HMQC experiments with NOE mixing times of 60 and 300 ms were recorded. This NMR data together with a crystal structure of the Dhh1 helicase core (pdb: 1s2m) provided the basis for the assignment process which was also based on the MAGIC algorithm.

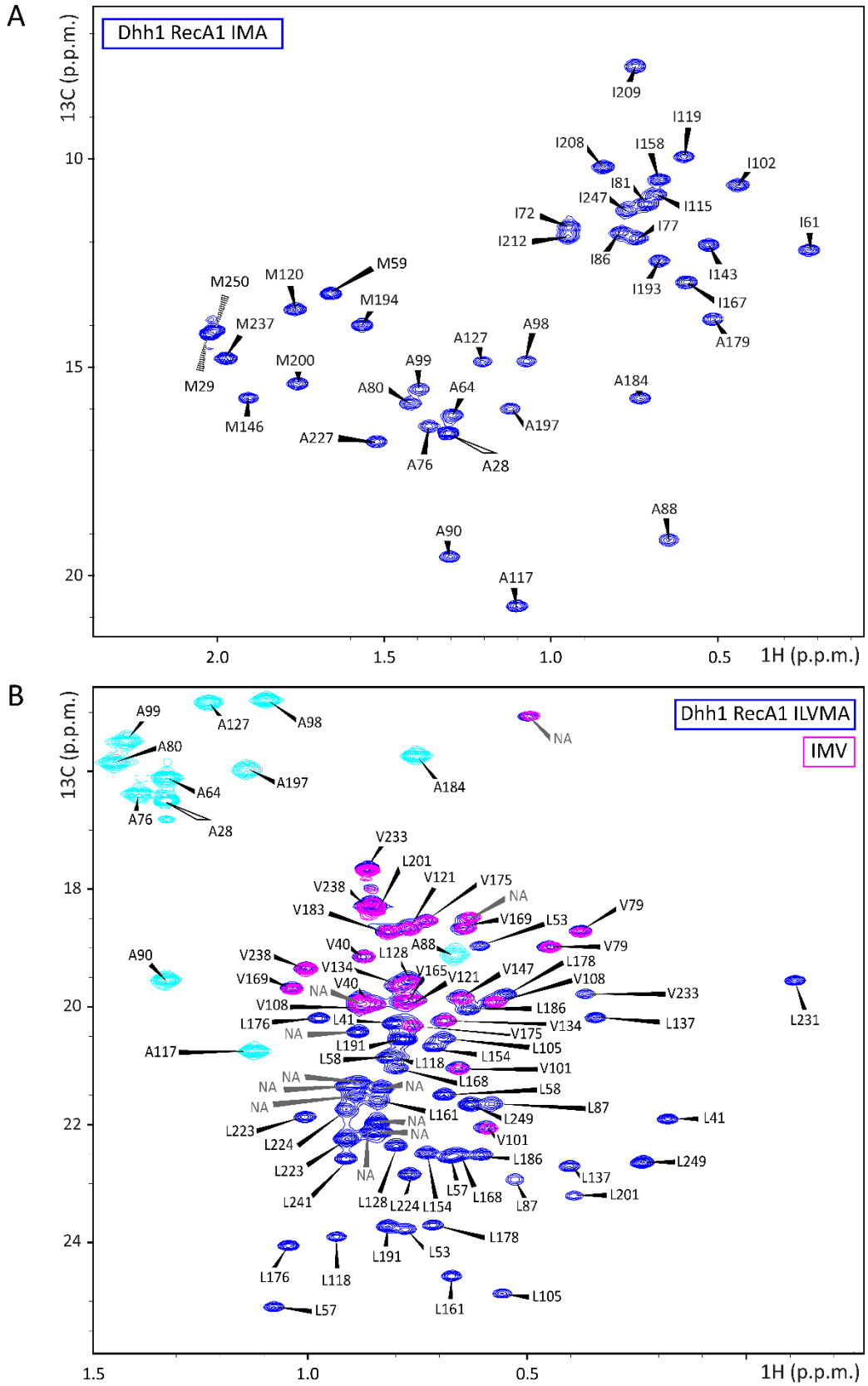

Figure 4.13: Methyl group assignment of the Dhh1 RecA1 domain obtained by the MAGIC algorithm. HMQC spectra of an **(A)** IMA-labeled sample and **(B)** a superposition of spectra obtained on IMV- (magenta) and ILVMA- (blue; for clarity, Ala resonances are in cyan) labeled samples. Assignments as obtained from MAGIC are indicated. The assignment was based on two HMQC-NOESY-HMQC experiments measured on an ILVMA-labeled sample with long and short NOE mixing times of 300 and 40 ms, respectively, and a 2.1 Å crystal structure of the Dhh1 helicase core (pdb: 1s2m). Methyl groups of Ala, Ile and Met residues are assigned to 100%. The assignment of the two most intense methionine resonances (dashed arrows) and the most intense alanine resonance (open arrow) was fixed to the terminal residues A28, M29, M250. Resonances that were not assigned (NA) by MAGIC are indicated with gray labels and arrows.

The strongest alanine resonances in the RecA2 methyl TROSY spectrum (**Figure 4.14 A**, open arrow) corresponds to a highly flexible residue. We thus assigned it to the N-terminal alanine A249, which originates from the cloning site. Note that the cloning site-derived residues A28 and A249 in the RecA1 and RecA2 domain, respectively, resonate at the same frequency (^{1}H: 1.33 ppm; ^{13}C: 16.4 ppm; compare **Figure 4.13** and **Figure 4.14**) as they are both part of the unstructured N-terminal GAM-tripeptide and thus have a very similar chemical environment. For a single Ala resonance, no NOE could be detected. We thus assigned it to the isolated A411 (**Figure 4.14**, dashed arrow), whose methyl group is separated by 8.6 Å from the closest methyl bearing amino acid. During the automated assignment process the assignments of A249 and A411 were fixed.

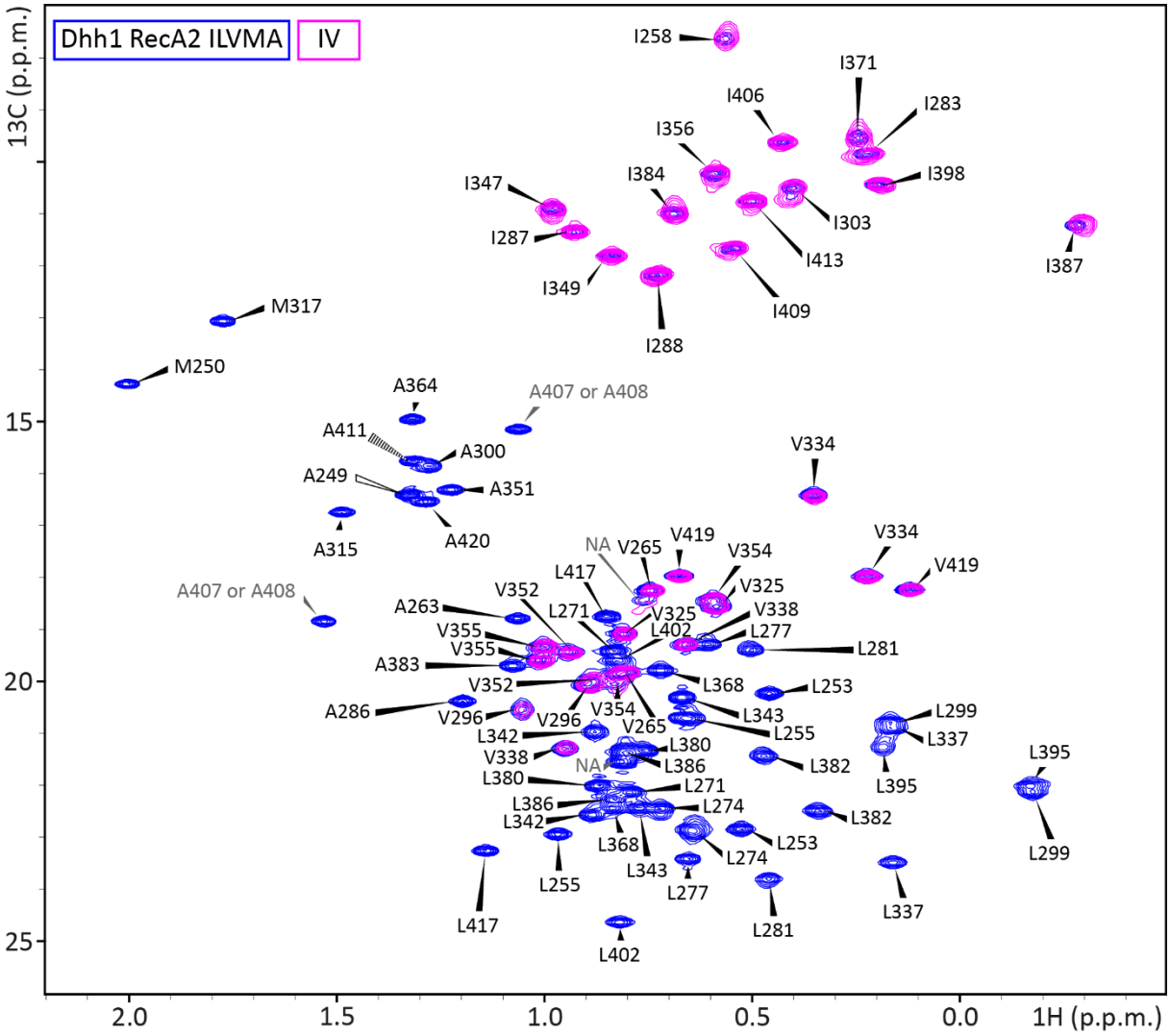

Figure 4.14: Methyl resonance assignment of the Dhh1 RecA1 domain. Superposition of spectra obtained on IV- (magenta) and ILVMA- (blue) labeled samples. Assignments as obtained from MAGIC are indicated. The assignment was based on two HMQC-NOESY-HMQC experiments measured on an ILVMA-labeled sample with long and short NOE mixing times of 300 and 60 ms, respectively, and a 2.1 Å crystal structure of the Dhh1 helicase core (pdb: 1s2m). Methyl groups of Ile, Met and Val residues are assigned to 100%. The assignment of the most intense alanine resonance (open arrow) and only alanine resonance for which no NOE could be detected (open arrow) was fixed to A249 and A411, respectively. Resonances that were not assigned (NA) or for which no unambiguous assignment could be obtained by MAGIC are indicated in gray.

All isoleucine (15), methionine (2) and valine (18) methyl groups have been assigned by MAGIC (**Figure 4.14**). For two out of the twelve alanine resonances an unambiguous assignment was not possible. These two resonances correspond to the A407 and A408 methyl groups, which are separated by 5.3 Å (**Figure 4.15**). Between these two methyl groups we could detect an NOE. However, the next closest methyl group (A263) is over 8 Å apart such that no NOE could be detected. Hence, the isolation of the two Ala methyl groups prevented the unambiguous assignment of the two corresponding resonances. Furthermore, 34 out of 35 visible leucine resonances (corresponding to 17 out of 19 leucine residues) have been assigned (**Figure 4.14**).

In **Figure 4.15**, the confidence of the computationally obtained assignments for the RecA1 and RecA2 domains is indicated. Individual methyl groups are depicted as spheres and colored from red to blue according to the completeness (low to high) of the NOE assignment for the respective NOESY strip. Additionally, the methyl-methyl connections, depicted as dashed lines, as constructed by MAGIC are color-coded according to their confidence score from red (low) to blue (high). It is important to note that most resonances that were assigned by MAGIC are of the high confidence class. Not unexpectedly, isolated methyl groups were either not assigned (gray) or the confidence level was very low (red).

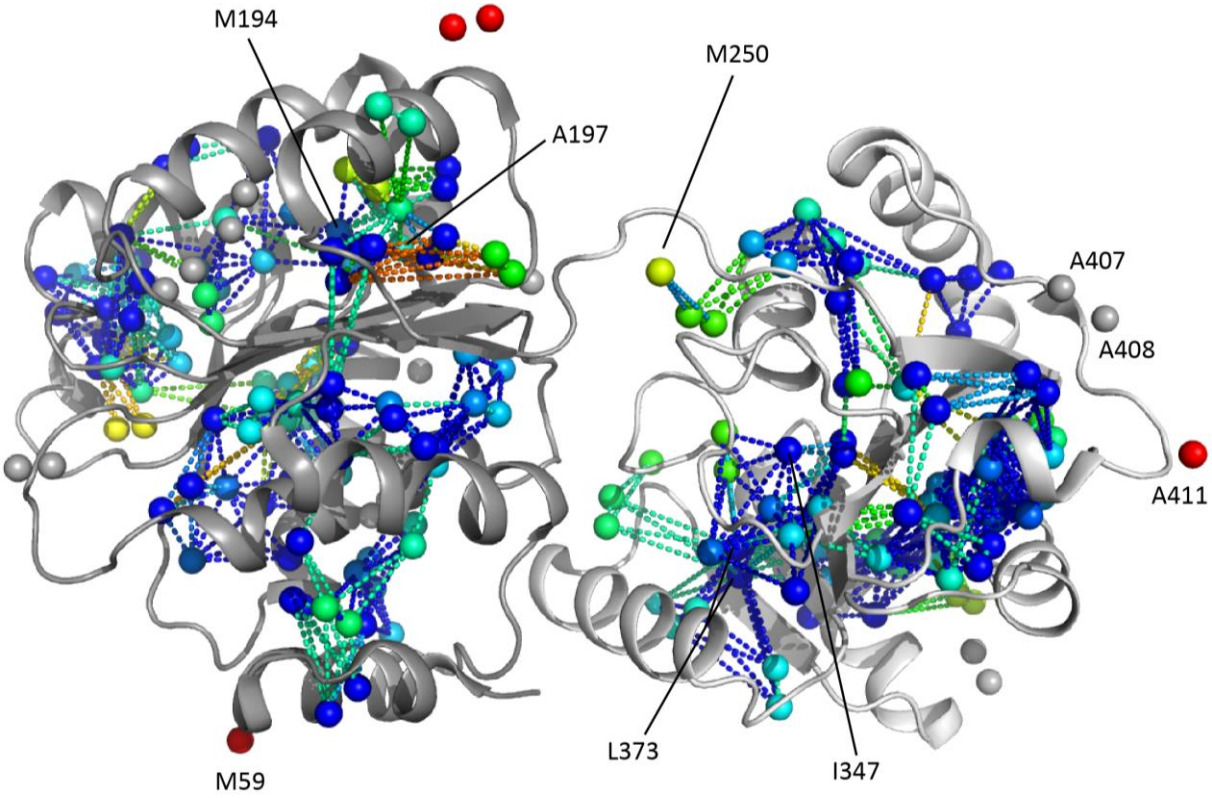

Figure 4.15: Confidence of MAGIC assignments for the Dhh1 methyl groups. The Dhh1 helicase core is displayed in a cartoon representation with the RecA1 in dark gray and the RecA2 in light gray (pdb: 1s2m). The orientation is as in **Figure 4.1 A**. Each methyl group is displayed as a sphere with a color code ranging from red to blue representing the completeness (0 – 100 %) of the NOE assignment for the related NOESY strip. Methyl-methyl connections are depicted as dashed lines and are colored according to the confidence score of the peak-peak connection constructed by MAGIC. The color gradient ranges from red (confidence score of 0) to blue (confidence score ≥ 2). Unassigned methyl groups are depicted as grey spheres. No electron density is observed for residues preceding N46 and succeeding; this includes A28 and M29 as well as V40 and L41.

4.3.5 *Residues in the RecA-like domains sense the phase-separated state*

To independently assess which residues are affected by the liquid-liquid phase separation process of the Dhh1 helicase core, we prepared NMR samples under non-phase separation and under phase separation conditions. In brief, we mixed ILVMA-labeled and unlabeled Dhh1 in a 1:3 molar ratio and concentrated the protein solution to 300 µM and induced phase separation by lowering the salt concentration. Methyl TROSY spectra of these samples were recorded immediately and compared with spectra obtained for ILVMA-labeled Dhh1 under non-phase separating conditions.

An overlay of spectra obtained under non-phase separation and under phase separation conditions revealed small but significant CSPs for a subset of resonances (**Figure 4.16**). The corresponding residues are located both in the RecA1 and RecA2 domain. Assignments obtained with the MAGIC approach (see above) identified them as M59, M194 and A197 in the RecA1 domain. While M59 is solvent-exposed, M194 and A197 are located in the DEAD-motif that is involved in ATP-binding and hydrolysis (**Figure 4.15**, see also **Figure 4.1**). Additionally, I347 and L337 in the RecA2 domain are affected by LLPS of the Dhh1 helicase core protein. Future research will explore how these two buried residues (**Figure 4.15**) sense the protein-protein contacts that take place at the LLPS boundary.

4.4 *DISCUSSION*

We recently showed that a synergistic network of protein-protein and protein-RNA interactions in the mRNA degradation machinery results in the formation of *in vitro* processing bodies (P-bodies)[252]. This network effects the mRNA decapping activity of the Dcp1:Dcp2 complex as long mRNAs are protected against decapping under phase separation conditions. These findings supported the hypothesis that P-bodies function as mRNA storage compartments[307,325,349,350]. Dhh1 is a conserved DEAD-box helicase with high cellular abundance that is involved in mRNA decay and P-body formation. *In vivo* P-body formation can be reproduced *in vitro* in a liquid-liquid phase separation process[103,252].

In a previous study, Dhh1 was reported to undergo *in vitro* phase separation in a strictly ATP- and RNA-dependent manner[164]. Here, we establish that the isolated helicase core of Dhh1 undergoes LLPS at concentrations as low as 25 µM and in the absence of ATP and RNA. As many other systems, apo Dhh1 phase separates at low salt concentrations, indicative for a phase separation process that is dominated by charge-charge and cation-π interactions[119]. We also find that Dhh1 phase separation is promoted by high pH (in the range of 7.4 as compared to 6.0). Notably, the concentrations we used are higher by a factor of 10-30 compared with LLPS experiments that used full-length, mCherry-tagged Dhh1 in the presence of ATP and poly(U) RNA[164].

