



PhasAGE

EXCELLENCE HUB ON
PHASE TRANSITIONS IN
AGING AND AGE-RELATED
DISORDERS

TRAINING SCHOOL 1

24 -28 MAY 2021

Computational methods to study Phase Separation

Lecture: Phase separation and emergent functions of intrinsically disordered proteins –

Peter Tompa

This content is shared under the license CC BY-NC-ND



This project received funding from
European Union's Horizon 2020
research and innovation programme
under grant agreement No 952334.



PhasAGE

EXCELLENCE HUB ON
PHASE TRANSITIONS IN AGING
AND AGE-RELATED DISORDERS

TRAINING SCHOOL

Computational Methods to Study Protein Phase Separation

24-28 May 2021 - Online event

Monday May 24th, 14:00 - 18:00 CEST / 13:00 - 17:00 GMT+1

Time (CEST)	Session title	Speaker	Type
14:00 - 15:00	PhasAGE: Excellence Hub on phase transitions in aging and age-related disorders	Sandra Macedo Ribeiro	L
15:00 - 16:45	Phase separation and emergent functions of Intrinsically Disordered Proteins	Peter Tompa	L
16:45 - 17:00	<i>Break</i>		
17:00 - 18:00	Group activity		G

- 1) How it started: three stories**
- 2) An outline of LLPS and emergent functions**
- 3) LLPS and disease**
- 4) Experimental techniques to study LLPS**
- 5) Databases and bioinformatics tools**

1) How it started: three stories



Germline P Granules Are Liquid Droplets That Localize by Controlled Dissolution/Condensation

Clifford P. Brangwynne,^{1,2,3} Christian R. Eckmann,¹ David S. Courson,³ Agata Rybarska,¹ Carsten Hoege,¹ Jöbin Gharakhani,^{2,3} Frank Jülicher,^{2,3} Anthony A. Hyman^{1,3*}

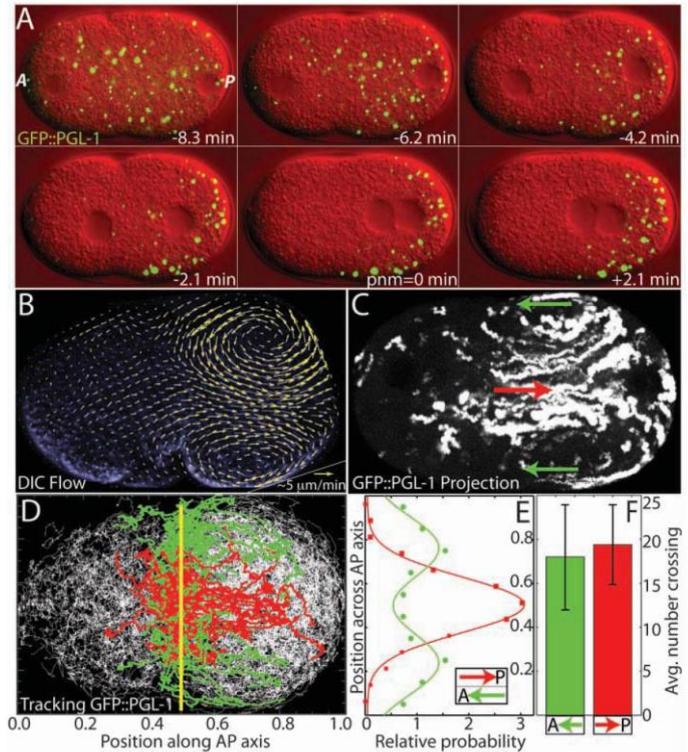
the embryo posterior; however, close to the cortex there was a flux of P granules into the anterior that was of similar magnitude to the posteriorly directed flux (Fig. 1, D to F). This behavior closely matched the overall flow behavior of cytoplasmic material such as yolk granules (6), quantified by particle imaging velocimetry (PIV) (Fig. 1, B and C) (9). P granules cannot preferentially localize to the posterior by convection in the surrounding cytoplasm alone. Thus, flows have little or no role

In sexually reproducing organisms, germ cells generate sperm and eggs. In *C. elegans*, the first germ cell is established when RNA and protein-rich P granules localize to the posterior of the one-cell embryo. It was shown that P granules exhibit liquid-like behaviors, including fusion, dripping, and wetting, and arise by condensation (phase separation). Polarity of the embryo is marked by biased localization at the posterior.

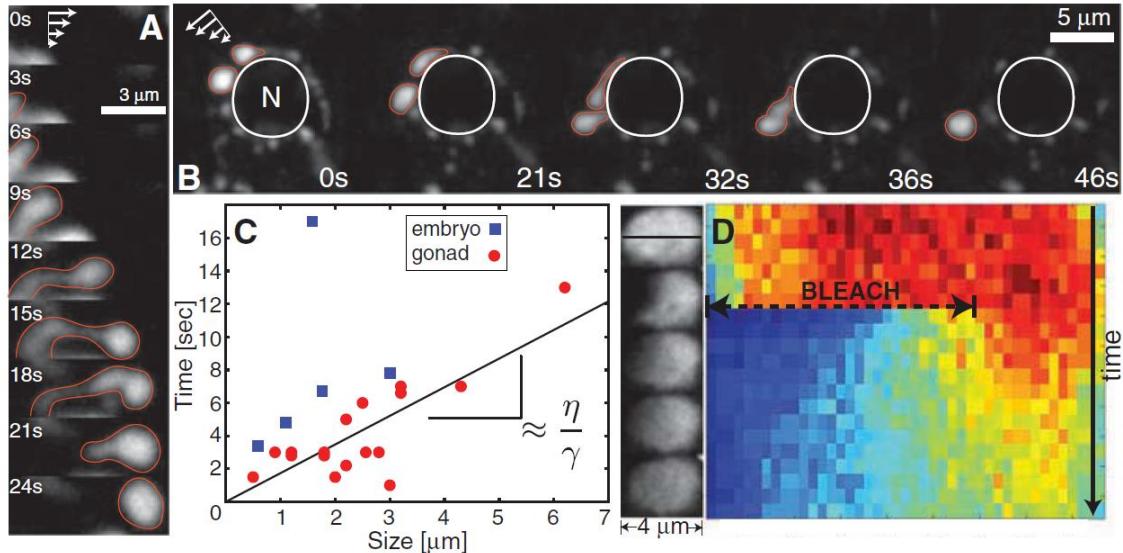


Preferential localization and movement of PGs

- GFP-PGL1, posterior end of embryo -



Liquid behavior of PGs



(A) jetting from nucleus, (B) dripping and fusion, (C) fusion and viscosity, (D) FRAP

LETTER

doi:10.1038/nature10879

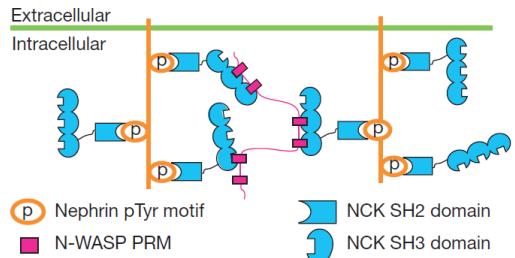
Phase transitions in the assembly of multivalent signalling proteins

Pilong Li^{1*}, Sudeep Banjade^{1*}, Hui-Chun Cheng^{1*}, Soyeon Kim¹, Baoyu Chen¹, Liang Guo², Marc Llaguno³, Javoris V. Hollingsworth⁴, David S. King⁵, Salman F. Banani¹, Paul S. Russo⁴, Qiu-Xing Jiang³, B. Tracy Nixon⁶ & Michael K. Rosen¹

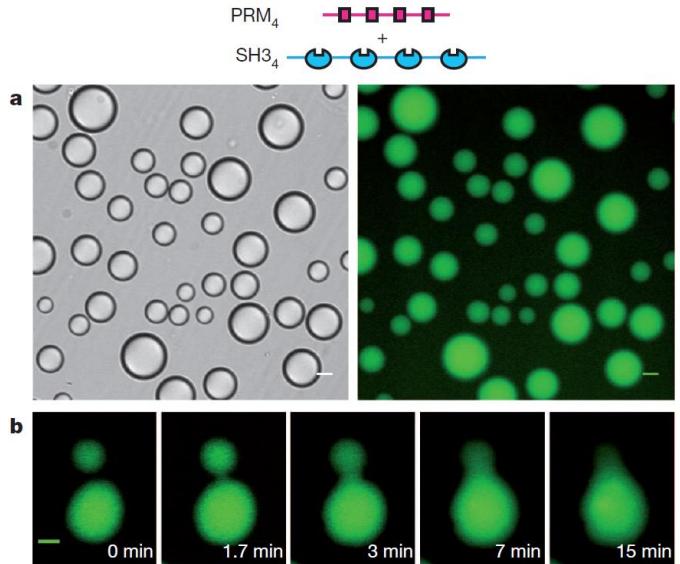
It is shown *in vitro* that mixtures of signaling proteins Wiskott-Aldrich syndrome protein (N-WASP), NCK and phosphorylated nephrin1 (complex for actin cytoskeletal rearrangements) phase separate by liquid-liquid demixing, generating μm -sized liquid droplets. The concentrations needed are directly related to the valency of the interacting species. The phase transition is regulated by nephrin phosphorylation and corresponds to a sharp increase in activity towards an actin nucleation factor, the Arp2/3 complex.



Nephrin-NCK-N-WASP system in cytoskeleton remodeling

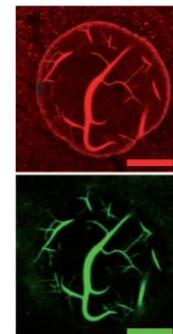
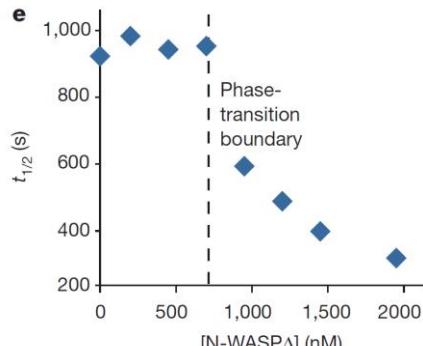
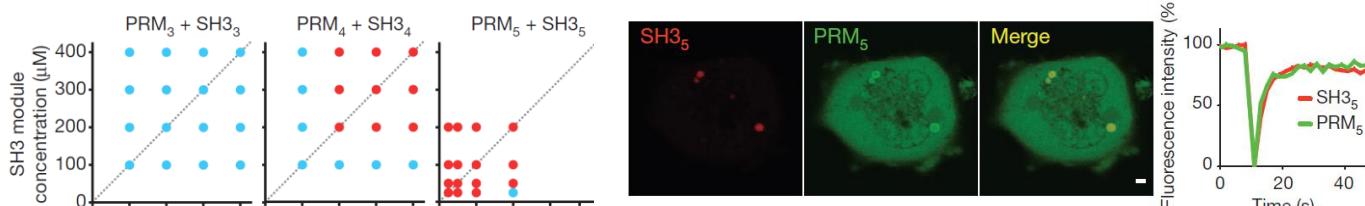


Multivalency
SH2 - pTyr, SH3 - PRM interactions:
signaling through Arp 2/3 to actin



- (A) Differential interference contrast (DIC) microscopy
(B) Wide-field fluorescence microscopy
(C) time-lapse imaging of fusion events

Nephrin-NCK-N-WASP system in cytoskeleton remodeling



$t_{1/2}$ of actin polymerization
(Arp2/3 and rhodamine-actin added)



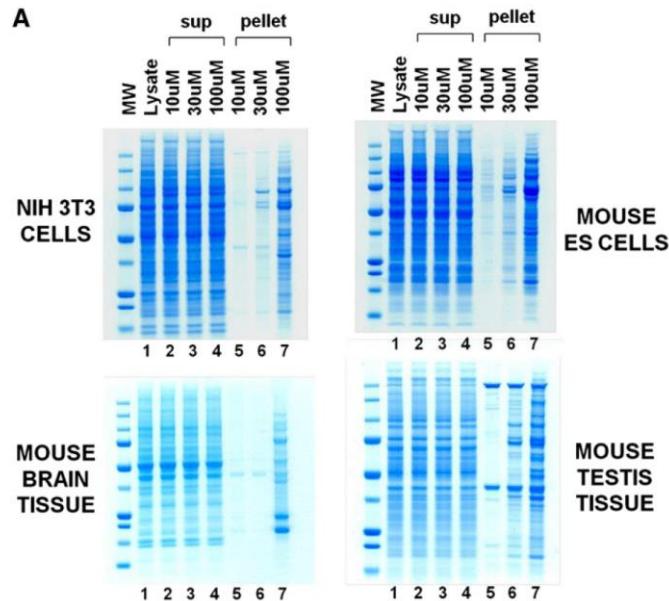
Cell-free Formation of RNA Granules: Low Complexity Sequence Domains Form Dynamic Fibers within Hydrogels

Masato Kato,¹ Tina W. Han,¹ Shanghai Xie,¹ Kevin Shi,¹ Xinlin Du,¹ Leeju C. Wu,¹ Hamid Mirzaei,¹ Elizabeth J. Goldsmith,¹ Jamie Longgood,¹ Jimin Pei,^{1,3} Nick V. Grishin,^{1,3} Douglas E. Frantz,⁴ Jay W. Schneider,² She Chen,⁵ Lin Li,⁵ Michael R. Sawaya,⁶ David Eisenberg,⁶ Robert Tycko,⁷ and Steven L. McKnight^{1,*}

Eukaryotic cells contain assemblies of RNAs and proteins termed RNA granules. Many proteins within RNA-binding domains and LCDs. Exposure of cell lysates to a biotinylated isoxazole (b-isox) precipitated hundreds of RNA-binding proteins with significant overlap to the constituents of RNA granules. LCDs are both necessary and sufficient for b-isox-mediated aggregation and can undergo a concentration-dependent phase transition to a hydrogel-like state. X-ray diffraction and EM showed that hydrogels are composed amyloid-like (but highly dynamic) fibers.