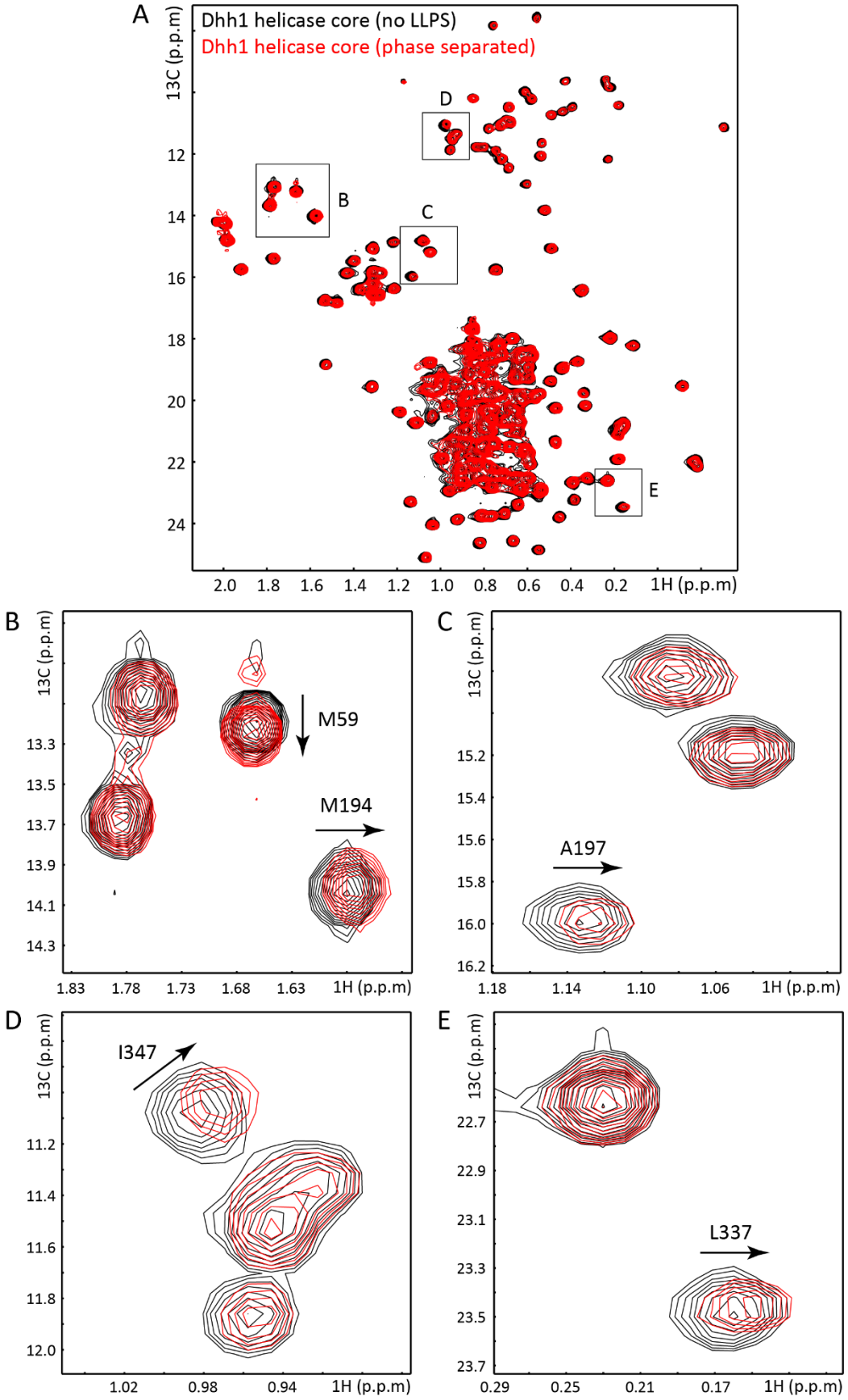

Figure 4.16: Residues in the RecA-like domains of Dhh1 sense the phase separated state. (A) Superposition of methyl TROSY spectra of Dhh1 under non-phase separation conditions (black) and after phase separation (red). A few resonances experience CSPs in the phase-separated state relative to the non-phase separated state. Regions containing such resonances are highlighted by boxes and are enlarged in the panels **(B)-(E)**. Assignments of perturbed resonances are indicated.

These observations indicate that Dhh1 phase separation receives contributions from (1) the helicase core, (2) the N- and C-terminal extensions and (3) RNA and ATP. Additionally, the mCherry-tag used in the previous report[164] might confer solubility and thereby increase the protein concentrations that are required to cross the phase boundary.

Using the homologous *C. thermophilum* Dhh1 protein with different phase separation behavior as a toolbox, we identified the RecA2 domain to be the driving force for phase separation of *S. cerevisiae* Dhh1. Thereby, two aspartate residues in the C-terminus of the *S. cerevisiae* RecA2 domain appeared to be essential for the phase separation process, as their mutation to prolines as in the *C. thermophilum* homolog abolished phase separation completely. These results demonstrate that even small changes in the amino acid sequence can have a drastic impact on the liquid-liquid phase separation behavior of a protein. Liquid droplets are highly dynamic entities, where subtle environmental changes or interference with the interaction network can rapidly influence the degree of LLPS.

We extended our LLPS studies from the Dhh1 helicase core to constructs that include its terminal extensions. Phase separation is increased relative to the helicase core for Dhh1 constructs that harbor either one or both disordered terminal extensions, whereof full-length Dhh1 shows the highest degree of LLPS. The importance of the prion-like N- and C-terminal extensions for efficient *in vitro* LLPS of full-length Dhh1 agrees with *in vivo* studies, where the deletion of the C-terminal extension resulted in a three-fold reduction of P-body assembly relative to the wild-type protein[146].

Moreover, phase separation of full-length Dhh1 is enhanced in the presence of 2.5 mM ATP relative to the apo protein. Remarkably, ATP has been suggested as a general hydrotrope at physiological concentrations between 5 and 10 mM[160]. As such, ATP is thought to prevent the formation of liquid droplets, to promote their disassembly and to confer solubility for a high fraction of the human proteome[160,162]. Our results indicate that this property of ATP might not apply to all protein aggregates or phase separated assemblies. Future experiments will address the question if ATP also acts on Dhh1 LLPS as a general hydrotrope at concentrations between 5-10 mM. It has been suggested that ATP-binding and ATP-dependent RNA-binding to Dhh1 result in a conformational rearrangement of the two RecA-like domains[318]. The increase in Dhh1 phase separation that we observe in the presence of ATP might thus be due to a different orientation of both RecA-like domain, which facilitates intermolecular contacts.

Dhh1 in isolation and Dhh1 in the presence of RNA undergo maximum phase separation under significantly different conditions. In the presence of RNA, we observe Dhh1 phase separation at high salt (75-100 mM) and low pH (6.0-6.5), while apo Dhh1 phase separates at low salt and high pH.

This argues for different molecular mechanisms that are responsible for Dhh1 phase separation in the free and in the RNA-bound form. Additionally, Dhh1 phase separation in the presence of RNA is higher relative to the apo form, indicative for an increase in intermolecular interactions in RNA-bound Dhh1. These results also clearly indicate the importance of performing phase separation assays in a two-dimensional grid of conditions, as drastic changes in the phase separation pattern cannot be detected if LLPS is monitored at a single condition. Conclusions that are drawn based on observations made for a single condition might thus be misleading.

The molecular basis by with RNA enhances Dhh1 phase separation is not clear so far. RNA might use three potential binding sites on Dhh1 to enhance phase separation: First, the canonical RNA interaction surface that is found in many DEAD-box helicases, where single-stranded RNA induces a closure of the two RecA-like domains in an ATP-dependent manner[319–322]. The closed state is supported by the RNA binding site that runs across both domains on the surface of the helicase core. It is thus tempting to speculate that in our phase separation experiments the single-stranded 30mer RNA also binds to the canonical RNA binding surface and potentially closes both domains. This domain reorientation can result in enhanced intermolecular contacts which further enhance LLPS. Second, surface patches on the RecA2 domain that bind RNA have been identified[105]. Interestingly, these patched are also responsible for the interaction with Pat1 and Edc3 FG and FDF motifs. This suggests a mutually exclusive interaction of Dhh1 with RNA, Pat1 and Edc3[105]. We find that Pat1 or Edc3 reduced Dhh1 phase separation. The molecular mechanism by which this takes place is likely through interference with intermolecular contacts that are important for LLPS. Third, the disordered C-terminal extension of Dhh1 contains 13 aromatic residues (His, Phe and Tyr) which could potentially engage in π-π-interactions with the RNA bases. Further NMR-driven studies will explore which of the three putative RNA binding sites is able to bind RNA and which interactions contribute to LLPS of the Dhh1:RNA complex.

To obtain residue specific information on the LLPS process of Dhh1 we here use methyl TROSY NMR techniques and show that a small number of methyl groups is involved in the initial process of LLPS. Based on initial assignments, these residues can be localized to both RecA domains, in agreement with the importance of both domains for LLPS. Future studies will reveal how these sites sense the phase-separated state and how they interact in detail to mediate LLPS.

In summary, we here show that the Dhh1 protein exploits a large number of intermolecular interactions to induce LLPS. These interactions include the folded RecA domains, the disordered N- and C-terminal tails as well as contacts with RNA and ATP. Interestingly, the LLPS behavior of Dhh1 can be modulated by Edc3 and Pat1, showing that the formation of cellular mRNP foci is a result of multiple constructive and destructive interactions.

CHAPTER 5 *Conclusion*

This thesis provides a rationale how several distinct inter- and intramolecular interactions, which on their own contribute only weakly to overall LLPS and do so only under non-physiological conditions, act together in an additive manner to mediate the formation of membrane-less organelles in a cellular setting. The presented results explicitly do not rule out the possibility that these interactions show cooperativity *in vivo*.

Importantly, a high-throughput assay has been developed in this thesis that allows for the rapid, reliable and resource-efficient screening of hundreds of conditions for LLPS making use of robotic small-volume pipetting and turbidity measurements.

Further, it has been demonstrated in this work that the same interactions that result in LLPS are preserved in and contribute to a second phase transition towards a more gel-like state. For the Edc3 protein, which is composed of disordered regions and well-folded domains, this maturation process is not associated with the formation of disease-related amyloid-like fibers that are frequently observed for phase-separating IDPs containing low-complexity regions or prion-like domains. These results indicate that interactions other than fibrillization can also contribute to the maturation of phase-separated liquid droplets.

Additionally, it was shown here that the enzymatic activity for the Dcp2 decapping enzyme is reduced when the protein and the capped RNA are localized to liquid droplets compared to when they are free in solution. An important aspect for future research will be to determine how the ATPase activity of Dhh1 is modulated under phase separation conditions and if Dhh1 possesses helicase activity in liquid droplets. Based on these results it will be possible in the future to propose models how cellular phase separation effects the function and activity of enzymes within membrane-less organelles.

In summary, this thesis contributes to the deeper understanding of the biological function of processing bodies and other cellular foci that form via liquid-liquid phase separation processes.

FUNDING

Work in Chapter 2 was supported by the IMPRS "From Molecules to Organisms" (to Stefan Schütz and Erik Nöldeke), the Max Planck Society and the European Research Council (ERC) under the European Union's Seventh Framework Programme (FP7/2007–2013; ERC grant agreement no. 616052 to Remco Sprangers).

Work in Chapter 3 was supported by the Max Planck Society and the European Research Council (ERC) under the European Union's Seventh Framework Programme (FP7/2007–2013), ERC Grant 616052 (to Remco Sprangers). Further funding was provided by Netherlands Organization for Scientific Research (NWO) (grants 723.014.003 to Markus Weingarth and grants 700.26.121 and 700.10.443 to Marc Baldus) and iNEXT (project number 653706), a Horizon 2020 program of the European Union.

Work in Chapter 4 was funded by the European Union's Seventh Framework Programme (FP7/2007–2013; ERC grant agreement no. 616052 to Remco Sprangers). I further acknowledge funding for a student assistant (Studentische Hilfskraft) by the Regensburg International Graduate School of Life Sciences (RIGeL).

ACKNOWLEDGEMENTS

Remco – my deep gratitude for having me as a PhD student, for your supervision and advice. I appreciate the way you do and teach science and I learned a lot from you in the past years. Besides being a great mentor, with your personality you make working in your lab also quite fun.

Ralf-Peter Jansen – thank you very much for giving advice and support as a member of my Thesis Advisory Committee and for coming to Regensburg for my defense.

Thilo Stehle – thank you for joining my TAC in Tübingen and for giving me the opportunity to get in touch with structural biology almost 10 years ago.

Ancilla, Daniela, Erik, Fabian, Jobst, Karolina, Magnus, Maxime, Milos, Natalia, Samira and all the other people who joined *the Sprangers and Wiesner Labs* for longer or shorter periods – thank you for the friendly working atmosphere, for joint lab meetings, fruitful off- and on-topic discussions, relaxing Happy Hours and distracting coffee breaks.

Iris Holdermann, Janina Petters, Johanna Stöfl and Mira Schütz-Stoffregen – special thanks for your excellent technical assistance throughout the years at the MPI and the University of Regensburg, respectively. Thank you for taking care about the lab and that we do not run out of anything - and for refilling what we took out. Thank you, Iris and Mira, for making the lab move to Regensburg as smooth as possible.

Anna – thank you for introducing me to the RNA world, for our joint time in the writing office and for proof-reading this thesis.

Carsten – thank you for taking extra care of our computers, despite of your own lab work.

Jan – thank you for being a real challenge at our daily table soccer matches.

Olga – special thanks for your help with protein expression and purification. It was great to have you as a Hiwi.

Philip – thank you for your help with Sparky, MAGIC and NMR in general. And for your cakes.

The 3rd floor people – thank you for enjoyable lunch breaks at the MPI, for great discussions and for input from the dark side of science, related to plants and stuff ;)

Sarah Danes, Dagmar Sigurdardottir and Kinga Ay – thank you for your commitment to the PhD programs (IMPRS and RIGeL) and for always being available for questions.

Vincent Truffault – thank you for your commitment in maintaining the NMR infrastructure at the MPI for Developmental Biology (Tübingen).

Werner Kremer and Simon Grunert – thanks for taking care about the magnets at the University of Regensburg.

Christian Liebig – thank you for Light Microscopy support at the MPI for Developmental Biology.

To all my friends, no matter if close by or far away – thank you for being a reliable support in everyday life, for giving advice and (especially) for everything that is not related to work. With you, life is even more worth living.

Ein großer Dank gebührt *meiner Familie* und insbesondere *meinen Eltern*. Ich möchte euch für eure fortwährende Unterstützung danken und dafür, dass ihr mich nie unter Druck gesetzt habt, möglichst schnell fertig zu werden. Danke für euer Verständnis, eure Zuversicht und Geduld, auch wenn „Biochemie“ für euch wohl immer etwas Abstraktes bleiben und auch diese Doktorarbeit daran vermutlich nichts ändern wird.

Thank you, *Mira*, my dear wife, for your love and your faith in us. For all your efforts, for carefully proof-reading this thesis and for your patience with me, not always knowing what comes next. I look forward to our growing little family - I love you truly!

Pascal – Danke, dass Du mir jeden Tag ein Lächeln ins Gesicht zauberst. Ich bin immer für Dich da!

REFERENCES

(1) Cramer, P. Structure and Function of RNA Polymerase II. In *Advances in Protein Chemistry*; Elsevier, 2004; Vol. 67, pp 1–42. https://doi.org/10.1016/S0065-3233(04)67001-X.

(2) Shatkin, A. J. Capping of Eucaryotic MRNAs. *Cell* **1976**, *9* (4 PT 2), 645–653.

(3) Andrecka, J.; Lewis, R.; Bruckner, F.; Lehmann, E.; Cramer, P.; Michaelis, J. Single-Molecule Tracking of MRNA Exiting from RNA Polymerase II. *Proc. Natl. Acad. Sci.* **2008**, *105* (1), 135–140. https://doi.org/10.1073/pnas.0703815105.

(4) Coppola, J. A.; Field, A. S.; Luse, D. S. Promoter-Proximal Pausing by RNA Polymerase II in Vitro: Transcripts Shorter than 20 Nucleotides Are Not Capped. *Proc. Natl. Acad. Sci. U. S. A.* **1983**, *80* (5), 1251–1255. https://doi.org/10.1073/pnas.80.5.1251.

(5) Proudfoot, N. Connecting Transcription to Messenger RNA Processing. *Trends Biochem. Sci.* **2000**, *25* (6), 290–293.

(6) Cho, E. J.; Takagi, T.; Moore, C. R.; Buratowski, S. MRNA Capping Enzyme Is Recruited to the Transcription Complex by Phosphorylation of the RNA Polymerase II Carboxy-Terminal Domain. *Genes Dev.* **1997**, *11* (24), 3319–3326.

(7) McCracken, S.; Fong, N.; Rosonina, E.; Yankulov, K.; Brothers, G.; Siderovski, D.; Hessel, A.; Foster, S.; Shuman, S.; Bentley, D. L. 5′-Capping Enzymes Are Targeted to Pre-MRNA by Binding to the Phosphorylated Carboxy-Terminal Domain of RNA Polymerase II. *Genes Dev.* **1997**, *11* (24), 3306–3318.

(8) Furuichi, Y.; LaFiandra, A.; Shatkin, A. J. 5′-Terminal Structure and MRNA Stability. *Nature* **1977**, *266* (5599), 235–239.

(9) Stevens, A. An Exoribonuclease from Saccharomyces Cerevisiae: Effect of Modifications of 5′ End Groups on the Hydrolysis of Substrates to 5′ Mononucleotides. *Biochem. Biophys. Res. Commun.* **1978**, *81* (2), 656–661.

(10) Ghosh, A.; Lima, C. D. Enzymology of RNA Cap Synthesis. *Wiley Interdiscip. Rev. RNA* **2010**, *1* (1), 152–172. https://doi.org/10.1002/wrna.19.

(11) Bird, J. G.; Zhang, Y.; Tian, Y.; Panova, N.; Barvík, I.; Greene, L.; Liu, M.; Buckley, B.; Krásný, L.; Lee, J. K.; et al. The Mechanism of RNA 5′ Capping with NAD+, NADH and Desphospho-CoA. *Nature* **2016**, *535* (7612), 444–447. https://doi.org/10.1038/nature18622.

(12) Walters, R. W.; Matheny, T.; Mizoue, L. S.; Rao, B. S.; Muhlrad, D.; Parker, R. Identification of NAD+ Capped MRNAs in Saccharomyces Cerevisiae. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114* (3), 480–485. https://doi.org/10.1073/pnas.1619369114.