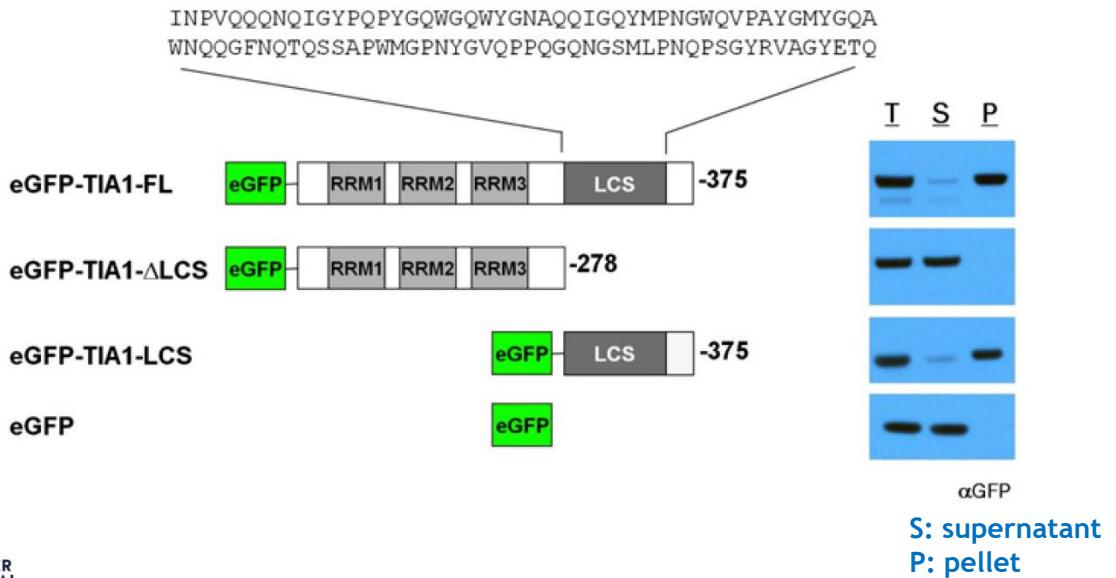
Cellular extracts treated with b-isox

A high-throughput drug screen for chemicals that promote mouse embryonic stem (ES) cell differentiation toward cardiomyocytes lead to the discovery of 5-aryl-isoxazole-3-carboxyamide. Neuronal progenitors further progressed to morphologically mature neurons. The biotinylated derivative (b-isox) was used to find its targets: it selectively precipitated proteins.

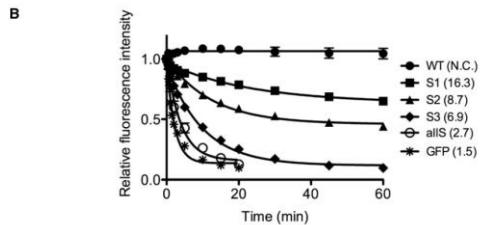
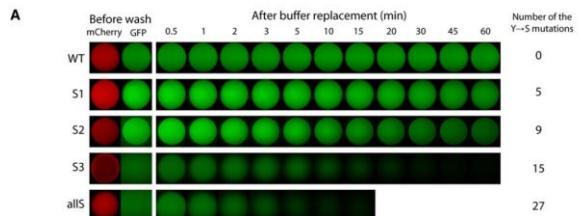
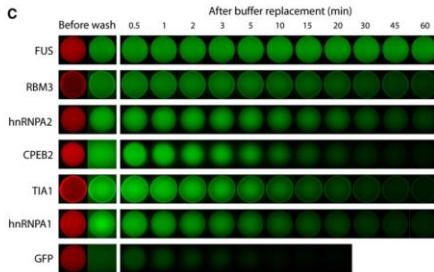
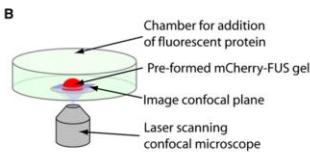


Enrichment for RNA granule proteins

Of 106 proteins found, 53 are among canonical RNA granule proteins (FUS, RBM3, hnRNPA2, CPEB2, TIA1, hnRNPA1, Atxn2, Ddx1, G3BP1, Npm1, Tdp-43, Caprin1). LCD drives b-isox-mediated precipitation.



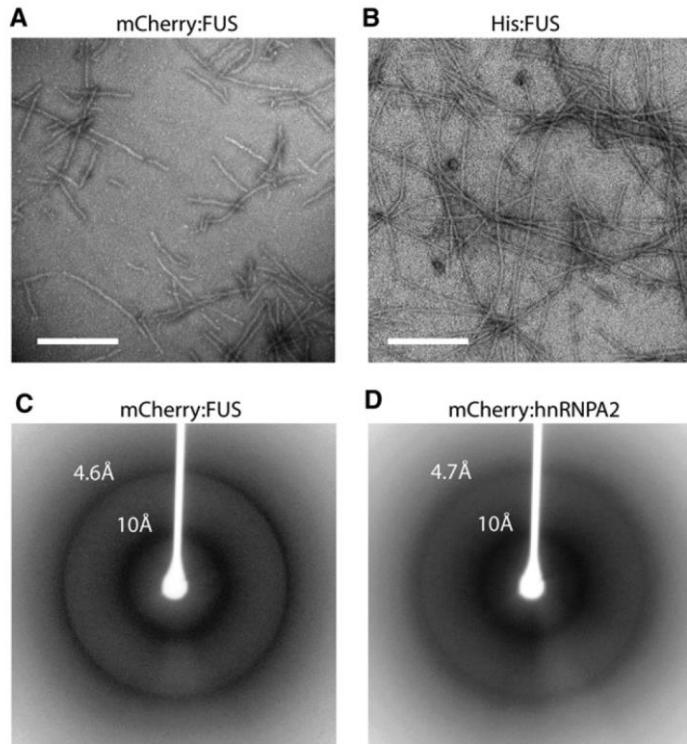
LCDs undergo gelation (without b-isox)