(13) Shi, Y. Mechanistic Insights into Precursor Messenger RNA Splicing by the Spliceosome. *Nat. Rev. Mol. Cell Biol.* **2017**, *18* (11), 655–670. https://doi.org/10.1038/nrm.2017.86.

(14) Fica, S. M.; Nagai, K. Cryo-Electron Microscopy Snapshots of the Spliceosome: Structural Insights into a Dynamic Ribonucleoprotein Machine. *Nat. Struct. Mol. Biol.* **2017**, *24* (10), 791–799. https://doi.org/10.1038/nsmb.3463.

(15) Wickens, M. How the Messenger Got Its Tail: Addition of Poly(A) in the Nucleus. *Trends Biochem. Sci.* **1990**, *15* (7), 277–281.

(16) Dheur, S.; Nykamp, K. R.; Viphakone, N.; Swanson, M. S.; Minvielle-Sebastia, L. Yeast MRNA Poly(A) Tail Length Control Can Be Reconstituted *in Vitro* in the Absence of Pab1p-Dependent Poly(A) Nuclease Activity. *J. Biol. Chem.* **2005**, *280* (26), 24532–24538. https://doi.org/10.1074/jbc.M504720200.

(17) Viphakone, N.; Voisinet-Hakil, F.; Minvielle-Sebastia, L. Molecular Dissection of MRNA Poly(A) Tail Length Control in Yeast. *Nucleic Acids Res.* **2008**, *36* (7), 2418–2433. https://doi.org/10.1093/nar/gkn080.

(18) Wahle, E. Poly(A) Tail Length Control Is Caused by Termination of Processive Synthesis. *J. Biol. Chem.* **1995**, *270* (6), 2800–2808. https://doi.org/10.1074/jbc.270.6.2800.

(19) Goldstrohm, A. C.; Wickens, M. Multifunctional Deadenylase Complexes Diversify MRNA Control. *Nat. Rev. Mol. Cell Biol.* **2008**, *9* (4), 337–344. https://doi.org/10.1038/nrm2370.

(20) Lima, S. A.; Chipman, L. B.; Nicholson, A. L.; Chen, Y.-H.; Yee, B. A.; Yeo, G. W.; Coller, J.; Pasquinelli, A. E. Short Poly(A) Tails Are a Conserved Feature of Highly Expressed Genes. *Nat. Struct. Mol. Biol.* **2017**, *24* (12), 1057–1063. https://doi.org/10.1038/nsmb.3499.

(21) Dávila López, M.; Samuelsson, T. Early Evolution of Histone MRNA 3′ End Processing. *RNA N. Y. N* **2008**, *14* (1), 1–10. https://doi.org/10.1261/rna.782308.

(22) Mignone, F.; Gissi, C.; Liuni, S.; Pesole, G. Untranslated Regions of MRNAs. *Genome Biol.* **2002**, *3* (3), REVIEWS0004.

(23) Araujo, P. R.; Yoon, K.; Ko, D.; Smith, A. D.; Qiao, M.; Suresh, U.; Burns, S. C.; Penalva, L. O. F. Before It Gets Started: Regulating Translation at the 5′ UTR. *Comp. Funct. Genomics* **2012**, *2012*, 475731. https://doi.org/10.1155/2012/475731.

(24) Vilela, C.; McCarthy, J. E. G. Regulation of Fungal Gene Expression via Short Open Reading Frames in the MRNA 5′untranslated Region. *Mol. Microbiol.* **2003**, *49* (4), 859–867.

(25) Barrett, L. W.; Fletcher, S.; Wilton, S. D. Regulation of Eukaryotic Gene Expression by the Untranslated Gene Regions and Other Non-Coding Elements. *Cell. Mol. Life Sci. CMLS* **2012**, *69* (21), 3613–3634. https://doi.org/10.1007/s00018-012-0990-9.

(26) Jackson, R. J. Cytoplasmic Regulation of MRNA Function: The Importance of the 3′ Untranslated Region. *Cell* **1993**, *74* (1), 9–14.

(27) Chatterjee, S.; Pal, J. K. Role of 5′- and 3′-Untranslated Regions of MRNAs in Human Diseases. *Biol. Cell* **2009**, *101* (5), 251–262. https://doi.org/10.1042/BC20080104.

(28) Babendure, J. R.; Babendure, J. L.; Ding, J.-H.; Tsien, R. Y. Control of Mammalian Translation by MRNA Structure near Caps. *RNA N. Y. N* **2006**, *12* (5), 851–861. https://doi.org/10.1261/rna.2309906.

(29) Zhou, W.; Edelman, G. M.; Mauro, V. P. Transcript Leader Regions of Two Saccharomyces Cerevisiae MRNAs Contain Internal Ribosome Entry Sites That Function in Living Cells. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98* (4), 1531–1536. https://doi.org/10.1073/pnas.98.4.1531.
(30) López-Lastra, M.; Rivas, A.; Barría, M. I. Protein Synthesis in Eukaryotes: The Growing Biological Relevance of Cap-Independent Translation Initiation. *Biol. Res.* **2005**, *38* (2–3). https://doi.org/10.4067/S0716-97602005000200003.
(31) Wilkie, G. S.; Dickson, K. S.; Gray, N. K. Regulation of MRNA Translation by 5'- and 3'-UTR-Binding Factors. *Trends Biochem. Sci.* **2003**, *28* (4), 182–188. https://doi.org/10.1016/S0968-0004(03)00051-3.
(32) Szostak, E.; Gebauer, F. Translational Control by 3'-UTR-Binding Proteins. *Brief. Funct. Genomics* **2013**, *12* (1), 58–65. https://doi.org/10.1093/bfgp/els056.
(33) Bartel, D. P. Metazoan MicroRNAs. *Cell* **2018**, *173* (1), 20–51. https://doi.org/10.1016/j.cell.2018.03.006.
(34) Huntzinger, E.; Izaurralde, E. Gene Silencing by MicroRNAs: Contributions of Translational Repression and MRNA Decay. *Nat. Rev. Genet.* **2011**, *12* (2), 99–110. https://doi.org/10.1038/nrg2936.
(35) Jinek, M.; Doudna, J. A. A Three-Dimensional View of the Molecular Machinery of RNA Interference. *Nature* **2009**, *457* (7228), 405–412. https://doi.org/10.1038/nature07755.
(36) Lewis, J. D.; Izaurralde, E. The Role of the Cap Structure in RNA Processing and Nuclear Export. *Eur. J. Biochem.* **1997**, *247* (2), 461–469.
(37) Eckner, R.; Ellmeier, W.; Birnstiel, M. L. Mature MRNA 3' End Formation Stimulates RNA Export from the Nucleus. *EMBO J.* **1991**, *10* (11), 3513–3522.
(38) Topisirovic, I.; Svitkin, Y. V.; Sonenberg, N.; Shatkin, A. J. Cap and Cap-Binding Proteins in the Control of Gene Expression. *Wiley Interdiscip. Rev. RNA* **2011**, *2* (2), 277–298. https://doi.org/10.1002/wrna.52.
(39) Deo, R. C.; Bonanno, J. B.; Sonenberg, N.; Burley, S. K. Recognition of Polyadenylate RNA by the Poly(A)-Binding Protein. *Cell* **1999**, *98* (6), 835–845.
(40) Wells, S. E.; Hillner, P. E.; Vale, R. D.; Sachs, A. B. Circularization of MRNA by Eukaryotic Translation Initiation Factors. *Mol. Cell* **1998**, *2* (1), 135–140.
(41) Gallie, D. R. The Cap and Poly(A) Tail Function Synergistically to Regulate MRNA Translational Efficiency. *Genes Dev.* **1991**, *5* (11), 2108–2116.
(42) Borman, A. M.; Michel, Y. M.; Kean, K. M. Biochemical Characterisation of Cap-Poly(A) Synergy in Rabbit Reticulocyte Lysates: The EIF4G-PABP Interaction Increases the Functional Affinity of EIF4E for the Capped MRNA 5'-End. *Nucleic Acids Res.* **2000**, *28* (21), 4068–4075. https://doi.org/10.1093/nar/28.21.4068.
(43) Jackson, R. J.; Hellen, C. U. T.; Pestova, T. V. The Mechanism of Eukaryotic Translation Initiation and Principles of Its Regulation. *Nat. Rev. Mol. Cell Biol.* **2010**, *11* (2), 113–127. https://doi.org/10.1038/nrm2838.
(44) Schwartz, D. C.; Parker, R. MRNA Decapping in Yeast Requires Dissociation of the Cap Binding Protein, Eukaryotic Translation Initiation Factor 4E. *Mol. Cell. Biol.* **2000**, *20* (21), 7933–7942.
(45) Cakmakci, N. G.; Lerner, R. S.; Wagner, E. J.; Zheng, L.; Marzluff, W. F. SLIP1, a Factor Required for Activation of Histone MRNA Translation by the Stem-Loop Binding Protein. *Mol. Cell. Biol.* **2008**, *28* (3), 1182–1194. https://doi.org/10.1128/MCB.01500-07.
(46) Chapat, C.; Jafarnejad, S. M.; Matta-Camacho, E.; Hesketh, G. G.; Gelbart, I. A.; Attig, J.; Gkogkas, C. G.; Alain, T.; Stern-Ginossar, N.; Fabian, M. R.; et al. Cap-Binding Protein 4EHP Effects Translation Silencing by MicroRNAs. *Proc. Natl. Acad. Sci.* **2017**, *114* (21), 5425–5430. https://doi.org/10.1073/pnas.1701488114.
(47) Kamenska, A.; Simpson, C.; Standart, N. EIF4E-Binding Proteins: New Factors, New Locations, New Roles. *Biochem. Soc. Trans.* **2014**, *42* (4), 1238–1245. https://doi.org/10.1042/BST20140063.
(48) Ozgur, S.; Chekulaeva, M.; Stoecklin, G. Human Pat1b Connects Deadenylation with MRNA Decapping and Controls the Assembly of Processing Bodies. *Mol. Cell. Biol.* **2010**, *30* (17), 4308–4323. https://doi.org/10.1128/MCB.00429-10.
(49) Charenton, C.; Gaudon-Plesse, C.; Fourati, Z.; Taverniti, V.; Back, R.; Kolesnikova, O.; Séraphin, B.; Graille, M. A Unique Surface on Pat1 C-Terminal Domain Directly Interacts with Dcp2 Decapping Enzyme and Xrn1 5'-3' MRNA Exonuclease in Yeast. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114* (45), E9493–E9501. https://doi.org/10.1073/pnas.1711680114.
(50) Sharova, L. V.; Sharov, A. A.; Nedorezov, T.; Piao, Y.; Shaik, N.; Ko, M. S. H. Database for MRNA Half-Life of 19 977 Genes Obtained by DNA Microarray Analysis of Pluripotent and Differentiating Mouse Embryonic Stem Cells. *DNA Res. Int. J. Rapid Publ. Rep. Genes Genomes* **2009**, *16* (1), 45–58. https://doi.org/10.1093/dnares/dsn030.
(51) Wang, Y.; Liu, C. L.; Storey, J. D.; Tibshirani, R. J.; Herschlag, D.; Brown, P. O. Precision and Functional Specificity in MRNA Decay. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99* (9), 5860–5865. https://doi.org/10.1073/pnas.092538799.
(52) Shyu, A. B.; Greenberg, M. E.; Belasco, J. G. The C-Fos Transcript Is Targeted for Rapid Decay by Two Distinct MRNA Degradation Pathways. *Genes Dev.* **1989**, *3* (1), 60–72.
(53) Lowenhaupt, K.; Lingrel, J. B. A Change in the Stability of Globin MRNA during the Induction of Murine Erythroleukemia Cells. *Cell* **1978**, *14* (2), 337–344.
(54) Muhlrad, D.; Parker, R. Premature Translational Termination Triggers MRNA Decapping. *Nature* **1994**, *370* (6490), 578–581. https://doi.org/10.1038/370578a0.
(55) Frischmeyer, P. A.; van Hoof, A.; O'Donnell, K.; Guerrerio, A. L.; Parker, R.; Dietz, H. C. An MRNA Surveillance Mechanism That Eliminates Transcripts Lacking Termination Codons. *Science* **2002**, *295* (5563), 2258–2261. https://doi.org/10.1126/science.1067338.
(56) van Hoof, A.; Frischmeyer, P. A.; Dietz, H. C.; Parker, R. Exosome-Mediated Recognition and Degradation of MRNAs Lacking a Termination Codon. *Science* **2002**, *295* (5563), 2262–2264. https://doi.org/10.1126/science.1067272.

(57) Doma, M. K.; Parker, R. Endonucleolytic Cleavage of Eukaryotic MRNAs with Stalls in Translation Elongation. *Nature* **2006**, *440* (7083), 561–564. https://doi.org/10.1038/nature04530.
(58) Harigaya, Y.; Parker, R. No-Go Decay: A Quality Control Mechanism for RNA in Translation: No-Go Decay. *Wiley Interdiscip. Rev. RNA* **2010**, *1* (1), 132–141. https://doi.org/10.1002/wrna.17.
(59) Parker, R. RNA Degradation in Saccharomyces Cerevisae. *Genetics* **2012**, *191* (3), 671–702. https://doi.org/10.1534/genetics.111.137265.
(60) Garneau, N. L.; Wilusz, J.; Wilusz, C. J. The Highways and Byways of MRNA Decay. *Nat. Rev. Mol. Cell Biol.* **2007**, *8* (2), 113–126. https://doi.org/10.1038/nrm2104.
(61) Decker, C. J.; Parker, R. A Turnover Pathway for Both Stable and Unstable MRNAs in Yeast: Evidence for a Requirement for Deadenylation. *Genes Dev.* **1993**, *7* (8), 1632–1643.
(62) Yamashita, A.; Chang, T.-C.; Yamashita, Y.; Zhu, W.; Zhong, Z.; Chen, C.-Y. A.; Shyu, A.-B. Concerted Action of Poly(A) Nucleases and Decapping Enzyme in Mammalian MRNA Turnover. *Nat. Struct. Mol. Biol.* **2005**, *12* (12), 1054–1063. https://doi.org/10.1038/nsmb1016.
(63) Brown, C. E.; Sachs, A. B. Poly(A) Tail Length Control in Saccharomyces Cerevisiae Occurs by Message-Specific Deadenylation. *Mol. Cell. Biol.* **1998**, *18* (11), 6548–6559.
(64) Wolf, J.; Passmore, L. A. MRNA Deadenylation by Pan2-Pan3. *Biochem. Soc. Trans.* **2014**, *42* (1), 184–187. https://doi.org/10.1042/BST20130211.
(65) Yi, H.; Park, J.; Ha, M.; Lim, J.; Chang, H.; Kim, V. N. PABP Cooperates with the CCR4-NOT Complex to Promote MRNA Deadenylation and Block Precocious Decay. *Mol. Cell* **2018**, *70* (6), 1081-1088.e5. https://doi.org/10.1016/j.molcel.2018.05.009.
(66) Tucker, M.; Valencia-Sanchez, M. A.; Staples, R. R.; Chen, J.; Denis, C. L.; Parker, R. The Transcription Factor Associated Ccr4 and Caf1 Proteins Are Components of the Major Cytoplasmic MRNA Deadenylase in Saccharomyces Cerevisiae. *Cell* **2001**, *104* (3), 377–386.
(67) Tucker, M.; Staples, R. R.; Valencia-Sanchez, M. A.; Muhlrad, D.; Parker, R. Ccr4p Is the Catalytic Subunit of a Ccr4p/Pop2p/Notp MRNA Deadenylase Complex in Saccharomyces Cerevisiae. *EMBO J.* **2002**, *21* (6), 1427–1436. https://doi.org/10.1093/emboj/21.6.1427.
(68) Song, M.-G.; Kiledjian, M. 3' Terminal Oligo U-Tract-Mediated Stimulation of Decapping. *RNA N. Y. N* **2007**, *13* (12), 2356–2365. https://doi.org/10.1261/rna.765807.
(69) Rissland, O. S.; Norbury, C. J. Decapping Is Preceded by 3' Uridylation in a Novel Pathway of Bulk MRNA Turnover. *Nat. Struct. Mol. Biol.* **2009**, *16* (6), 616–623. https://doi.org/10.1038/nsmb.1601.
(70) Lim, J.; Ha, M.; Chang, H.; Kwon, S. C.; Simanshu, D. K.; Patel, D. J.; Kim, V. N. Uridylation by TUT4 and TUT7 Marks MRNA for Degradation. *Cell* **2014**, *159* (6), 1365–1376. https://doi.org/10.1016/j.cell.2014.10.055.
(71) Sachs, A. B.; Davis, R. W. The Poly(A) Binding Protein Is Required for Poly(A) Shortening and 60S Ribosomal Subunit-Dependent Translation Initiation. *Cell* **1989**, *58* (5), 857–867.
(72) Webster, M. W.; Chen, Y.-H.; Stowell, J. A. W.; Alhusaini, N.; Sweet, T.; Graveley, B. R.; Coller, J.; Passmore, L. A. MRNA Deadenylation Is Coupled to Translation Rates by the Differential Activities of Ccr4-Not Nucleases. *Mol. Cell* **2018**, *70* (6), 1089-1100.e8. https://doi.org/10.1016/j.molcel.2018.05.033.
(73) Presnyak, V.; Alhusaini, N.; Chen, Y.-H.; Martin, S.; Morris, N.; Kline, N.; Olson, S.; Weinberg, D.; Baker, K. E.; Graveley, B. R.; et al. Codon Optimality Is a Major Determinant of MRNA Stability. *Cell* **2015**, *160* (6), 1111–1124. https://doi.org/10.1016/j.cell.2015.02.029.
(74) Radhakrishnan, A.; Chen, Y.-H.; Martin, S.; Alhusaini, N.; Green, R.; Coller, J. The DEAD-Box Protein Dhh1p Couples MRNA Decay and Translation by Monitoring Codon Optimality. *Cell* **2016**, *167* (1), 122-132.e9. https://doi.org/10.1016/j.cell.2016.08.053.
(75) Carroll, J. S.; Munchel, S. E.; Weis, K. The DExD/H Box ATPase Dhh1 Functions in Translational Repression, MRNA Decay, and Processing Body Dynamics. *J. Cell Biol.* **2011**, *194* (4), 527–537. https://doi.org/10.1083/jcb.201007151.
(76) Cheng, J.; Maier, K. C.; Avsec, Ž.; Rus, P.; Gagneur, J. Cis-Regulatory Elements Explain Most of the MRNA Stability Variation across Genes in Yeast. *RNA N. Y. N* **2017**, *23* (11), 1648–1659. https://doi.org/10.1261/rna.062224.117.
(77) Mitchell, P.; Petfalski, E.; Shevchenko, A.; Mann, M.; Tollervey, D. The Exosome: A Conserved Eukaryotic RNA Processing Complex Containing Multiple 3'-->5' Exoribonucleases. *Cell* **1997**, *91* (4), 457–466.
(78) Liu, Q.; Greimann, J. C.; Lima, C. D. Reconstitution, Activities, and Structure of the Eukaryotic RNA Exosome. *Cell* **2006**, *127* (6), 1223–1237. https://doi.org/10.1016/j.cell.2006.10.037.
(79) Dziembowski, A.; Lorentzen, E.; Conti, E.; Séraphin, B. A Single Subunit, Dis3, Is Essentially Responsible for Yeast Exosome Core Activity. *Nat. Struct. Mol. Biol.* **2007**, *14* (1), 15–22. https://doi.org/10.1038/nsmb1184.
(80) Anderson, J. S.; Parker, R. P. The 3' to 5' Degradation of Yeast MRNAs Is a General Mechanism for MRNA Turnover That Requires the SKI2 DEVH Box Protein and 3' to 5' Exonucleases of the Exosome Complex. *EMBO J.* **1998**, *17* (5), 1497–1506. https://doi.org/10.1093/emboj/17.5.1497.
(81) Zinder, J. C.; Lima, C. D. Targeting RNA for Processing or Destruction by the Eukaryotic RNA Exosome and Its Cofactors. *Genes Dev.* **2017**, *31* (2), 88–100. https://doi.org/10.1101/gad.294769.116.
(82) Chen, N.; Walsh, M. A.; Liu, Y.; Parker, R.; Song, H. Crystal Structures of Human DcpS in Ligand-Free and M7GDP-Bound Forms Suggest a Dynamic Mechanism for Scavenger MRNA Decapping. *J. Mol. Biol.* **2005**, *347* (4), 707–718. https://doi.org/10.1016/j.jmb.2005.01.062.
(83) Wang, Z.; Kiledjian, M. Functional Link between the Mammalian Exosome and MRNA Decapping. *Cell* **2001**, *107* (6), 751–762.
(84) Gu, M.; Fabrega, C.; Liu, S.-W.; Liu, H.; Kiledjian, M.; Lima, C. D. Insights into the Structure, Mechanism, and Regulation of Scavenger MRNA Decapping Activity. *Mol. Cell* **2004**, *14* (1), 67–80.