FUS LCD hydrogel
Retention of other GFP-LCDs

FUS LCD hydrogel, GFP-FUS LCD retention
Mutation of Tyr residues

b-isoX causes reversible amyloid formation



hydrogel droplets visualized by TEM

Kato... McKnight (2012) *Cell* 149: 753

2) An outline of LLPS and emergent functions

REVIEW

CELLULAR BIOPHYSICS

Liquid phase condensation in cell physiology and disease

Yongdae Shin and Clifford P. Brangwynne*

Shin, Brangwynne (2017) *Science* 357: eaaf4382



Annual Review of Physical Chemistry

Biomolecular Phase Separation:
From Molecular Driving Forces
to Macroscopic Properties

Gregory L. Dignon,^{1,2} Robert B. Best,³
and Jeetain Mittal¹

REVIEWS



Biomolecular condensates at the
nexus of cellular stress, protein
aggregation disease and ageing

Simon Alberti¹ and Anthony A. Hyman²

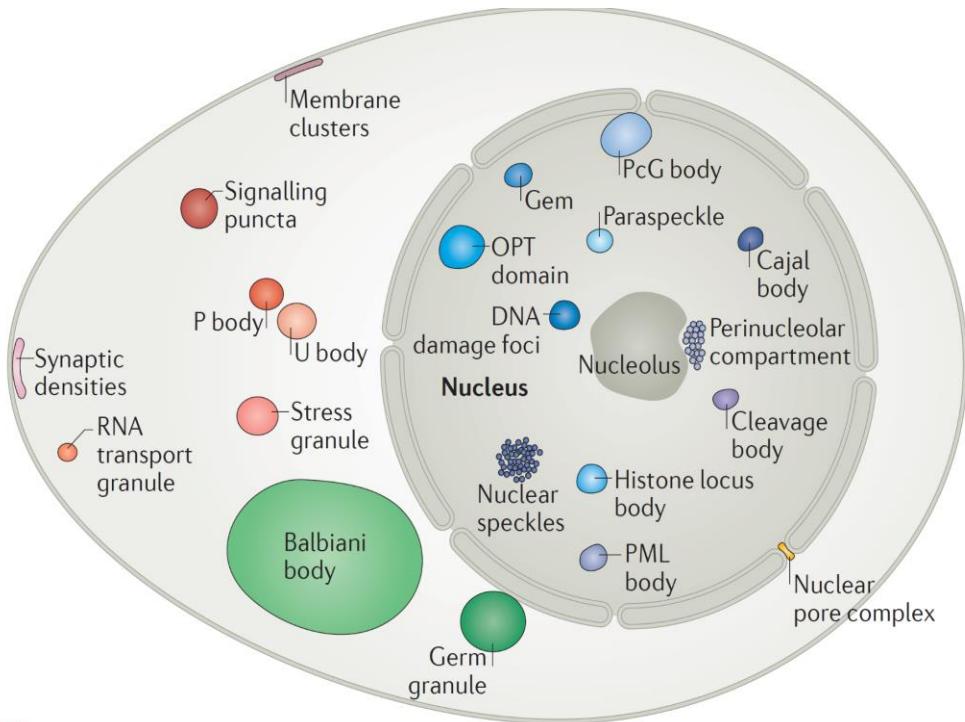
Alberti, Hyman (2021) *Nat. Rev. MCB* 22: 196



PhasAGE

Membraneless organelles

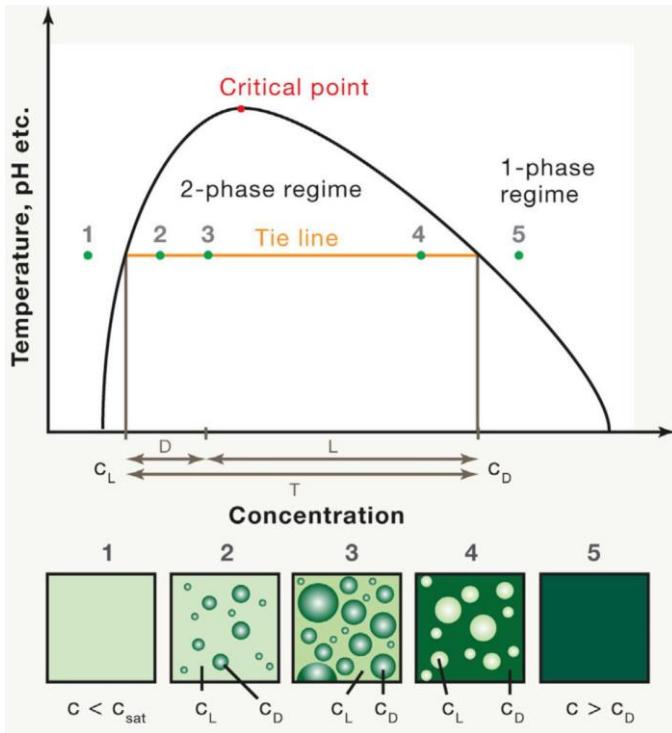
- biomolecular condensates, RNP bodies, LLPS -



Their mechanism of formation

- spontaneous demixing, LLPS -

- 1) phase diagram
- 2) binodal/coexistence line
- 3) saturation concentration (C_{sat})



Polymer physics (thermodynamics) of LLPS

- e.g. Flory-Huggins formalism -

Free energy of mixing per lattice site (Φ - volume fraction)

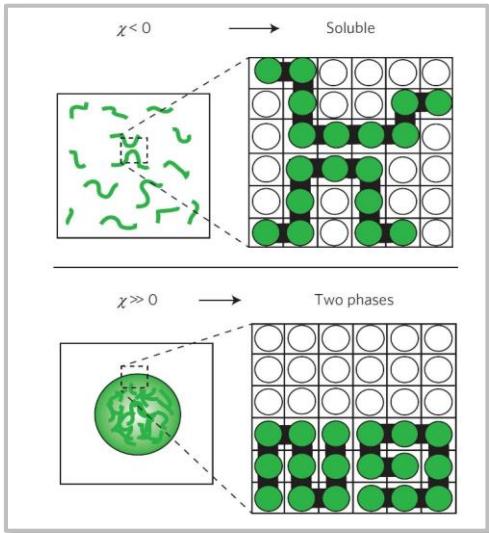
$$\frac{F}{k_B T} = \frac{\phi}{N} \ln \phi + (1-\phi) \ln(1-\phi) + \chi \phi(1-\phi)$$

Entropy Enthalpy

Chain-chain vs. chain-solvent interaction
(χ - Flory prmt.)