(85) Liu, H.; Rodgers, N. D.; Jiao, X.; Kiledjian, M. The Scavenger MRNA Decapping Enzyme DcpS Is a Member of the HIT Family of Pyrophosphatases. *EMBO J.* **2002**, *21* (17), 4699–4708.
(86) Muhlrad, D.; Decker, C. J.; Parker, R. Deadenylation of the Unstable MRNA Encoded by the Yeast MFA2 Gene Leads to Decapping Followed by 5'-->3' Digestion of the Transcript. *Genes Dev.* **1994**, *8* (7), 855–866.
(87) Schwartz, D. C.; Parker, R. MRNA Decapping in Yeast Requires Dissociation of the Cap Binding Protein, Eukaryotic Translation Initiation Factor 4E. *Mol. Cell. Biol.* **2000**, *20* (21), 7933–7942. https://doi.org/10.1128/MCB.20.21.7933-7942.2000.
(88) Coller, J.; Parker, R. General Translational Repression by Activators of MRNA Decapping. *Cell* **2005**, *122* (6), 875–886. https://doi.org/10.1016/j.cell.2005.07.012.
(89) van Dijk, E.; Cougot, N.; Meyer, S.; Babajko, S.; Wahle, E.; Seraphin, B. Human Dcp2: A Catalytically Active MRNA Decapping Enzyme Located in Specific Cytoplasmic Structures. *EMBO J* **2002**, *21* (24), 6915–6924.
(90) She, M.; Decker, C. J.; Chen, N.; Tumati, S.; Parker, R.; Song, H. Crystal Structure and Functional Analysis of Dcp2p from Schizosaccharomyces Pombe. *Nat. Struct. Mol. Biol.* **2006**, *13* (1), 63–70. https://doi.org/10.1038/nsmb1033.
(91) Stevens, A. Purification and Characterization of a Saccharomyces Cerevisiae Exoribonuclease Which Yields 5'-Mononucleotides by a 5' Leads to 3' Mode of Hydrolysis. *J. Biol. Chem.* **1980**, *255* (7), 3080–3085.
(92) Jones, C. I.; Zabolotskaya, M. V.; Newbury, S. F. The 5' → 3' Exoribonuclease XRN1/Pacman and Its Functions in Cellular Processes and Development. *Wiley Interdiscip. Rev. RNA* **2012**, *3* (4), 455–468. https://doi.org/10.1002/wrna.1109.
(93) Nagarajan, V. K.; Jones, C. I.; Newbury, S. F.; Green, P. J. XRN 5'→3' Exoribonucleases: Structure, Mechanisms and Functions. *Biochim. Biophys. Acta* **2013**, *1829* (6–7), 590–603. https://doi.org/10.1016/j.bbagrm.2013.03.005.
(94) Wypijewska, A.; Bojarska, E.; Lukaszewicz, M.; Stepinski, J.; Jemielity, J.; Davis, R. E.; Darzynkiewicz, E. 7-Methylguanosine Diphosphate (m(7)GDP) Is Not Hydrolyzed but Strongly Bound by Decapping Scavenger (DcpS) Enzymes and Potently Inhibits Their Activity. *Biochemistry* **2012**, *51* (40), 8003–8013. https://doi.org/10.1021/bi300781g.
(95) Nissan, T.; Rajyaguru, P.; She, M.; Song, H.; Parker, R. Decapping Activators in Saccharomyces Cerevisiae Act by Multiple Mechanisms. *Mol. Cell* **2010**, *39* (5), 773–783. https://doi.org/10.1016/j.molcel.2010.08.025.
(96) Wurm, J. P.; Holdermann, I.; Overbeck, J. H.; Mayer, P. H. O.; Sprangers, R. Changes in Conformational Equilibria Regulate the Activity of the Dcp2 Decapping Enzyme. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114* (23), 6034–6039. https://doi.org/10.1073/pnas.1704496114.
(97) She, M.; Decker, C. J.; Sundramurthy, K.; Liu, Y.; Chen, N.; Parker, R.; Song, H. Crystal Structure of Dcp1p and Its Functional Implications in MRNA Decapping. *Nat. Struct. Mol. Biol.* **2004**, *11* (3), 249–256. https://doi.org/10.1038/nsmb730.
(98) She, M.; Decker, C. J.; Svergun, D. I.; Round, A.; Chen, N.; Muhlrad, D.; Parker, R.; Song, H. Structural Basis of Dcp2 Recognition and Activation by Dcp1. *Mol. Cell* **2008**, *29* (3), 337–349. https://doi.org/10.1016/j.molcel.2008.01.002.
(99) Wurm, J. P.; Overbeck, J.; Sprangers, R. The S. Pombe MRNA Decapping Complex Recruits Cofactors and an Edc1-like Activator through a Single Dynamic Surface. *RNA N. Y. N* **2016**, *22* (9), 1360–1372. https://doi.org/10.1261/rna.057315.116.
(100) Borja, M. S.; Piotukh, K.; Freund, C.; Gross, J. D. Dcp1 Links Coactivators of MRNA Decapping to Dcp2 by Proline Recognition. *RNA N. Y. N* **2011**, *17* (2), 278–290. https://doi.org/10.1261/rna.2382011.
(101) Fromm, S. A.; Truffault, V.; Kamenz, J.; Braun, J. E.; Hoffmann, N. A.; Izaurralde, E.; Sprangers, R. The Structural Basis of Edc3- and Scd6-Mediated Activation of the Dcp1:Dcp2 MRNA Decapping Complex. *EMBO J.* **2012**, *31* (2), 279–290. https://doi.org/10.1038/emboj.2011.408.
(102) Kshirsagar, M.; Parker, R. Identification of Edc3p as an Enhancer of MRNA Decapping in Saccharomyces Cerevisiae. *Genetics* **2004**, *166* (2), 729–739.
(103) Fromm, S. A.; Kamenz, J.; Nöldeke, E. R.; Neu, A.; Zocher, G.; Sprangers, R. In Vitro Reconstitution of a Cellular Phase-Transition Process That Involves the MRNA Decapping Machinery. *Angew. Chem. Int. Ed Engl.* **2014**, *53* (28), 7354–7359. https://doi.org/10.1002/anie.201402885.
(104) Harigaya, Y.; Jones, B. N.; Muhlrad, D.; Gross, J. D.; Parker, R. Identification and Analysis of the Interaction between Edc3 and Dcp2 in Saccharomyces Cerevisiae. *Mol Cell Biol* **2010**, *30* (6), 1446–1456. https://doi.org/10.1128/MCB.01305-09.
(105) Sharif, H.; Ozgur, S.; Sharma, K.; Basquin, C.; Urlaub, H.; Conti, E. Structural Analysis of the Yeast Dhh1-Pat1 Complex Reveals How Dhh1 Engages Pat1, Edc3 and RNA in Mutually Exclusive Interactions. *Nucleic Acids Res.* **2013**, *41* (17), 8377–8390. https://doi.org/10.1093/nar/gkt600.
(106) Tritschler, F.; Braun, J. E.; Eulalio, A.; Truffault, V.; Izaurralde, E.; Weichenrieder, O. Structural Basis for the Mutually Exclusive Anchoring of P Body Components EDC3 and Tral to the DEAD Box Protein DDX6/Me31B. *Mol Cell* **2009**, *33* (5), 661–668. https://doi.org/10.1016/j.molcel.2009.02.014.
(107) Tritschler, F.; Eulalio, A.; Helms, S.; Schmidt, S.; Coles, M.; Weichenrieder, O.; Izaurralde, E.; Truffault, V. Similar Modes of Interaction Enable Trailer Hitch and EDC3 to Associate with DCP1 and Me31B in Distinct Protein Complexes. *Mol. Cell. Biol.* **2008**, *28* (21), 6695–6708. https://doi.org/10.1128/MCB.00759-08.
(108) Brandmann, T.; Fakim, H.; Padamsi, Z.; Youn, J.-Y.; Gingras, A.-C.; Fabian, M. R.; Jinek, M. Molecular Architecture of LSM14 Interactions Involved in the Assembly of MRNA Silencing Complexes. *EMBO J.* **2018**, *37* (7). https://doi.org/10.15252/embj.201797869.
(109) Wu, D.; Muhlrad, D.; Bowler, M. W.; Jiang, S.; Liu, Z.; Parker, R.; Song, H. Lsm2 and Lsm3 Bridge the Interaction of the Lsm1-7 Complex with Pat1 for Decapping Activation. *Cell Res.* **2014**, *24* (2), 233–246. https://doi.org/10.1038/cr.2013.152.

(110) Chowdhury, A.; Mukhopadhyay, J.; Tharun, S. The Decapping Activator Lsm1p-7p-Pat1p Complex Has the Intrinsic Ability to Distinguish between Oligoadenylated and Polyadenylated RNAs. *RNA N. Y. N* **2007**, *13* (7), 998–1016. https://doi.org/10.1261/rna.502507.
(111) Tharun, S. Lsm1-7-Pat1 Complex: A Link between 3' and 5'-Ends in MRNA Decay? *RNA Biol.* **2009**, *6* (3), 228–232. https://doi.org/10.4161/rna.6.3.8282.
(112) Chang, C.-T.; Bercovich, N.; Loh, B.; Jonas, S.; Izaurralde, E. The Activation of the Decapping Enzyme DCP2 by DCP1 Occurs on the EDC4 Scaffold and Involves a Conserved Loop in DCP1. *Nucleic Acids Res.* **2014**, *42* (8), 5217–5233. https://doi.org/10.1093/nar/gku129.
(113) Sheth, U.; Parker, R. Decapping and Decay of Messenger RNA Occur in Cytoplasmic Processing Bodies. *Science* **2003**, *300* (5620), 805–808. https://doi.org/10.1126/science.1082320.
(114) Parker, R.; Sheth, U. P Bodies and the Control of MRNA Translation and Degradation. *Mol Cell* **2007**, *25* (5), 635–646. https://doi.org/10.1016/j.molcel.2007.02.011.
(115) Hyman, A. A.; Weber, C. A.; Julicher, F. Liquid-Liquid Phase Separation in Biology. *Annu Rev Cell Dev Biol* **2014**, *30*, 39–58. https://doi.org/10.1146/annurev-cellbio-100913-013325.
(116) Kroschwald, S.; Maharana, S.; Mateju, D.; Malinovska, L.; Nuske, E.; Poser, I.; Richter, D.; Alberti, S. Promiscuous Interactions and Protein Disaggregases Determine the Material State of Stress-Inducible RNP Granules. *Elife* **2015**, *4*, e06807. https://doi.org/10.7554/eLife.06807.
(117) Brangwynne, C. P.; Eckmann, C. R.; Courson, D. S.; Rybarska, A.; Hoege, C.; Gharakhani, J.; Julicher, F.; Hyman, A. A. Germline P Granules Are Liquid Droplets That Localize by Controlled Dissolution/Condensation. *Science* **2009**, *324* (5935), 1729–1732. https://doi.org/10.1126/science.1172046.
(118) Banani, S. F.; Lee, H. O.; Hyman, A. A.; Rosen, M. K. Biomolecular Condensates: Organizers of Cellular Biochemistry. *Nat. Rev. Mol. Cell Biol.* **2017**, *18* (5), 285–298. https://doi.org/10.1038/nrm.2017.7.
(119) Brangwynne, C. P.; Tompa, P.; Pappu, R. V. Polymer Physics of Intracellular Phase Transitions. *Nat. Phys.* **2015**, *11* (11), 899–904. https://doi.org/10.1038/nphys3532.
(120) Nott, T. J.; Petsalaki, E.; Farber, P.; Jervis, D.; Fussner, E.; Plochowietz, A.; Craggs, T. D.; Bazett-Jones, D. P.; Pawson, T.; Forman-Kay, J. D.; et al. Phase Transition of a Disordered Nuage Protein Generates Environmentally Responsive Membraneless Organelles. *Mol Cell* **2015**, *57* (5), 936–947. https://doi.org/10.1016/j.molcel.2015.01.013.
(121) Patel, A.; Lee, H. O.; Jawerth, L.; Maharana, S.; Jahnel, M.; Hein, M. Y.; Stoynov, S.; Mahamid, J.; Saha, S.; Franzmann, T. M.; et al. A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. *Cell* **2015**, *162* (5), 1066–1077. https://doi.org/10.1016/j.cell.2015.07.047.
(122) Zhang, H.; Elbaum-Garfinkle, S.; Langdon, E. M.; Taylor, N.; Occhipinti, P.; Bridges, A. A.; Brangwynne, C. P.; Gladfelter, A. S. RNA Controls PolyQ Protein Phase Transitions. *Mol Cell* **2015**, *60* (2), 220–230. https://doi.org/10.1016/j.molcel.2015.09.017.
(123) Molliex, A.; Temirov, J.; Lee, J.; Coughlin, M.; Kanagaraj, A. P.; Kim, H. J.; Mittag, T.; Taylor, J. P. Phase Separation by Low Complexity Domains Promotes Stress Granule Assembly and Drives Pathological Fibrillization. *Cell* **2015**, *163* (1), 123–133. https://doi.org/10.1016/j.cell.2015.09.015.
(124) Elbaum-Garfinkle, S.; Kim, Y.; Szczepaniak, K.; Chen, C. C.; Eckmann, C. R.; Myong, S.; Brangwynne, C. P. The Disordered P Granule Protein LAF-1 Drives Phase Separation into Droplets with Tunable Viscosity and Dynamics. *Proc Natl Acad Sci U A* **2015**, *112* (23), 7189–7194. https://doi.org/10.1073/pnas.1504822112.
(125) Mitrea, D. M.; Kriwacki, R. W. Phase Separation in Biology; Functional Organization of a Higher Order. *Cell Commun Signal* **2016**, *14* (1), 1. https://doi.org/10.1186/s12964-015-0125-7.
(126) Brangwynne, C. P.; Mitchison, T. J.; Hyman, A. A. Active Liquid-like Behavior of Nucleoli Determines Their Size and Shape in Xenopus Laevis Oocytes. *Proc Natl Acad Sci U A* **2011**, *108* (11), 4334–4339. https://doi.org/10.1073/pnas.1017150108.
(127) Nizami, Z.; Deryusheva, S.; Gall, J. G. The Cajal Body and Histone Locus Body. *Cold Spring Harb. Perspect. Biol.* **2010**, *2* (7), a000653. https://doi.org/10.1101/cshperspect.a000653.
(128) Fox, A. H.; Lamond, A. I. Paraspeckles. *Cold Spring Harb. Perspect. Biol.* **2010**, *2* (7), a000687. https://doi.org/10.1101/cshperspect.a000687.
(129) Duronio, R. J.; Marzluff, W. F. Coordinating Cell Cycle-Regulated Histone Gene Expression through Assembly and Function of the Histone Locus Body. *RNA Biol.* **2017**, *14* (6), 726–738. https://doi.org/10.1080/15476286.2016.1265198.
(130) Eskiw, C. H.; Dellaire, G.; Mymryk, J. S.; Bazett-Jones, D. P. Size, Position and Dynamic Behavior of PML Nuclear Bodies Following Cell Stress as a Paradigm for Supramolecular Trafficking and Assembly. *J. Cell Sci.* **2003**, *116* (Pt 21), 4455–4466. https://doi.org/10.1242/jcs.00758.
(131) Frey, S.; Richter, R. P.; Gorlich, D. FG-Rich Repeats of Nuclear Pore Proteins Form a Three-Dimensional Meshwork with Hydrogel-like Properties. *Science* **2006**, *314* (5800), 815–817. https://doi.org/10.1126/science.1132516.
(132) Kedersha, N.; Stoecklin, G.; Ayodele, M.; Yacono, P.; Lykke-Andersen, J.; Fritzler, M. J.; Scheuner, D.; Kaufman, R. J.; Golan, D. E.; Anderson, P. Stress Granules and Processing Bodies Are Dynamically Linked Sites of MRNP Remodeling. *J Cell Biol* **2005**, *169* (6), 871–884. https://doi.org/10.1083/jcb.200502088.
(133) Boke, E.; Ruer, M.; Wühr, M.; Coughlin, M.; Lemaitre, R.; Gygi, S. P.; Alberti, S.; Drechsel, D.; Hyman, A. A.; Mitchison, T. J. Amyloid-like Self-Assembly of a Cellular Compartment. *Cell* **2016**, *166* (3), 637–650. https://doi.org/10.1016/j.cell.2016.06.051.
(134) Li, P.; Banjade, S.; Cheng, H. C.; Kim, S.; Chen, B.; Guo, L.; Llaguno, M.; Hollingsworth, J. V.; King, D. S.; Banani, S. F.; et al. Phase Transitions in the Assembly of Multivalent Signalling Proteins. *Nature* **2012**, *483* (7389), 336–340. https://doi.org/10.1038/nature10879.

(135) Schwarz-Romond, T.; Fiedler, M.; Shibata, N.; Butler, P. J. G.; Kikuchi, A.; Higuchi, Y.; Bienz, M. The DIX Domain of Dishevelled Confers Wnt Signaling by Dynamic Polymerization. *Nat. Struct. Mol. Biol.* **2007**, *14* (6), 484–492. https://doi.org/10.1038/nsmb1247.
(136) Banjade, S.; Rosen, M. K. Phase Transitions of Multivalent Proteins Can Promote Clustering of Membrane Receptors. *eLife* **2014**, *3*. https://doi.org/10.7554/eLife.04123.
(137) An, S.; Kumar, R.; Sheets, E. D.; Benkovic, S. J. Reversible Compartmentalization of de Novo Purine Biosynthetic Complexes in Living Cells. *Science* **2008**, *320* (5872), 103–106. https://doi.org/10.1126/science.1152241.
(138) Saito, M.; Hess, D.; Eglinger, J.; Fritsch, A. W.; Kreysing, M.; Weinert, B. T.; Choudhary, C.; Matthias, P. Acetylation of Intrinsically Disordered Regions Regulates Phase Separation. *Nat. Chem. Biol.* **2019**, *15* (1), 51–61. https://doi.org/10.1038/s41589-018-0180-7.
(139) Wippich, F.; Bodenmiller, B.; Trajkovska, M. G.; Wanka, S.; Aebersold, R.; Pelkmans, L. Dual Specificity Kinase DYRK3 Couples Stress Granule Condensation/Dissolution to MTORC1 Signaling. *Cell* **2013**, *152* (4), 791–805. https://doi.org/10.1016/j.cell.2013.01.033.
(140) Monahan, Z.; Ryan, V. H.; Janke, A. M.; Burke, K. A.; Rhoads, S. N.; Zerze, G. H.; O'Meally, R.; Dignon, G. L.; Conicella, A. E.; Zheng, W.; et al. Phosphorylation of the FUS Low-Complexity Domain Disrupts Phase Separation, Aggregation, and Toxicity. *EMBO J.* **2017**, *36* (20), 2951–2967. https://doi.org/10.15252/embj.201696394.
(141) Aumiller, W. M.; Keating, C. D. Phosphorylation-Mediated RNA/Peptide Complex Coacervation as a Model for Intracellular Liquid Organelles. *Nat Chem* **2016**, *8* (2), 129–137. https://doi.org/10.1038/nchem.2414.
(142) Falahati, H.; Wieschaus, E. Independent Active and Thermodynamic Processes Govern the Nucleolus Assembly in Vivo. *Proc. Natl. Acad. Sci.* **2017**, *114* (6), 1335–1340. https://doi.org/10.1073/pnas.1615395114.
(143) Banjade, S.; Wu, Q.; Mittal, A.; Peeples, W. B.; Pappu, R. V.; Rosen, M. K. Conserved Interdomain Linker Promotes Phase Separation of the Multivalent Adaptor Protein Nck. *Proc Natl Acad Sci U A* **2015**, *112* (47), E6426-35. https://doi.org/10.1073/pnas.1508778112.
(144) Decker, C. J.; Teixeira, D.; Parker, R. Edc3p and a Glutamine/Asparagine-Rich Domain of Lsm4p Function in Processing Body Assembly in Saccharomyces Cerevisiae. *J Cell Biol* **2007**, *179* (3), 437–449. https://doi.org/10.1083/jcb.200704147.
(145) Peskett, T. R.; Rau, F.; O'Driscoll, J.; Patani, R.; Lowe, A. R.; Saibil, H. R. A Liquid to Solid Phase Transition Underlying Pathological Huntingtin Exon1 Aggregation. *Mol. Cell* **2018**, *70* (4), 588-601.e6. https://doi.org/10.1016/j.molcel.2018.04.007.
(146) Reijns, M. A. M.; Alexander, R. D.; Spiller, M. P.; Beggs, J. D. A Role for Q/N-Rich Aggregation-Prone Regions in P-Body Localization. *J. Cell Sci.* **2008**, *121* (Pt 15), 2463–2472. https://doi.org/10.1242/jcs.024976.
(147) Hughes, M. P.; Sawaya, M. R.; Boyer, D. R.; Goldschmidt, L.; Rodriguez, J. A.; Cascio, D.; Chong, L.; Gonen, T.; Eisenberg, D. S. Atomic Structures of Low-Complexity Protein Segments Reveal Kinked β Sheets That Assemble Networks. *Science* **2018**, *359* (6376), 698–701. https://doi.org/10.1126/science.aan6398.
(148) Murray, D. T.; Kato, M.; Lin, Y.; Thurber, K. R.; Hung, I.; McKnight, S. L.; Tycko, R. Structure of FUS Protein Fibrils and Its Relevance to Self-Assembly and Phase Separation of Low-Complexity Domains. *Cell* **2017**, *171* (3), 615-627.e16. https://doi.org/10.1016/j.cell.2017.08.048.
(149) Ryan, V. H.; Dignon, G. L.; Zerze, G. H.; Chabata, C. V.; Silva, R.; Conicella, A. E.; Amaya, J.; Burke, K. A.; Mittal, J.; Fawzi, N. L. Mechanistic View of HnRNPA2 Low-Complexity Domain Structure, Interactions, and Phase Separation Altered by Mutation and Arginine Methylation. *Mol. Cell* **2018**, *69* (3), 465-479.e7. https://doi.org/10.1016/j.molcel.2017.12.022.
(150) Banani, S. F.; Rice, A. M.; Peeples, W. B.; Lin, Y.; Jain, S.; Parker, R.; Rosen, M. K. Compositional Control of Phase-Separated Cellular Bodies. *Cell* **2016**. https://doi.org/10.1016/j.cell.2016.06.010.
(151) Lin, Y.; Protter, D. S.; Rosen, M. K.; Parker, R. Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins. *Mol Cell* **2015**, *60* (2), 208–219. https://doi.org/10.1016/j.molcel.2015.08.018.
(152) Sachdev, R.; Hondele, M.; Linsenmeier, M.; Vallotton, P.; Mugler, C. F.; Arosio, P.; Weis, K. Pat1 Promotes Processing Body Assembly by Enhancing the Phase Separation of the DEAD-Box ATPase Dhh1 and RNA. *eLife* **2019**, *8*. https://doi.org/10.7554/eLife.41415.
(153) Ramachandran, V.; Shah, K. H.; Herman, P. K. The CAMP-Dependent Protein Kinase Signaling Pathway Is a Key Regulator of P Body Foci Formation. *Mol. Cell* **2011**, *43* (6), 973–981. https://doi.org/10.1016/j.molcel.2011.06.032.
(154) Shin, Y.; Berry, J.; Pannucci, N.; Haataja, M. P.; Toettcher, J. E.; Brangwynne, C. P. Spatiotemporal Control of Intracellular Phase Transitions Using Light-Activated OptoDroplets. *Cell* **2017**, *168* (1–2), 159-171.e14. https://doi.org/10.1016/j.cell.2016.11.054.
(155) Kato, M.; Han, T. W.; Xie, S.; Shi, K.; Du, X.; Wu, L. C.; Mirzaei, H.; Goldsmith, E. J.; Longgood, J.; Pei, J.; et al. Cell-Free Formation of RNA Granules: Low Complexity Sequence Domains Form Dynamic Fibers within Hydrogels. *Cell* **2012**, *149* (4), 753–767. https://doi.org/10.1016/j.cell.2012.04.017.
(156) Murakami, T.; Qamar, S.; Lin, J. Q.; Schierle, G. S. K.; Rees, E.; Miyashita, A.; Costa, A. R.; Dodd, R. B.; Chan, F. T. S.; Michel, C. H.; et al. ALS/FTD Mutation-Induced Phase Transition of FUS Liquid Droplets and Reversible Hydrogels into Irreversible Hydrogels Impairs RNP Granule Function. *Neuron* **2015**, *88* (4), 678–690. https://doi.org/10.1016/j.neuron.2015.10.030.
(157) Kroschwald, S.; Maharana, S.; Simon, A. Hexanediol: A Chemical Probe to Investigate the Material Properties of Membrane-Less Compartments. *Matters* **2017**. https://doi.org/10.19185/matters.201702000010.
(158) Ader, C.; Frey, S.; Maas, W.; Schmidt, H. B.; Görlich, D.; Baldus, M. Amyloid-like Interactions within Nucleoporin FG Hydrogels. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107* (14), 6281–6285. https://doi.org/10.1073/pnas.0910163107.

(159) Maharana, S.; Wang, J.; Papadopoulos, D. K.; Richter, D.; Pozniakovsky, A.; Poser, I.; Bickle, M.; Rizk, S.; Guillén-Boixet, J.; Franzmann, T. M.; et al. RNA Buffers the Phase Separation Behavior of Prion-like RNA Binding Proteins. *Science* **2018**, *360* (6391), 918–921. https://doi.org/10.1126/science.aar7366.
(160) Patel, A.; Malinovska, L.; Saha, S.; Wang, J.; Alberti, S.; Krishnan, Y.; Hyman, A. A. ATP as a Biological Hydrotrope. *Science* **2017**, *356* (6339), 753–756. https://doi.org/10.1126/science.aaf6846.
(161) Kang, J.; Lim, L.; Song, J. ATP Enhances at Low Concentrations but Dissolves at High Concentrations Liquid-Liquid Phase Separation (LLPS) of ALS/FTD-Causing FUS. *Biochem. Biophys. Res. Commun.* **2018**, *504* (2), 545–551. https://doi.org/10.1016/j.bbrc.2018.09.014.
(162) Sridharan, S.; Kurzawa, N.; Werner, T.; Günthner, I.; Helm, D.; Huber, W.; Bantscheff, M.; Savitski, M. M. Proteome-Wide Solubility and Thermal Stability Profiling Reveals Distinct Regulatory Roles for ATP. *Nat. Commun.* **2019**, *10* (1), 1155. https://doi.org/10.1038/s41467-019-09107-y.
(163) Buchan, J. R.; Kolaitis, R. M.; Taylor, J. P.; Parker, R. Eukaryotic Stress Granules Are Cleared by Autophagy and Cdc48/VCP Function. *Cell* **2013**, *153* (7), 1461–1474. https://doi.org/10.1016/j.cell.2013.05.037.
(164) Mugler, C. F.; Hondele, M.; Heinrich, S.; Sachdev, R.; Vallotton, P.; Koek, A. Y.; Chan, L. Y.; Weis, K. ATPase Activity of the DEAD-Box Protein Dhh1 Controls Processing Body Formation. *eLife* **2016**, *5*. https://doi.org/10.7554/eLife.18746.
(165) Gardner, K. H.; Kay, L. E. Production and Incorporation of ^{15}N, ^{13}C, ^{2}H (^{1}H-Δ1 Methyl) Isoleucine into Proteins for Multidimensional NMR Studies. *J. Am. Chem. Soc.* **1997**, *119* (32), 7599–7600. https://doi.org/10.1021/ja9706514.
(166) LeMaster, D. M. Chiral Beta and Random Fractional Deuteration for the Determination of Protein Sidechain Conformation by NMR. *FEBS Lett.* **1987**, *223* (1), 191–196.
(167) Markus, M. A.; Dayie, K. T.; Matsudaira, P.; Wagner, G. Effect of Deuteration on the Amide Proton Relaxation Rates in Proteins. Heteronuclear NMR Experiments on Villin 14T. *J. Magn. Reson. B* **1994**, *105* (2), 192–195.
(168) Venters, R. A.; Farmer, B. T.; Fierke, C. A.; Spicer, L. D. Characterizing the Use of Perdeuteration in NMR Studies of Large Proteins: 13C, 15N and 1H Assignments of Human Carbonic Anhydrase II. *J. Mol. Biol.* **1996**, *264* (5), 1101–1116.
(169) Sattler, M.; Fesik, S. W. Use of Deuterium Labeling in NMR: Overcoming a Sizeable Problem. *Struct. Lond. Engl. 1993* **1996**, *4* (11), 1245–1249.
(170) Venters, R. A.; Huang, C. C.; Farmer, B. T.; Trolard, R.; Spicer, L. D.; Fierke, C. A. High-Level 2H/13C/15N Labeling of Proteins for NMR Studies. *J. Biomol. NMR* **1995**, *5* (4), 339–344.
(171) Grzesiek, S.; Anglister, J.; Ren, H.; Bax, A. Carbon-13 Line Narrowing by Deuterium Decoupling in Deuterium/Carbon-13/Nitrogen-15 Enriched Proteins. Application to Triple Resonance 4D J Connectivity of Sequential Amides. *J. Am. Chem. Soc.* **1993**, *115* (10), 4369–4370. https://doi.org/10.1021/ja00063a068.
(172) Yamazaki, T.; Lee, W.; Arrowsmith, C. H.; Muhandiram, D. R.; Kay, L. E. A Suite of Triple Resonance NMR Experiments for the Backbone Assignment of 15N, 13C, 2H Labeled Proteins with High Sensitivity. *J. Am. Chem. Soc.* **1994**, *116* (26), 11655–11666. https://doi.org/10.1021/ja00105a005.
(173) Yamazaki, T.; Lee, W.; Revington, M.; Mattiello, D. L.; Dahlquist, F. W.; Arrowsmith, C. H.; Kay, L. E. An HNCA Pulse Scheme for the Backbone Assignment of 15N,13C,2H-Labeled Proteins: Application to a 37-KDa Trp Repressor-DNA Complex. *J. Am. Chem. Soc.* **1994**, *116* (14), 6464–6465. https://doi.org/10.1021/ja00093a069.
(174) Griffey, R. H.; Redfield, A. G. Proton-Detected Heteronuclear Edited and Correlated Nuclear Magnetic Resonance and Nuclear Overhauser Effect in Solution. *Q. Rev. Biophys.* **1987**, *19* (1–2), 51–82.
(175) Gueron, M.; Leroy, J. L.; Griffey, R. H. Proton Nuclear Magnetic Relaxation of Nitrogen-15-Labeled Nucleic Acids via Dipolar Coupling and Chemical Shift Anisotropy. *J. Am. Chem. Soc.* **1983**, *105* (25), 7262–7266. https://doi.org/10.1021/ja00363a009.
(176) Pervushin, K.; Riek, R.; Wider, G.; Wüthrich, K. Attenuated T2 Relaxation by Mutual Cancellation of Dipole-Dipole Coupling and Chemical Shift Anisotropy Indicates an Avenue to NMR Structures of Very Large Biological Macromolecules in Solution. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94* (23), 12366–12371.
(177) Riek, R.; Wider, G.; Pervushin, K.; Wüthrich, K. Polarization Transfer by Cross-Correlated Relaxation in Solution NMR with Very Large Molecules. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96* (9), 4918–4923.
(178) Wider, G.; Wüthrich, K. NMR Spectroscopy of Large Molecules and Multimolecular Assemblies in Solution. *Curr. Opin. Struct. Biol.* **1999**, *9* (5), 594–601.
(179) Riek, R.; Pervushin, K.; Wüthrich, K. TROSY and CRINEPT: NMR with Large Molecular and Supramolecular Structures in Solution. *Trends Biochem. Sci.* **2000**, *25* (10), 462–468.
(180) Pervushin, K.; Riek, R.; Wider, G.; Wüthrich, K. Transverse Relaxation-Optimized Spectroscopy (TROSY) for NMR Studies of Aromatic Spin Systems in ^{13}C-Labeled Proteins. *J. Am. Chem. Soc.* **1998**, *120* (25), 6394–6400. https://doi.org/10.1021/ja980742g.
(181) Tugarinov, V.; Hwang, P. M.; Ollerenshaw, J. E.; Kay, L. E. Cross-Correlated Relaxation Enhanced 1H-13C NMR Spectroscopy of Methyl Groups in Very High Molecular Weight Proteins and Protein Complexes. *J. Am. Chem. Soc.* **2003**, *125* (34), 10420–10428. https://doi.org/10.1021/ja030153x.
(182) Bodenhausen, G.; Ruben, D. J. Natural Abundance Nitrogen-15 NMR by Enhanced Heteronuclear Spectroscopy. *Chem. Phys. Lett.* **1980**, *69* (1), 185–189. https://doi.org/10.1016/0009-2614(80)80041-8.
(183) Takeuchi, K.; Arthanari, H.; Wagner, G. Perspective: Revisiting the Field Dependence of TROSY Sensitivity. *J. Biomol. NMR* **2016**, *66* (4), 221–225. https://doi.org/10.1007/s10858-016-0075-4.
(184) Sprangers, R.; Kay, L. E. Quantitative Dynamics and Binding Studies of the 20S Proteasome by NMR. *Nature* **2007**, *445* (7128), 618–622. https://doi.org/10.1038/nature05512.