$$\chi = \frac{z}{k_B T} \left[u_{ps} - \frac{1}{2} (u_{pp} + u_{ss}) \right]$$

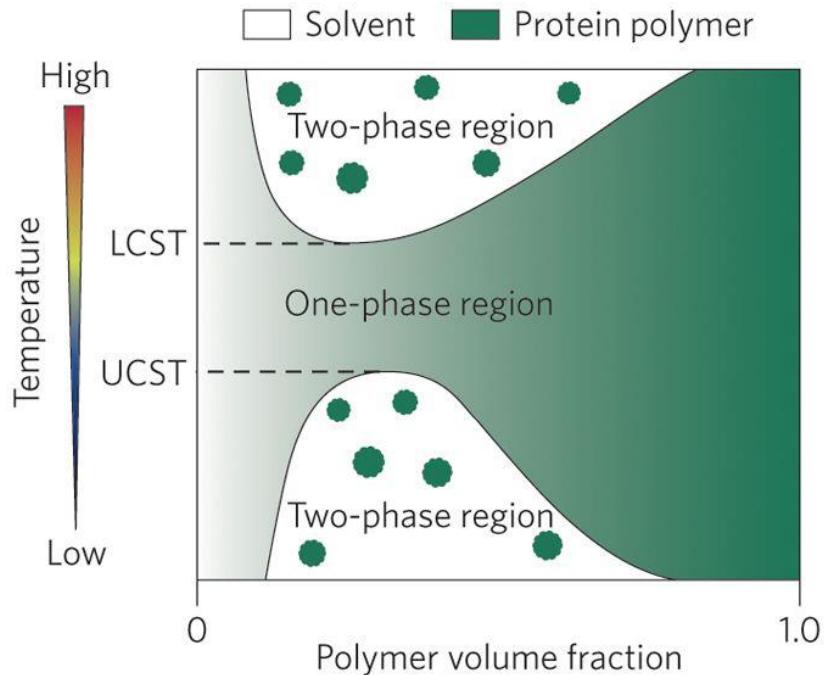
Homotypic, homopolymer



Homogeneous homopolymers

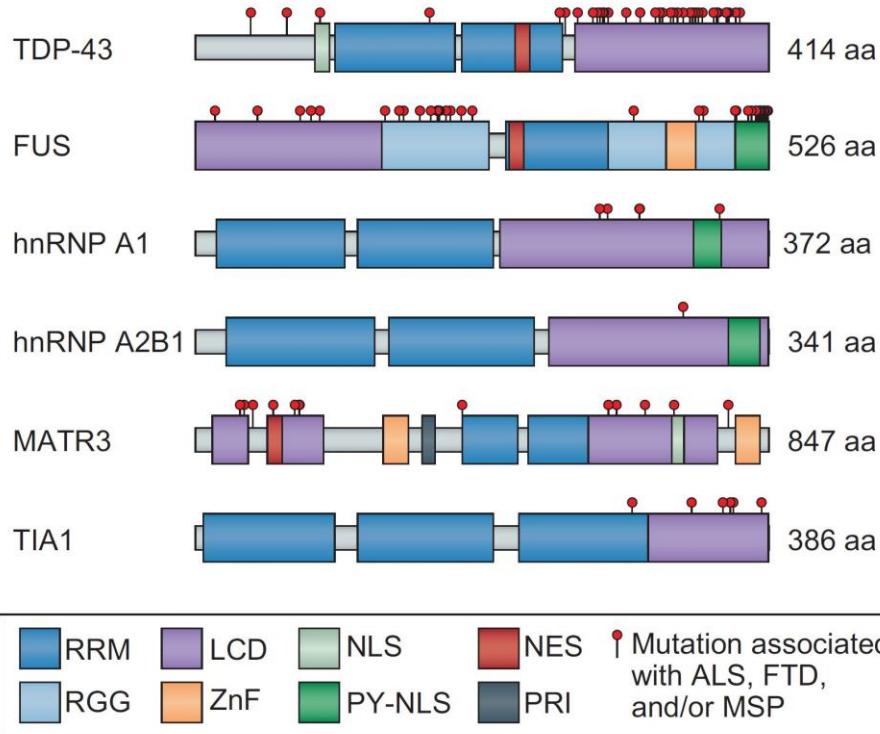
Sometimes opposite behavior

Temperature dependence of attractive term



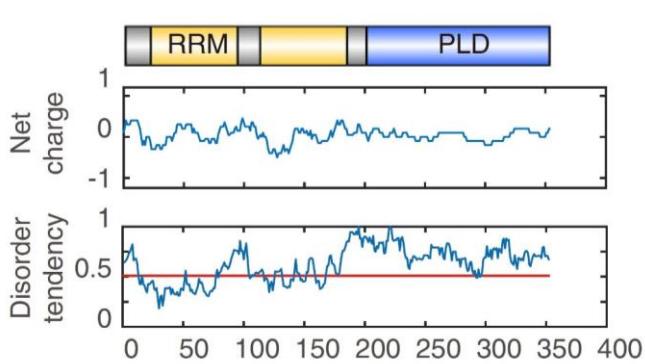
Structural disorder and RNA binding in LLPS proteins

- LCD: low-complexity IDR -

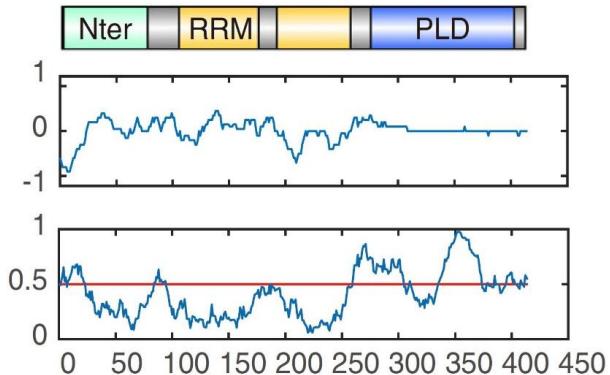


Structural disorder in LLPS proteins

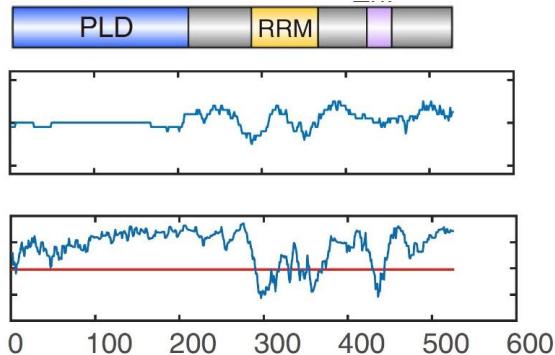
hnRNP A2B1



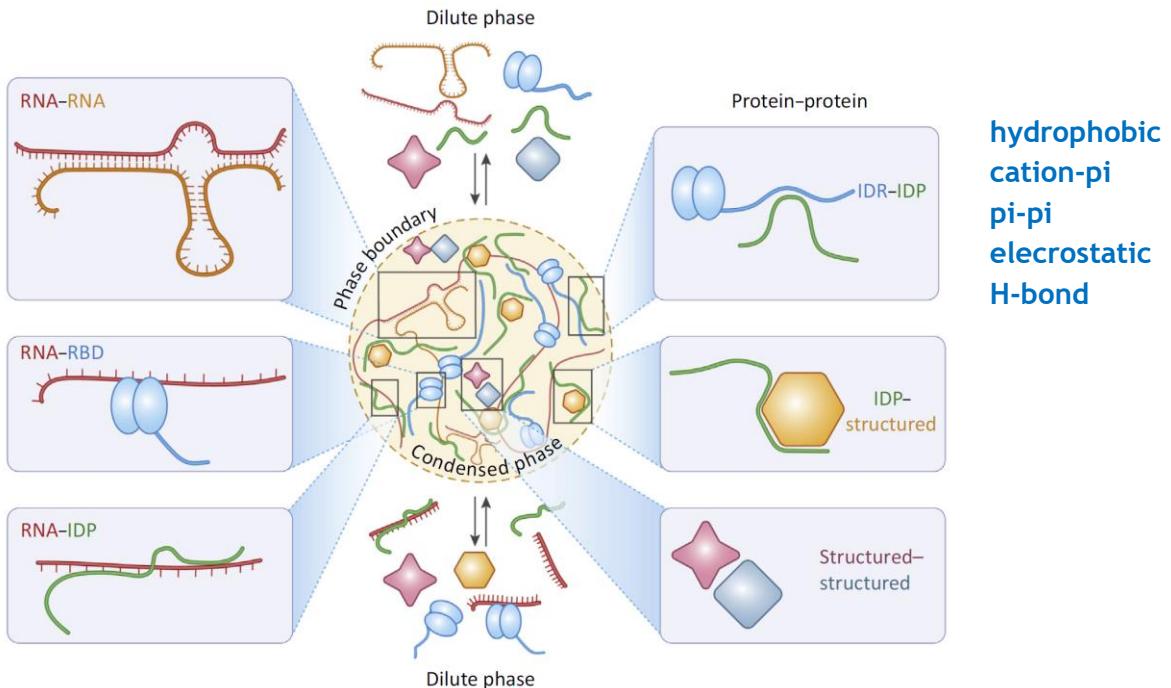
TDP-43



FUS

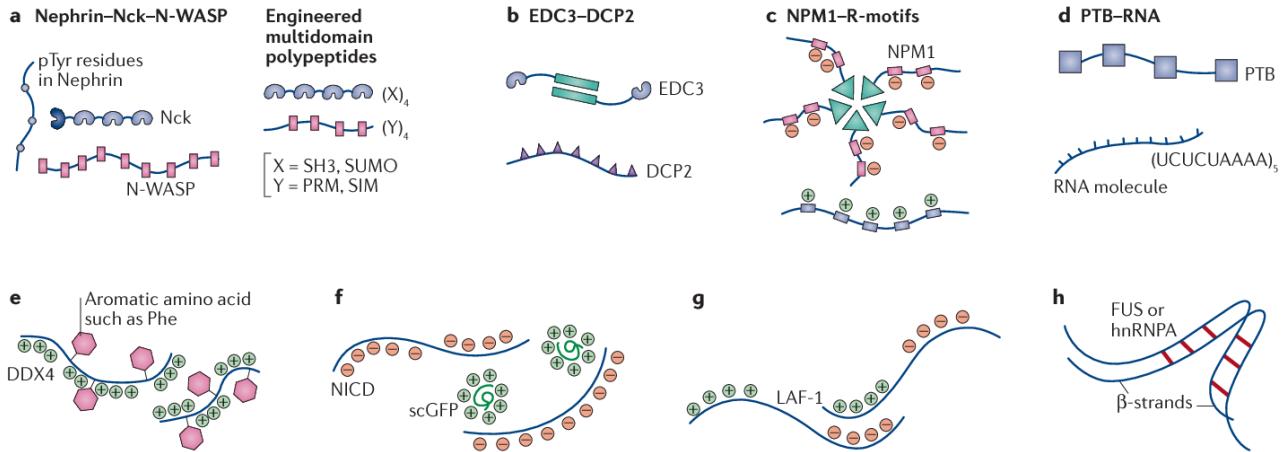


Different types of interactions driving LLPS



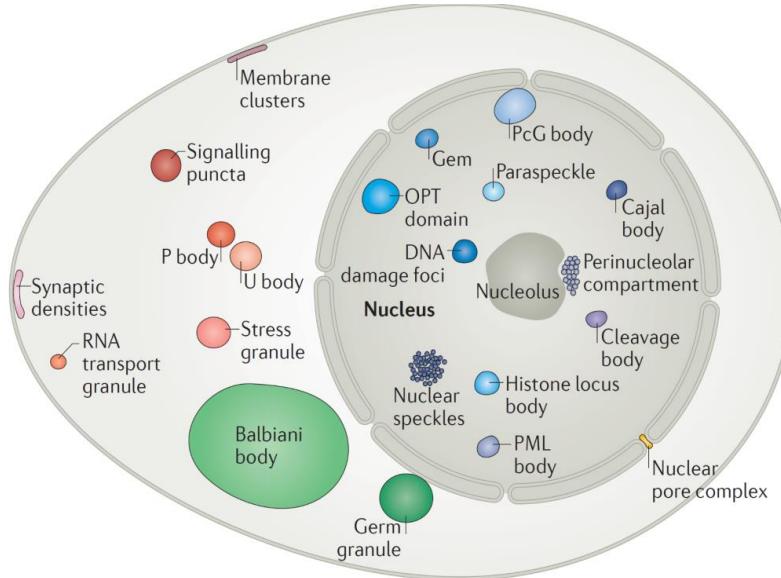
Features: (i) dynamics, (ii) strength, (iii) specificity, (iv) stoichiometry, (v) structure

Multivalency is basic to LLPS

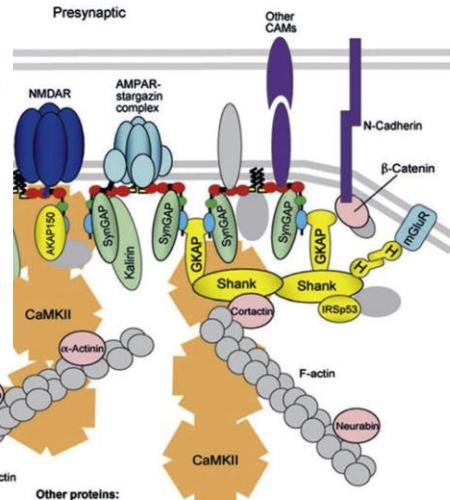


- b) Enhancer of mRNA-decapping protein 3 (EDC3) decapping enzyme subunit 2 (DCP2)
d) polypyrimidine tract (RNA) binding protein (PTB)
f) nephrin intracellular domain (NICD) and supercharged GFP
g) P-granule LAF-1