(185) Religa, T. L.; Ruschak, A. M.; Rosenzweig, R.; Kay, L. E. Site-Directed Methyl Group Labeling as an NMR Probe of Structure and Dynamics in Supramolecular Protein Systems: Applications to the Proteasome and to the ClpP Protease. *J. Am. Chem. Soc.* **2011**, *133* (23), 9063–9068. https://doi.org/10.1021/ja202259a.
(186) Theillet, F.-X.; Liokatis, S.; Jost, J. O.; Bekei, B.; Rose, H. M.; Binolfi, A.; Schwarzer, D.; Selenko, P. Site-Specific Mapping and Time-Resolved Monitoring of Lysine Methylation by High-Resolution NMR Spectroscopy. *J. Am. Chem. Soc.* **2012**, *134* (18), 7616–7619. https://doi.org/10.1021/ja301895f.
(187) Kerfah, R.; Plevin, M. J.; Sounier, R.; Gans, P.; Boisbouvier, J. Methyl-Specific Isotopic Labeling: A Molecular Tool Box for Solution NMR Studies of Large Proteins. *Curr. Opin. Struct. Biol.* **2015**, *32*, 113–122. https://doi.org/10.1016/j.sbi.2015.03.009.
(188) Monneau, Y. R.; Ishida, Y.; Rossi, P.; Saio, T.; Tzeng, S.-R.; Inouye, M.; Kalodimos, C. G. Exploiting E. Coli Auxotrophs for Leucine, Valine, and Threonine Specific Methyl Labeling of Large Proteins for NMR Applications. *J. Biomol. NMR* **2016**, *65* (2), 99–108. https://doi.org/10.1007/s10858-016-0041-1.
(189) Miyanoiri, Y.; Ishida, Y.; Takeda, M.; Terauchi, T.; Inouye, M.; Kainosho, M. Highly Efficient Residue-Selective Labeling with Isotope-Labeled Ile, Leu, and Val Using a New Auxotrophic E. Coli Strain. *J. Biomol. NMR* **2016**, *65* (2), 109–119. https://doi.org/10.1007/s10858-016-0042-0.
(190) Isaacson, R. L.; Simpson, P. J.; Liu, M.; Cota, E.; Zhang, X.; Freemont, P.; Matthews, S. A New Labeling Method for Methyl Transverse Relaxation-Optimized Spectroscopy NMR Spectra of Alanine Residues. *J. Am. Chem. Soc.* **2007**, *129* (50), 15428–15429. https://doi.org/10.1021/ja0761784.
(191) Ayala, I.; Sounier, R.; Usé, N.; Gans, P.; Boisbouvier, J. An Efficient Protocol for the Complete Incorporation of Methyl-Protonated Alanine in Perdeuterated Protein. *J. Biomol. NMR* **2009**, *43* (2), 111–119. https://doi.org/10.1007/s10858-008-9294-7.
(192) Ruschak, A. M.; Velyvis, A.; Kay, L. E. A Simple Strategy for ^{13}C, ^{1}H Labeling at the Ile-Γ2 Methyl Position in Highly Deuterated Proteins. *J. Biomol. NMR* **2010**, *48* (3), 129–135. https://doi.org/10.1007/s10858-010-9449-1.
(193) Ayala, I.; Hamelin, O.; Amero, C.; Pessey, O.; Plevin, M. J.; Gans, P.; Boisbouvier, J. An Optimized Isotopic Labelling Strategy of Isoleucine-Γ2 Methyl Groups for Solution NMR Studies of High Molecular Weight Proteins. *Chem. Commun. Camb. Engl.* **2012**, *48* (10), 1434–1436. https://doi.org/10.1039/c1cc12932e.
(194) Hajduk, P. J.; Augeri, D. J.; Mack, J.; Mendoza, R.; Yang, J.; Betz, S. F.; Fesik, S. W. NMR-Based Screening of Proteins Containing ^{13} C-Labeled Methyl Groups. *J. Am. Chem. Soc.* **2000**, *122* (33), 7898–7904. https://doi.org/10.1021/ja000350l.
(195) Tugarinov, V.; Kanelis, V.; Kay, L. E. Isotope Labeling Strategies for the Study of High-Molecular-Weight Proteins by Solution NMR Spectroscopy. *Nat. Protoc.* **2006**, *1* (2), 749–754. https://doi.org/10.1038/nprot.2006.101.
(196) Kerfah, R.; Plevin, M. J.; Pessey, O.; Hamelin, O.; Gans, P.; Boisbouvier, J. Scrambling Free Combinatorial Labeling of Alanine-β, Isoleucine-Δ1, Leucine-ProS and Valine-ProS Methyl Groups for the Detection of Long Range NOEs. *J. Biomol. NMR* **2015**, *61* (1), 73–82. https://doi.org/10.1007/s10858-014-9887-2.
(197) Fischer, M.; Kloiber, K.; Häusler, J.; Ledolter, K.; Konrat, R.; Schmid, W. Synthesis of a 13C-Methyl-Group-Labeled Methionine Precursor as a Useful Tool for Simplifying Protein Structural Analysis by NMR Spectroscopy. *Chembiochem Eur. J. Chem. Biol.* **2007**, *8* (6), 610–612. https://doi.org/10.1002/cbic.200600551.
(198) DellaVecchia, M. J.; Merritt, W. K.; Peng, Y.; Kirby, T. W.; DeRose, E. F.; Mueller, G. A.; Van Houten, B.; London, R. E. NMR Analysis of [Methyl-13C]Methionine UvrB from Bacillus Caldotenax Reveals UvrB-Domain 4 Heterodimer Formation in Solution. *J. Mol. Biol.* **2007**, *373* (2), 282–295. https://doi.org/10.1016/j.jmb.2007.07.045.
(199) Zheng, X.; Mueller, G. A.; DeRose, E. F.; London, R. E. Solution Characterization of [Methyl-(13)C]Methionine HIV-1 Reverse Transcriptase by NMR Spectroscopy. *Antiviral Res.* **2009**, *84* (3), 205–214. https://doi.org/10.1016/j.antiviral.2009.07.021.
(200) Gelis, I.; Bonvin, A. M. J. J.; Keramisanou, D.; Koukaki, M.; Gouridis, G.; Karamanou, S.; Economou, A.; Kalodimos, C. G. Structural Basis for Signal-Sequence Recognition by the Translocase Motor SecA as Determined by NMR. *Cell* **2007**, *131* (4), 756–769. https://doi.org/10.1016/j.cell.2007.09.039.
(201) Sinha, K.; Jen-Jacobson, L.; Rule, G. S. Specific Labeling of Threonine Methyl Groups for NMR Studies of Protein-Nucleic Acid Complexes. *Biochemistry* **2011**, *50* (47), 10189–10191. https://doi.org/10.1021/bi201496d.
(202) Velyvis, A.; Ruschak, A. M.; Kay, L. E. An Economical Method for Production of (2)H, (13)CH3-Threonine for Solution NMR Studies of Large Protein Complexes: Application to the 670 KDa Proteasome. *PloS One* **2012**, *7* (9), e43725. https://doi.org/10.1371/journal.pone.0043725.
(203) Saio, T.; Guan, X.; Rossi, P.; Economou, A.; Kalodimos, C. G. Structural Basis for Protein Antiaggregation Activity of the Trigger Factor Chaperone. *Science* **2014**, *344* (6184), 1250494. https://doi.org/10.1126/science.1250494.
(204) Goto, N. K.; Gardner, K. H.; Mueller, G. A.; Willis, R. C.; Kay, L. E. A Robust and Cost-Effective Method for the Production of Val, Leu, Ile (Delta 1) Methyl-Protonated 15N-, 13C-, 2H-Labeled Proteins. *J. Biomol. NMR* **1999**, *13* (4), 369–374.
(205) Gross, J. D.; Gelev, V. M.; Wagner, G. A Sensitive and Robust Method for Obtaining Intermolecular NOEs between Side Chains in Large Protein Complexes. *J. Biomol. NMR* **2003**, *25* (3), 235–242.
(206) Tugarinov, V.; Kay, L. E. Ile, Leu, and Val Methyl Assignments of the 723-Residue Malate Synthase G Using a New Labeling Strategy and Novel NMR Methods. *J. Am. Chem. Soc.* **2003**, *125* (45), 13868–13878. https://doi.org/10.1021/ja030345s.
(207) Tugarinov, V.; Kay, L. E. An Isotope Labeling Strategy for Methyl TROSY Spectroscopy. *J. Biomol. NMR* **2004**, *28* (2), 165–172. https://doi.org/10.1023/B:JNMR.0000013824.93994.1f.

(208) Lichtenecker, R. J.; Coudevylle, N.; Konrat, R.; Schmid, W. Selective Isotope Labelling of Leucine Residues by Using α-Ketoacid Precursor Compounds. *Chembiochem Eur. J. Chem. Biol.* **2013**, *14* (7), 818–821. https://doi.org/10.1002/cbic.201200737.
(209) Lichtenecker, R. J.; Weinhäupl, K.; Reuther, L.; Schörghuber, J.; Schmid, W.; Konrat, R. Independent Valine and Leucine Isotope Labeling in Escherichia Coli Protein Overexpression Systems. *J. Biomol. NMR* **2013**, *57* (3), 205–209. https://doi.org/10.1007/s10858-013-9786-y.
(210) Gans, P.; Hamelin, O.; Sounier, R.; Ayala, I.; Durá, M. A.; Amero, C. D.; Noirclerc-Savoye, M.; Franzetti, B.; Plevin, M. J.; Boisbouvier, J. Stereospecific Isotopic Labeling of Methyl Groups for NMR Spectroscopic Studies of High-Molecular-Weight Proteins. *Angew. Chem. Int. Ed Engl.* **2010**, *49* (11), 1958–1962. https://doi.org/10.1002/anie.200905660.
(211) Miyanoiri, Y.; Takeda, M.; Okuma, K.; Ono, A. M.; Terauchi, T.; Kainosho, M. Differential Isotope-Labeling for Leu and Val Residues in a Protein by E. Coli Cellular Expression Using Stereo-Specifically Methyl Labeled Amino Acids. *J. Biomol. NMR* **2013**, *57* (3), 237–249. https://doi.org/10.1007/s10858-013-9784-0.
(212) Godoy-Ruiz, R.; Guo, C.; Tugarinov, V. Alanine Methyl Groups as NMR Probes of Molecular Structure and Dynamics in High-Molecular-Weight Proteins. *J. Am. Chem. Soc.* **2010**, *132* (51), 18340–18350. https://doi.org/10.1021/ja1083656.
(213) Huang, C.; Rossi, P.; Saio, T.; Kalodimos, C. G. Structural Basis for the Antifolding Activity of a Molecular Chaperone. *Nature* **2016**, *537* (7619), 202–206. https://doi.org/10.1038/nature18965.
(214) Cvetkovic, M. A.; Wurm, J. P.; Audin, M. J.; Schütz, S.; Sprangers, R. The Rrp4-Exosome Complex Recruits and Channels Substrate RNA by a Unique Mechanism. *Nat. Chem. Biol.* **2017**, *13* (5), 522–528. https://doi.org/10.1038/nchembio.2328.
(215) Neu, A.; Neu, U.; Fuchs, A.-L.; Schlager, B.; Sprangers, R. An Excess of Catalytically Required Motions Inhibits the Scavenger Decapping Enzyme. *Nat. Chem. Biol.* **2015**, *11* (9), 697–704. https://doi.org/10.1038/nchembio.1866.
(216) Sattler, M.; Schleucher, J.; Griesinger, C. Heteronuclear Multidimensional NMR Experiments for the Structure Determination of Proteins in Solution Employing Pulsed Field Gradients. *Prog. Nucl. Magn. Reson. Spectrosc.* **1999**, *34* (2), 93–158. https://doi.org/Doi 10.1016/S0079-6565(98)00025-9.
(217) Gardner, K. H.; Zhang, X.; Gehring, K.; Kay, L. E. Solution NMR Studies of a 42 KDa *Escherichia C Oli* Maltose Binding Protein/β-Cyclodextrin Complex: Chemical Shift Assignments and Analysis. *J. Am. Chem. Soc.* **1998**, *120* (45), 11738–11748. https://doi.org/10.1021/ja982019w.
(218) Hilty, C.; Fernández, C.; Wider, G.; Wüthrich, K. Side Chain NMR Assignments in the Membrane Protein OmpX Reconstituted in DHPC Micelles. *J. Biomol. NMR* **2002**, *23* (4), 289–301.
(219) Tugarinov, V.; Kay, L. E. Side Chain Assignments of Ile Delta 1 Methyl Groups in High Molecular Weight Proteins: An Application to a 46 Ns Tumbling Molecule. *J. Am. Chem. Soc.* **2003**, *125* (19), 5701–5706. https://doi.org/10.1021/ja021452+.
(220) Tugarinov, V.; Venditti, V.; Marius Clore, G. A NMR Experiment for Simultaneous Correlations of Valine and Leucine/Isoleucine Methyls with Carbonyl Chemical Shifts in Proteins. *J. Biomol. NMR* **2014**, *58* (1), 1–8. https://doi.org/10.1007/s10858-013-9803-1.
(221) Sheppard, D.; Guo, C.; Tugarinov, V. Methyl-Detected "out-and-Back" NMR Experiments for Simultaneous Assignments of Alabeta and Ilegamma2 Methyl Groups in Large Proteins. *J. Biomol. NMR* **2009**, *43* (4), 229–238. https://doi.org/10.1007/s10858-009-9305-3.
(222) Bax, A.; Delaglio, F.; Grzesiek, S.; Vuister, G. W. Resonance Assignment of Methionine Methyl Groups and Chi 3 Angular Information from Long-Range Proton-Carbon and Carbon-Carbon J Correlation in a Calmodulin-Peptide Complex. *J. Biomol. NMR* **1994**, *4* (6), 787–797.
(223) Tugarinov, V.; Muhandiram, R.; Ayed, A.; Kay, L. E. Four-Dimensional NMR Spectroscopy of a 723-Residue Protein: Chemical Shift Assignments and Secondary Structure of Malate Synthase g. *J. Am. Chem. Soc.* **2002**, *124* (34), 10025–10035.
(224) Audin, M. J. C.; Dorn, G.; Fromm, S. A.; Reiss, K.; Schütz, S.; Vorländer, M. K.; Sprangers, R. The Archaeal Exosome: Identification and Quantification of Site-Specific Motions That Correlate with Cap and RNA Binding. *Angew. Chem. Int. Ed Engl.* **2013**, *52* (32), 8312–8316. https://doi.org/10.1002/anie.201302811.
(225) Karagöz, G. E.; Duarte, A. M. S.; Ippel, H.; Uetrecht, C.; Sinnige, T.; van Rosmalen, M.; Hausmann, J.; Heck, A. J. R.; Boelens, R.; Rüdiger, S. G. D. N-Terminal Domain of Human Hsp90 Triggers Binding to the Cochaperone P23. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108* (2), 580–585. https://doi.org/10.1073/pnas.1011867108.
(226) Ogunjimi, A. A.; Wiesner, S.; Briant, D. J.; Varelas, X.; Sicheri, F.; Forman-Kay, J.; Wrana, J. L. The Ubiquitin Binding Region of the Smurf HECT Domain Facilitates Polyubiquitylation and Binding of Ubiquitylated Substrates. *J. Biol. Chem.* **2010**, *285* (9), 6308–6315. https://doi.org/10.1074/jbc.M109.044537.
(227) Sprangers, R.; Gribun, A.; Hwang, P. M.; Houry, W. A.; Kay, L. E. Quantitative NMR Spectroscopy of Supramolecular Complexes: Dynamic Side Pores in ClpP Are Important for Product Release. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (46), 16678–16683. https://doi.org/10.1073/pnas.0507370102.
(228) Amero, C.; Asunción Durá, M.; Noirclerc-Savoye, M.; Perollier, A.; Gallet, B.; Plevin, M. J.; Vernet, T.; Franzetti, B.; Boisbouvier, J. A Systematic Mutagenesis-Driven Strategy for Site-Resolved NMR Studies of Supramolecular Assemblies. *J. Biomol. NMR* **2011**, *50* (3), 229–236. https://doi.org/10.1007/s10858-011-9513-5.
(229) Religa, T. L.; Sprangers, R.; Kay, L. E. Dynamic Regulation of Archaeal Proteasome Gate Opening as Studied by TROSY NMR. *Science* **2010**, *328* (5974), 98–102. https://doi.org/10.1126/science.1184991.

(230) Fesik, S. W.; Zuiderweg, E. R. P. Heteronuclear Three-Dimensional Nmr Spectroscopy. A Strategy for the Simplification of Homonuclear Two-Dimensional NMR Spectra. *J. Magn. Reson. 1969* **1988**, *78* (3), 588–593. https://doi.org/10.1016/0022-2364(88)90144-8.
(231) Clore, G. M.; Kay, L. E.; Bax, A.; Gronenborn, A. M. Four-Dimensional 13C/13C-Edited Nuclear Overhauser Enhancement Spectroscopy of a Protein in Solution: Application to Interleukin 1 Beta. *Biochemistry* **1991**, *30* (1), 12–18.
(232) Zuiderweg, E. R. P.; Petros, A. M.; Fesik, S. W.; Olejniczak, E. T. Four-Dimensional [13C,1H,13C,1H] HMQC-NOE-HMQC NMR Spectroscopy: Resolving Tertiary Nuclear Overhauser Effect Distance Constraints in the Spectra of Larger Proteins. *J. Am. Chem. Soc.* **1991**, *113* (1), 370–372. https://doi.org/10.1021/ja00001a060.
(233) Mobli, M.; Hoch, J. C. Nonuniform Sampling and Non-Fourier Signal Processing Methods in Multidimensional NMR. *Prog. Nucl. Magn. Reson. Spectrosc.* **2014**, *83*, 21–41. https://doi.org/10.1016/j.pnmrs.2014.09.002.
(234) Hyberts, S. G.; Arthanari, H.; Robson, S. A.; Wagner, G. Perspectives in Magnetic Resonance: NMR in the Post-FFT Era. *J. Magn. Reson.* **2014**, *241*, 60–73. https://doi.org/10.1016/j.jmr.2013.11.014.
(235) Vuister, G. W.; Clore, G. M.; Gronenborn, A. M.; Powers, R.; Garrett, D. S.; Tschudin, R.; Bax, A. Increased Resolution and Improved Spectral Quality in Four-Dimensional 13C/13C-Separated HMQC-NOESY-HMQC Spectra Using Pulsed Field Gradients. *J. Magn. Reson. B* **1993**, *101* (2), 210–213. https://doi.org/10.1006/jmrb.1993.1035.
(236) Zwahlen, C.; Gardner, K. H.; Sarma, S. P.; Horita, D. A.; Byrd, R. A.; Kay, L. E. An NMR Experiment for Measuring Methyl−Methyl NOEs in ^{13} C-Labeled Proteins with High Resolution. *J. Am. Chem. Soc.* **1998**, *120* (30), 7617–7625. https://doi.org/10.1021/ja981205z.
(237) Xu, Y.; Matthews, S. MAP-XSII: An Improved Program for the Automatic Assignment of Methyl Resonances in Large Proteins. *J. Biomol. NMR* **2013**, *55* (2), 179–187. https://doi.org/10.1007/s10858-012-9700-z.
(238) Chao, F.-A.; Kim, J.; Xia, Y.; Milligan, M.; Rowe, N.; Veglia, G. FLAMEnGO 2.0: An Enhanced Fuzzy Logic Algorithm for Structure-Based Assignment of Methyl Group Resonances. *J. Magn. Reson. San Diego Calif 1997* **2014**, *245*, 17–23. https://doi.org/10.1016/j.jmr.2014.04.012.
(239) Pritišanac, I.; Degiacomi, M. T.; Alderson, T. R.; Carneiro, M. G.; Ab, E.; Siegal, G.; Baldwin, A. J. Automatic Assignment of Methyl-NMR Spectra of Supramolecular Machines Using Graph Theory. *J. Am. Chem. Soc.* **2017**, *139* (28), 9523–9533. https://doi.org/10.1021/jacs.6b11358.
(240) Monneau, Y. R.; Rossi, P.; Bhaumik, A.; Huang, C.; Jiang, Y.; Saleh, T.; Xie, T.; Xing, Q.; Kalodimos, C. G. Automatic Methyl Assignment in Large Proteins by the MAGIC Algorithm. *J. Biomol. NMR* **2017**, *69* (4), 215–227. https://doi.org/10.1007/s10858-017-0149-y.
(241) Venditti, V.; Fawzi, N. L.; Clore, G. M. Automated Sequence- and Stereo-Specific Assignment of Methyl-Labeled Proteins by Paramagnetic Relaxation and Methyl-Methyl Nuclear Overhauser Enhancement Spectroscopy. *J. Biomol. NMR* **2011**, *51* (3), 319–328. https://doi.org/10.1007/s10858-011-9559-4.
(242) Lescanne, M.; Skinner, S. P.; Blok, A.; Timmer, M.; Cerofolini, L.; Fragai, M.; Luchinat, C.; Ubbink, M. Methyl Group Assignment Using Pseudocontact Shifts with PARAssign. *J. Biomol. NMR* **2017**, *69* (4), 183–195. https://doi.org/10.1007/s10858-017-0136-3.
(243) Stoffregen, M. C.; Schwer, M. M.; Renschler, F. A.; Wiesner, S. Methionine Scanning as an NMR Tool for Detecting and Analyzing Biomolecular Interaction Surfaces. *Struct. Lond. Engl. 1993* **2012**, *20* (4), 573–581. https://doi.org/10.1016/j.str.2012.02.012.
(244) Cunningham, B. C.; Wells, J. A. High-Resolution Epitope Mapping of HGH-Receptor Interactions by Alanine-Scanning Mutagenesis. *Science* **1989**, *244* (4908), 1081–1085.
(245) Clackson, T.; Wells, J. A. A Hot Spot of Binding Energy in a Hormone-Receptor Interface. *Science* **1995**, *267* (5196), 383–386.
(246) Brooks, D. J.; Fresco, J. R.; Lesk, A. M.; Singh, M. Evolution of Amino Acid Frequencies in Proteins over Deep Time: Inferred Order of Introduction of Amino Acids into the Genetic Code. *Mol. Biol. Evol.* **2002**, *19* (10), 1645–1655. https://doi.org/10.1093/oxfordjournals.molbev.a003988.
(247) McCaldon, P.; Argos, P. Oligopeptide Biases in Protein Sequences and Their Use in Predicting Protein Coding Regions in Nucleotide Sequences. *Proteins* **1988**, *4* (2), 99–122. https://doi.org/10.1002/prot.340040204.
(248) Miller, S.; Janin, J.; Lesk, A. M.; Chothia, C. Interior and Surface of Monomeric Proteins. *J. Mol. Biol.* **1987**, *196* (3), 641–656.
(249) Schanda, P.; Kupce, E.; Brutscher, B. SOFAST-HMQC Experiments for Recording Two-Dimensional Heteronuclear Correlation Spectra of Proteins within a Few Seconds. *J. Biomol. NMR* **2005**, *33* (4), 199–211. https://doi.org/10.1007/s10858-005-4425-x.
(250) Amero, C.; Schanda, P.; Durá, M. A.; Ayala, I.; Marion, D.; Franzetti, B.; Brutscher, B.; Boisbouvier, J. Fast Two-Dimensional NMR Spectroscopy of High Molecular Weight Protein Assemblies. *J. Am. Chem. Soc.* **2009**, *131* (10), 3448–3449. https://doi.org/10.1021/ja809880p.
(251) Lorentzen, E.; Dziembowski, A.; Lindner, D.; Seraphin, B.; Conti, E. RNA Channelling by the Archaeal Exosome. *EMBO Rep.* **2007**, *8* (5), 470–476. https://doi.org/10.1038/sj.embor.7400945.
(252) Schütz, S.; Nöldeke, E. R.; Sprangers, R. A Synergistic Network of Interactions Promotes the Formation of in Vitro Processing Bodies and Protects MRNA against Decapping. *Nucleic Acids Res.* **2017**, *45* (11), 6911–6922. https://doi.org/10.1093/nar/gkx353.
(253) Keating, C. D. Aqueous Phase Separation as a Possible Route to Compartmentalization of Biological Molecules. *Acc Chem Res* **2012**, *45* (12), 2114–2124. https://doi.org/10.1021/ar200294y.
(254) Toretsky, J. A.; Wright, P. E. Assemblages: Functional Units Formed by Cellular Phase Separation. *J Cell Biol* **2014**, *206* (5), 579–588. https://doi.org/10.1083/jcb.201404124.

(255) Decker, C. J.; Parker, R. P-Bodies and Stress Granules: Possible Roles in the Control of Translation and MRNA Degradation. *Cold Spring Harb Perspect Biol* **2012**, *4* (9), a012286. https://doi.org/10.1101/cshperspect.a012286.
(256) Uversky, V. N.; Kuznetsova, I. M.; Turoverov, K. K.; Zaslavsky, B. Intrinsically Disordered Proteins as Crucial Constituents of Cellular Aqueous Two Phase Systems and Coacervates. *FEBS Lett* **2015**, *589* (1), 15–22. https://doi.org/10.1016/j.febslet.2014.11.028.
(257) Walters, R. W.; Muhlrad, D.; Garcia, J.; Parker, R. Differential Effects of Ydj1 and Sis1 on Hsp70-Mediated Clearance of Stress Granules in Saccharomyces Cerevisiae. *RNA* **2015**, *21* (9), 1660–1671. https://doi.org/10.1261/rna.053116.115.
(258) Ramaswami, M.; Taylor, J. P.; Parker, R. Altered Ribostasis: RNA-Protein Granules in Degenerative Disorders. *Cell* **2013**, *154* (4), 727–736. https://doi.org/10.1016/j.cell.2013.07.038.
(259) Teixeira, D.; Sheth, U.; Valencia-Sanchez, M. A.; Brengues, M.; Parker, R. Processing Bodies Require RNA for Assembly and Contain Nontranslating MRNAs. *RNA* **2005**, *11* (4), 371–382. https://doi.org/10.1261/rna.7258505.
(260) Ingelfinger, D.; Arndt-Jovin, D. J.; Luhrmann, R.; Achsel, T. The Human LSm1-7 Proteins Colocalize with the MRNA-Degrading Enzymes Dcp1/2 and Xrnl in Distinct Cytoplasmic Foci. *RNA* **2002**, *8* (12), 1489–1501.
(261) Wang, C. Y.; Chen, W. L.; Wang, S. W. Pdc1 Functions in the Assembly of P Bodies in Schizosaccharomyces Pombe. *Mol Cell Biol* **2013**, *33* (6), 1244–1253. https://doi.org/10.1128/MCB.01583-12.
(262) Brengues, M.; Teixeira, D.; Parker, R. Movement of Eukaryotic MRNAs between Polysomes and Cytoplasmic Processing Bodies. *Science* **2005**, *310* (5747), 486–489. https://doi.org/10.1126/science.1115791.
(263) Fuchs, A.-L.; Neu, A.; Sprangers, R. A General Method for Rapid and Cost-Efficient Large-Scale Production of 5' Capped RNA. *RNA N. Y. N* **2016**, *22* (9), 1454–1466. https://doi.org/10.1261/rna.056614.116.
(264) Audin, M. J. C.; Wurm, J. P.; Cvetkovic, M. A.; Sprangers, R. The Oligomeric Architecture of the Archaeal Exosome Is Important for Processive and Efficient RNA Degradation. *Nucleic Acids Res.* **2016**, *44* (6), 2962–2973. https://doi.org/10.1093/nar/gkw062.
(265) Ramos, A.; Varani, G. A New Method to Detect Long-Range Protein-RNA Contacts: NMR Detection of Electron-Proton Relaxation Induced by Nitroxide Spin-Labeled RNA. *J. Am. Chem. Soc.* **1998**, *120* (42), 10992–10993. https://doi.org/Doi 10.1021/Ja982496e.
(266) Schindelin, J.; Arganda-Carreras, I.; Frise, E.; Kaynig, V.; Longair, M.; Pietzsch, T.; Preibisch, S.; Rueden, C.; Saalfeld, S.; Schmid, B.; et al. Fiji: An Open-Source Platform for Biological-Image Analysis. *Nat Methods* **2012**, *9* (7), 676–682. https://doi.org/10.1038/nmeth.2019.
(267) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. NMRPipe: A Multidimensional Spectral Processing System Based on UNIX Pipes. *J Biomol NMR* **1995**, *6* (3), 277–293.
(268) Conicella, A. E.; Zerze, G. H.; Mittal, J.; Fawzi, N. L. ALS Mutations Disrupt Phase Separation Mediated by Alpha-Helical Structure in the TDP-43 Low-Complexity C-Terminal Domain. *Structure* **2016**, *24* (9), 1537–1549. https://doi.org/10.1016/j.str.2016.07.007.
(269) Burke, K. A.; Janke, A. M.; Rhine, C. L.; Fawzi, N. L. Residue-by-Residue View of In Vitro FUS Granules That Bind the C-Terminal Domain of RNA Polymerase II. *Mol Cell* **2015**, *60* (2), 231–241. https://doi.org/10.1016/j.molcel.2015.09.006.
(270) Shevtsov, S. P.; Dundr, M. Nucleation of Nuclear Bodies by RNA. *Nat Cell Biol* **2011**, *13* (2), 167–173. https://doi.org/10.1038/ncb2157.
(271) Ostareck, D. H.; Naarmann-de Vries, I. S.; Ostareck-Lederer, A. DDX6 and Its Orthologs as Modulators of Cellular and Viral RNA Expression: Functions of DDX6 and Its Orthologs. *Wiley Interdiscip. Rev. RNA* **2014**, *5* (5), 659–678. https://doi.org/10.1002/wrna.1237.
(272) Marguerat, S.; Schmidt, A.; Codlin, S.; Chen, W.; Aebersold, R.; Bahler, J. Quantitative Analysis of Fission Yeast Transcriptomes and Proteomes in Proliferating and Quiescent Cells. *Cell* **2012**, *151* (3), 671–683. https://doi.org/10.1016/j.cell.2012.09.019.
(273) Shin, Y.; Brangwynne, C. P. Liquid Phase Condensation in Cell Physiology and Disease. *Science* **2017**, *357* (6357). https://doi.org/10.1126/science.aaf4382.
(274) Tompa, P.; Schad, E.; Tantos, A.; Kalmar, L. Intrinsically Disordered Proteins: Emerging Interaction Specialists. *Curr Opin Struct Biol* **2015**, *35*, 49–59. https://doi.org/10.1016/j.sbi.2015.08.009.
(275) Bhattacharyya, S. N.; Habermacher, R.; Martine, U.; Closs, E. I.; Filipowicz, W. Relief of MicroRNA-Mediated Translational Repression in Human Cells Subjected to Stress. *Cell* **2006**, *125* (6), 1111–1124. https://doi.org/10.1016/j.cell.2006.04.031.
(276) Wu, H.; Fuxreiter, M. The Structure and Dynamics of Higher-Order Assemblies: Amyloids, Signalosomes, and Granules. *Cell* **2016**, *165* (5), 1055–1066. https://doi.org/10.1016/j.cell.2016.05.004.
(277) Mitrea, D. M.; Chandra, B.; Ferrolino, M. C.; Gibbs, E. B.; Tolbert, M.; White, M. R.; Kriwacki, R. W. Methods for Physical Characterization of Phase-Separated Bodies and Membrane-Less Organelles. *J. Mol. Biol.* **2018**, *430* (23), 4773–4805. https://doi.org/10.1016/j.jmb.2018.07.006.
(278) Baldwin, A. J.; Kay, L. E. NMR Spectroscopy Brings Invisible Protein States into Focus. *Nat. Chem. Biol.* **2009**, *5* (11), 808–814. https://doi.org/10.1038/nchembio.238.
(279) Wright, P. E.; Dyson, H. J. Intrinsically Disordered Proteins in Cellular Signalling and Regulation. *Nat. Rev. Mol. Cell Biol.* **2015**, *16* (1), 18–29. https://doi.org/10.1038/nrm3920.
(280) Hu, K.-N.; Tycko, R. What Can Solid State NMR Contribute to Our Understanding of Protein Folding? *Biophys. Chem.* **2010**, *151* (1–2), 10–21. https://doi.org/10.1016/j.bpc.2010.05.009.
(281) Renault, M.; Cukkemane, A.; Baldus, M. Solid-State NMR Spectroscopy on Complex Biomolecules. *Angew. Chem. Int. Ed Engl.* **2010**, *49* (45), 8346–8357. https://doi.org/10.1002/anie.201002823.

(282) Kühn, J.; Wong, L. E.; Pirkuliyeva, S.; Schulz, K.; Schwiegk, C.; Fünfgeld, K. G.; Keppler, S.; Batista, F. D.; Urlaub, H.; Habeck, M.; et al. The Adaptor Protein CIN85 Assembles Intracellular Signaling Clusters for B Cell Activation. *Sci. Signal.* **2016**, *9* (434), ra66. https://doi.org/10.1126/scisignal.aad6275.
(283) Brady, J. P.; Farber, P. J.; Sekhar, A.; Lin, Y.-H.; Huang, R.; Bah, A.; Nott, T. J.; Chan, H. S.; Baldwin, A. J.; Forman-Kay, J. D.; et al. Structural and Hydrodynamic Properties of an Intrinsically Disordered Region of a Germ Cell-Specific Protein on Phase Separation. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114* (39), E8194–E8203. https://doi.org/10.1073/pnas.1706197114.
(284) Ackermann, B. E.; Debelouchina, G. T. Heterochromatin Protein HP1α Gelation Dynamics Revealed by Solid-State NMR Spectroscopy. *Angew. Chem. Int. Ed Engl.* **2019**. https://doi.org/10.1002/anie.201901141.
(285) Kennedy, S. B.; deAzevedo, E. R.; Petka, W. A.; Russell, T. P.; Tirrell, D. A.; Hong, M. Dynamic Structure of a Protein Hydrogel: A Solid-State NMR Study. *Macromolecules* **2001**, *34* (25), 8675–8685. https://doi.org/10.1021/ma010768j.
(286) Labokha, A. A.; Gradmann, S.; Frey, S.; Hülsmann, B. B.; Urlaub, H.; Baldus, M.; Görlich, D. Systematic Analysis of Barrier-Forming FG Hydrogels from Xenopus Nuclear Pore Complexes. *EMBO J.* **2013**, *32* (2), 204–218. https://doi.org/10.1038/emboj.2012.302.
(287) Nagy-Smith, K.; Moore, E.; Schneider, J.; Tycko, R. Molecular Structure of Monomorphic Peptide Fibrils within a Kinetically Trapped Hydrogel Network. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112* (32), 9816–9821. https://doi.org/10.1073/pnas.1509313112.
(288) Ward, M. E.; Daniëls, M. A.; van Kappel, E. C.; Maurice, M. M.; Baldus, M. Investigations of Dynamic Amyloid-like Structures of the Wnt Signalling Pathway by Solid-State NMR. *Chem. Commun. Camb. Engl.* **2018**, *54* (32), 3959–3962. https://doi.org/10.1039/C8CC01346B.
(289) Xiang, S.; Kato, M.; Wu, L. C.; Lin, Y.; Ding, M.; Zhang, Y.; Yu, Y.; McKnight, S. L. The LC Domain of HnRNPA2 Adopts Similar Conformations in Hydrogel Polymers, Liquid-like Droplets, and Nuclei. *Cell* **2015**, *163* (4), 829–839. https://doi.org/10.1016/j.cell.2015.10.040.
(290) Luo, Y.; Na, Z.; Slavoff, S. A. P-Bodies: Composition, Properties, and Functions. *Biochemistry* **2018**, *57* (17), 2424–2431. https://doi.org/10.1021/acs.biochem.7b01162.
(291) Ling, S. H. M.; Decker, C. J.; Walsh, M. A.; She, M.; Parker, R.; Song, H. Crystal Structure of Human Edc3 and Its Functional Implications. *Mol. Cell. Biol.* **2008**, *28* (19), 5965–5976. https://doi.org/10.1128/MCB.00761-08.
(292) Andronesi, O. C.; Becker, S.; Seidel, K.; Heise, H.; Young, H. S.; Baldus, M. Determination of Membrane Protein Structure and Dynamics by Magic-Angle-Spinning Solid-State NMR Spectroscopy. *J. Am. Chem. Soc.* **2005**, *127* (37), 12965–12974. https://doi.org/10.1021/ja0530164.
(293) Mance, D.; Sinnige, T.; Kaplan, M.; Narasimhan, S.; Daniëls, M.; Houben, K.; Baldus, M.; Weingarth, M. An Efficient Labelling Approach to Harness Backbone and Side-Chain Protons in (1) H-Detected Solid-State NMR Spectroscopy. *Angew. Chem. Int. Ed Engl.* **2015**, *54* (52), 15799–15803. https://doi.org/10.1002/anie.201509170.
(294) Medeiros-Silva, J.; Mance, D.; Daniëls, M.; Jekhmane, S.; Houben, K.; Baldus, M.; Weingarth, M. 1 H-Detected Solid-State NMR Studies of Water-Inaccessible Proteins In Vitro and In Situ. *Angew. Chem. Int. Ed Engl.* **2016**, *55* (43), 13606–13610. https://doi.org/10.1002/anie.201606594.
(295) Morris, G. A.; Freeman, R. Enhancement of Nuclear Magnetic Resonance Signals by Polarization Transfer. *J. Am. Chem. Soc.* **1979**, *101* (3), 760–762. https://doi.org/10.1021/ja00497a058.
(296) Baldus, M.; Meier, B. H. Total Correlation Spectroscopy in the Solid State. The Use of Scalar Couplings to Determine the Through-Bond Connectivity. *J. Magn. Reson. Ser. AJ Magn Reson Ser A* **1996**, *121* (1), 65–69.
(297) Weingarth, M.; Bodenhausen, G.; Tekely, P. Low-Power Decoupling at High Spinning Frequencies in High Static Fields. *J. Magn. Reson. San Diego Calif 1997* **2009**, *199* (2), 238–241. https://doi.org/10.1016/j.jmr.2009.04.015.
(298) Lange, A.; Luca, S.; Baldus, M. Structural Constraints from Proton-Mediated Rare-Spin Correlation Spectroscopy in Rotating Solids. *J. Am. Chem. Soc.* **2002**, *124* (33), 9704–9705.
(299) Schmid, N.; Eichenberger, A. P.; Choutko, A.; Riniker, S.; Winger, M.; Mark, A. E.; van Gunsteren, W. F. Definition and Testing of the GROMOS Force-Field Versions 54A7 and 54B7. *Eur. Biophys. J. EBJ* **2011**, *40* (7), 843–856. https://doi.org/10.1007/s00249-011-0700-9.
(300) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **2008**, *4* (3), 435–447. https://doi.org/10.1021/ct700301q.
(301) Sali, A.; Potterton, L.; Yuan, F.; van Vlijmen, H.; Karplus, M. Evaluation of Comparative Protein Modeling by MODELLER. *Proteins* **1995**, *23* (3), 318–326. https://doi.org/10.1002/prot.340230306.
(302) Linder, P.; Lasko, P. F.; Ashburner, M.; Leroy, P.; Nielsen, P. J.; Nishi, K.; Schnier, J.; Slonimski, P. P. Birth of the D-E-A-D Box. *Nature* **1989**, *337* (6203), 121–122. https://doi.org/10.1038/337121a0.
(303) Fairman-Williams, M. E.; Guenther, U.-P.; Jankowsky, E. SF1 and SF2 Helicases: Family Matters. *Curr. Opin. Struct. Biol.* **2010**, *20* (3), 313–324. https://doi.org/10.1016/j.sbi.2010.03.011.
(304) Strahl-Bolsinger, S.; Tanner, W. A Yeast Gene Encoding a Putative RNA Helicase of the "DEAD"-Box Family. *Yeast Chichester Engl.* **1993**, *9* (4), 429–432. https://doi.org/10.1002/yea.320090414.
(305) Chang, L.-C.; Lee, F.-J. S. The RNA Helicase Dhh1p Cooperates with Rbp1p to Promote Porin MRNA Decay via Its Non-Conserved C-Terminal Domain. *Nucleic Acids Res.* **2012**, *40* (3), 1331–1344. https://doi.org/10.1093/nar/gkr803.
(306) Tseng-Rogenski, S. S.-I.; Chong, J.-L.; Thomas, C. B.; Enomoto, S.; Berman, J.; Chang, T.-H. Functional Conservation of Dhh1p, a Cytoplasmic DExD/H-Box Protein Present in Large Complexes. *Nucleic Acids Res.* **2003**, *31* (17), 4995–5002.
(307) Presnyak, V.; Coller, J. The DHH1/RCKp54 Family of Helicases: An Ancient Family of Proteins That Promote Translational Silencing. *Biochim. Biophys. Acta* **2013**, *1829* (8), 817–823. https://doi.org/10.1016/j.bbagrm.2013.03.006.