Not every cellular assembly is LLPS though



Banani et al. (2017) *NRMCB*

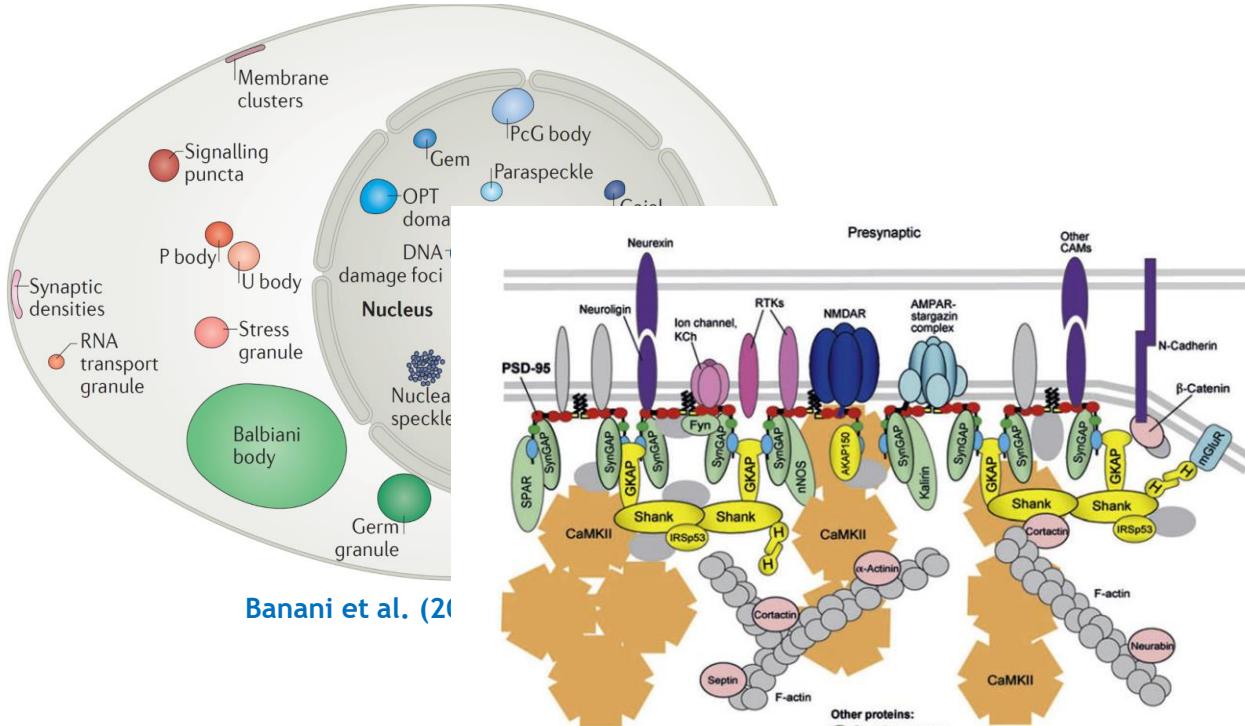


PSD-95:
● PDZ domain
● SH3 domain
● GUK domain

Other proteins:
● Signaling enzymes
● Scaffold proteins
● Actin-binding proteins
● Adhesion molecules
● Unnamed proteins

Pancsa (2019) *BBA*

Not every cellular assembly is LLPS though



- 1) Assembly by PS?
- 2) Cooperativity
- 3) Structure (beyond 1st neighbor)

Pancsa (2019) BBA

Functional consequences of LLPS

- 1) an emergent property, might not manifest at the level of individual proteins
- 2) means two different things (remember Gene Ontology MF, BP and CC)
 - the way an LLPS droplet functions (MF)
 - what it does in the cell (BP)



VIB-VUB CENTER
FOR STRUCTURAL
BIOLOGY

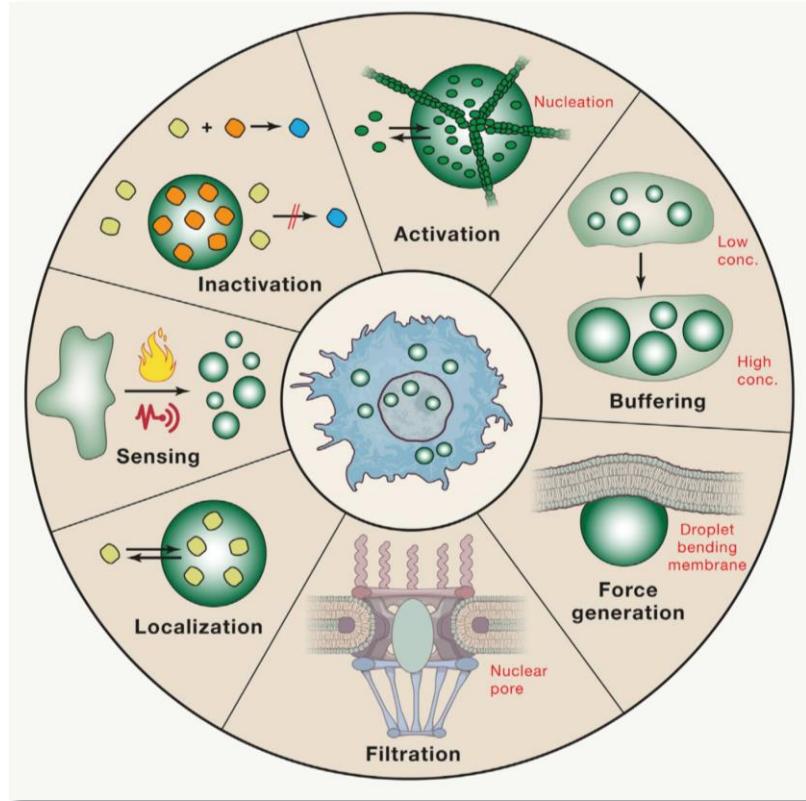


VRIJE
UNIVERSITEIT
BRUSSEL



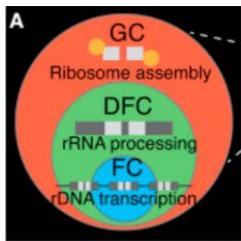
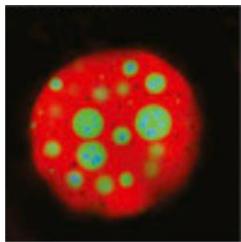
Functional consequences of LLPS

(Molecular Function ontology)