(308) Maekawa, H.; Nakagawa, T.; Uno, Y.; Kitamura, K.; Shimoda, C. The Ste13+ Gene Encoding a Putative RNA Helicase Is Essential for Nitrogen Starvation-Induced G1 Arrest and Initiation of Sexual Development in the Fission Yeast Schizosaccharomyces Pombe. *Mol. Gen. Genet. MGG* **1994**, *244* (5), 456–464.
(309) Navarro, R. E.; Shim, E. Y.; Kohara, Y.; Singson, A.; Blackwell, T. K. Cgh-1, a Conserved Predicted RNA Helicase Required for Gametogenesis and Protection from Physiological Germline Apoptosis in C. Elegans. *Dev. Camb. Engl.* **2001**, *128* (17), 3221–3232.
(310) de Valoir, T.; Tucker, M. A.; Belikoff, E. J.; Camp, L. A.; Bolduc, C.; Beckingham, K. A Second Maternally Expressed Drosophila Gene Encodes a Putative RNA Helicase of the "DEAD Box" Family. *Proc. Natl. Acad. Sci. U. S. A.* **1991**, *88* (6), 2113–2117.
(311) Ladomery, M.; Wade, E.; Sommerville, J. Xp54, the Xenopus Homologue of Human RNA Helicase P54, Is an Integral Component of Stored MRNP Particles in Oocytes. *Nucleic Acids Res.* **1997**, *25* (5), 965–973.
(312) Akao, Y.; Marukawa, O.; Morikawa, H.; Nakao, K.; Kamei, M.; Hachiya, T.; Tsujimoto, Y. The Rck/P54 Candidate Proto-Oncogene Product Is a 54-Kilodalton D-E-A-D Box Protein Differentially Expressed in Human and Mouse Tissues. *Cancer Res.* **1995**, *55* (15), 3444–3449.
(313) Akao, Y.; Yoshida, H.; Matsumoto, K.; Matsui, T.; Hogetu, K.; Tanaka, N.; Usukura, J. A Tumour-Associated DEAD-Box Protein, Rck/P54 Exhibits RNA Unwinding Activity toward c-Myc RNAs in Vitro. *Genes Cells Devoted Mol. Cell. Mech.* **2003**, *8* (8), 671–676.
(314) Lu, D.; Yunis, J. J. Cloning, Expression and Localization of an RNA Helicase Gene from a Human Lymphoid Cell Line with Chromosomal Breakpoint 11q23.3. *Nucleic Acids Res.* **1992**, *20* (8), 1967–1972.
(315) Roca, A. I.; Cox, M. M. RecA Protein: Structure, Function, and Role in Recombinational DNA Repair. *Prog. Nucleic Acid Res. Mol. Biol.* **1997**, *56*, 129–223.
(316) Sondheimer, N.; Lindquist, S. Rnq1: An Epigenetic Modifier of Protein Function in Yeast. *Mol. Cell* **2000**, *5* (1), 163–172. https://doi.org/10.1016/S1097-2765(00)80412-8.
(317) Derkatch, I. L.; Uptain, S. M.; Outeiro, T. F.; Krishnan, R.; Lindquist, S. L.; Liebman, S. W. Effects of Q/N-Rich, PolyQ, and Non-PolyQ Amyloids on the de Novo Formation of the [PSI+] Prion in Yeast and Aggregation of Sup35 in Vitro. *Proc. Natl. Acad. Sci.* **2004**, *101* (35), 12934–12939. https://doi.org/10.1073/pnas.0404968101.
(318) Cheng, Z.; Coller, J.; Parker, R.; Song, H. Crystal Structure and Functional Analysis of DEAD-Box Protein Dhh1p. *RNA N. Y. N* **2005**, *11* (8), 1258–1270. https://doi.org/10.1261/rna.2920905.
(319) Sengoku, T.; Nureki, O.; Nakamura, A.; Kobayashi, S.; Yokoyama, S. Structural Basis for RNA Unwinding by the DEAD-Box Protein Drosophila Vasa. *Cell* **2006**, *125* (2), 287–300. https://doi.org/10.1016/j.cell.2006.01.054.
(320) Del Campo, M.; Lambowitz, A. M. Structure of the Yeast DEAD Box Protein Mss116p Reveals Two Wedges That Crimp RNA. *Mol. Cell* **2009**, *35* (5), 598–609. https://doi.org/10.1016/j.molcel.2009.07.032.
(321) von Moeller, H.; Basquin, C.; Conti, E. The MRNA Export Protein DBP5 Binds RNA and the Cytoplasmic Nucleoporin NUP214 in a Mutually Exclusive Manner. *Nat. Struct. Mol. Biol.* **2009**, *16* (3), 247–254. https://doi.org/10.1038/nsmb.1561.
(322) Collins, R.; Karlberg, T.; Lehtiö, L.; Schütz, P.; van den Berg, S.; Dahlgren, L.-G.; Hammarström, M.; Weigelt, J.; Schüler, H. The DEXD/H-Box RNA Helicase DDX19 Is Regulated by an α-Helical Switch. *J. Biol. Chem.* **2009**, *284* (16), 10296–10300. https://doi.org/10.1074/jbc.C900018200.
(323) Dutta, A.; Zheng, S.; Jain, D.; Cameron, C. E.; Reese, J. C. Intermolecular Interactions within the Abundant DEAD-Box Protein Dhh1 Regulate Its Activity *in Vivo*. *J. Biol. Chem.* **2011**, *286* (31), 27454–27470. https://doi.org/10.1074/jbc.M111.220251.
(324) Minshall, N.; Thom, G.; Standart, N. A Conserved Role of a DEAD Box Helicase in MRNA Masking. *RNA N. Y. N* **2001**, *7* (12), 1728–1742.
(325) Weston, A.; Sommerville, J. Xp54 and Related (DDX6-like) RNA Helicases: Roles in Messenger RNP Assembly, Translation Regulation and RNA Degradation. *Nucleic Acids Res.* **2006**, *34* (10), 3082–3094. https://doi.org/10.1093/nar/gkl409.
(326) Nakamura, A.; Amikura, R.; Hanyu, K.; Kobayashi, S. Me31B Silences Translation of Oocyte-Localizing RNAs through the Formation of Cytoplasmic RNP Complex during Drosophila Oogenesis. *Dev. Camb. Engl.* **2001**, *128* (17), 3233–3242.
(327) Hillebrand, J.; Pan, K.; Kokaram, A.; Barbee, S.; Parker, R.; Ramaswami, M. The Me31B DEAD-Box Helicase Localizes to Postsynaptic Foci and Regulates Expression of a CaMKII Reporter MRNA in Dendrites of Drosophila Olfactory Projection Neurons. *Front. Neural Circuits* **2010**, *4*, 121. https://doi.org/10.3389/fncir.2010.00121.
(328) Kramer, S.; Queiroz, R.; Ellis, L.; Hoheisel, J. D.; Clayton, C.; Carrington, M. The RNA Helicase DHH1 Is Central to the Correct Expression of Many Developmentally Regulated MRNAs in Trypanosomes. *J. Cell Sci.* **2010**, *123* (Pt 5), 699–711. https://doi.org/10.1242/jcs.058511.
(329) Hata, H.; Mitsui, H.; Liu, H.; Bai, Y.; Denis, C. L.; Shimizu, Y.; Sakai, A. Dhh1p, a Putative RNA Helicase, Associates with the General Transcription Factors Pop2p and Ccr4p from Saccharomyces Cerevisiae. *Genetics* **1998**, *148* (2), 571–579.
(330) Coller, J. M.; Tucker, M.; Sheth, U.; Valencia-Sanchez, M. A.; Parker, R. The DEAD Box Helicase, Dhh1p, Functions in MRNA Decapping and Interacts with Both the Decapping and Deadenylase Complexes. *RNA N. Y. N* **2001**, *7* (12), 1717–1727.
(331) Basquin, J.; Roudko, V. V.; Rode, M.; Basquin, C.; Séraphin, B.; Conti, E. Architecture of the Nuclease Module of the Yeast Ccr4-Not Complex: The Not1-Caf1-Ccr4 Interaction. *Mol. Cell* **2012**, *48* (2), 207–218. https://doi.org/10.1016/j.molcel.2012.08.014.

(332) Mathys, H.; Basquin, J.; Ozgur, S.; Czarnocki-Cieciura, M.; Bonneau, F.; Aartse, A.; Dziembowski, A.; Nowotny, M.; Conti, E.; Filipowicz, W. Structural and Biochemical Insights to the Role of the CCR4-NOT Complex and DDX6 ATPase in MicroRNA Repression. *Mol. Cell* **2014**, *54* (5), 751–765. https://doi.org/10.1016/j.molcel.2014.03.036.
(333) Chen, Y.; Boland, A.; Kuzuoğlu-Öztürk, D.; Bawankar, P.; Loh, B.; Chang, C.-T.; Weichenrieder, O.; Izaurralde, E. A DDX6-CNOT1 Complex and W-Binding Pockets in CNOT9 Reveal Direct Links between MiRNA Target Recognition and Silencing. *Mol. Cell* **2014**, *54* (5), 737–750. https://doi.org/10.1016/j.molcel.2014.03.034.
(334) Ozgur, S.; Basquin, J.; Kamenska, A.; Filipowicz, W.; Standart, N.; Conti, E. Structure of a Human 4E-T/DDX6/CNOT1 Complex Reveals the Different Interplay of DDX6-Binding Proteins with the CCR4-NOT Complex. *Cell Rep.* **2015**, *13* (4), 703–711. https://doi.org/10.1016/j.celrep.2015.09.033.
(335) Teixeira, D.; Parker, R. Analysis of P-Body Assembly in Saccharomyces Cerevisiae. *Mol. Biol. Cell* **2007**, *18* (6), 2274–2287. https://doi.org/10.1091/mbc.e07-03-0199.
(336) Gibson, D. G.; Benders, G. A.; Andrews-Pfannkoch, C.; Denisova, E. A.; Baden-Tillson, H.; Zaveri, J.; Stockwell, T. B.; Brownley, A.; Thomas, D. W.; Algire, M. A.; et al. Complete Chemical Synthesis, Assembly, and Cloning of a Mycoplasma Genitalium Genome. *Science* **2008**, *319* (5867), 1215–1220. https://doi.org/10.1126/science.1151721.
(337) Gibson, D. G.; Young, L.; Chuang, R.-Y.; Venter, J. C.; Hutchison, C. A.; Smith, H. O. Enzymatic Assembly of DNA Molecules up to Several Hundred Kilobases. *Nat. Methods* **2009**, *6* (5), 343–345. https://doi.org/10.1038/nmeth.1318.
(338) Lee, W.; Tonelli, M.; Markley, J. L. NMRFAM-SPARKY: Enhanced Software for Biomolecular NMR Spectroscopy. *Bioinforma. Oxf. Engl.* **2015**, *31* (8), 1325–1327. https://doi.org/10.1093/bioinformatics/btu830.
(339) Quiroz, F. G.; Chilkoti, A. Sequence Heuristics to Encode Phase Behaviour in Intrinsically Disordered Protein Polymers. *Nat. Mater.* **2015**, *14* (11), 1164–1171. https://doi.org/10.1038/nmat4418.
(340) Costantini, S.; Colonna, G.; Facchiano, A. M. Amino Acid Propensities for Secondary Structures Are Influenced by the Protein Structural Class. *Biochem. Biophys. Res. Commun.* **2006**, *342* (2), 441–451. https://doi.org/10.1016/j.bbrc.2006.01.159.
(341) Chou, P. Y.; Fasman, G. D. Prediction of Protein Conformation. *Biochemistry* **1974**, *13* (2), 222–245. https://doi.org/10.1021/bi00699a002.
(342) Guo, L.; Shorter, J. It's Raining Liquids: RNA Tunes Viscoelasticity and Dynamics of Membraneless Organelles. *Mol Cell* **2015**, *60* (2), 189–192. https://doi.org/10.1016/j.molcel.2015.10.006.
(343) Bono, F.; Ebert, J.; Lorentzen, E.; Conti, E. The Crystal Structure of the Exon Junction Complex Reveals How It Maintains a Stable Grip on MRNA. *Cell* **2006**, *126* (4), 713–725. https://doi.org/10.1016/j.cell.2006.08.006.
(344) Andersen, C. B. F.; Ballut, L.; Johansen, J. S.; Chamieh, H.; Nielsen, K. H.; Oliveira, C. L. P.; Pedersen, J. S.; Séraphin, B.; Le Hir, H.; Andersen, G. R. Structure of the Exon Junction Core Complex with a Trapped DEAD-Box ATPase Bound to RNA. *Science* **2006**, *313* (5795), 1968–1972. https://doi.org/10.1126/science.1131981.
(345) Nielsen, K. H.; Chamieh, H.; Andersen, C. B. F.; Fredslund, F.; Hamborg, K.; Le Hir, H.; Andersen, G. R. Mechanism of ATP Turnover Inhibition in the EJC. *RNA N. Y. N* **2009**, *15* (1), 67–75. https://doi.org/10.1261/rna.1283109.
(346) Mallam, A. L.; Sidote, D. J.; Lambowitz, A. M. Molecular Insights into RNA and DNA Helicase Evolution from the Determinants of Specificity for a DEAD-Box RNA Helicase. *eLife* **2014**, *3*. https://doi.org/10.7554/eLife.04630.
(347) Yuwen, T.; Brady, J. P.; Kay, L. E. Probing Conformational Exchange in Weakly Interacting, Slowly Exchanging Protein Systems via Off-Resonance R1ρ Experiments: Application to Studies of Protein Phase Separation. *J. Am. Chem. Soc.* **2018**, *140* (6), 2115–2126. https://doi.org/10.1021/jacs.7b09576.
(348) Yuwen, T.; Bah, A.; Brady, J. P.; Ferrage, F.; Bouvignies, G.; Kay, L. E. Measuring Solvent Hydrogen Exchange Rates by Multifrequency Excitation 15N CEST: Application to Protein Phase Separation. *J. Phys. Chem. B* **2018**. https://doi.org/10.1021/acs.jpcb.8b06820.
(349) Aizer, A.; Kalo, A.; Kafri, P.; Shraga, A.; Ben-Yishay, R.; Jacob, A.; Kinor, N.; Shav-Tal, Y. Quantifying MRNA Targeting to P-Bodies in Living Human Cells Reveals Their Dual Role in MRNA Decay and Storage. *J Cell Sci* **2014**, *127* (Pt 20), 4443–4456. https://doi.org/10.1242/jcs.152975.
(350) Lavut, A.; Raveh, D. Sequestration of Highly Expressed MRNAs in Cytoplasmic Granules, P-Bodies, and Stress Granules Enhances Cell Viability. *PLoS Genet* **2012**, *8* (2), e1002527. https://doi.org/10.1371/journal.pgen.1002527.