Function of nucleolus

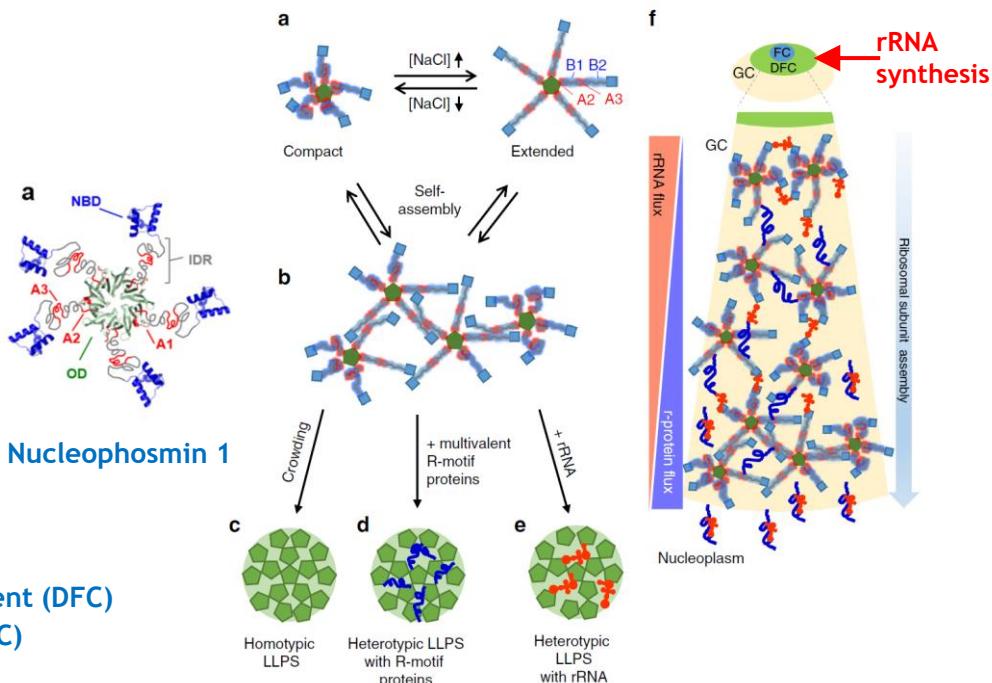
- ribosome biogenesis -



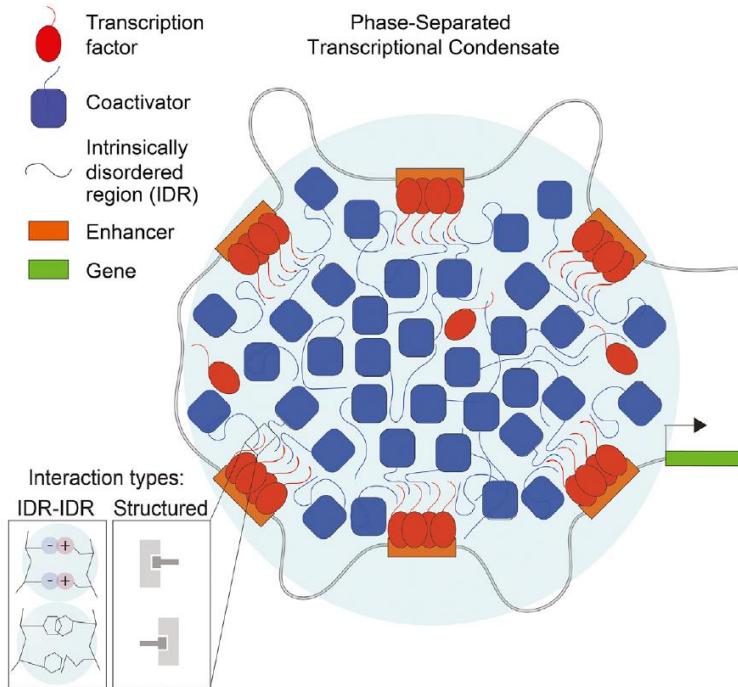
fibrillar center (FC)

dense fibrillar component (DFC)

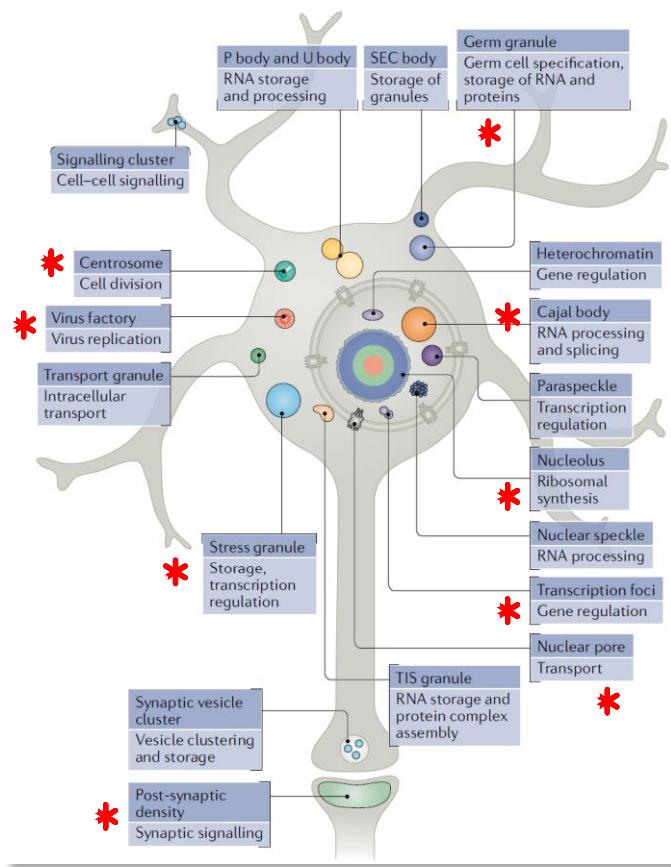
granular component (GC)



Function of super-enhancers (SEs) in transcription regulation



Functional consequences of LLPS (Biological Process ontology)



Alberti, Hyman (2019) *Nat. Rev. MCB* 22:196

3) LLPS and disease

LLPS may lead to disease

- ALS/FTD, Lou Gehrig's disease



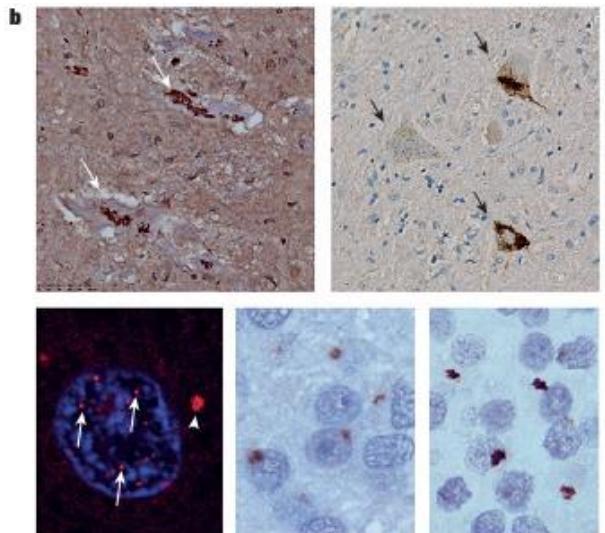
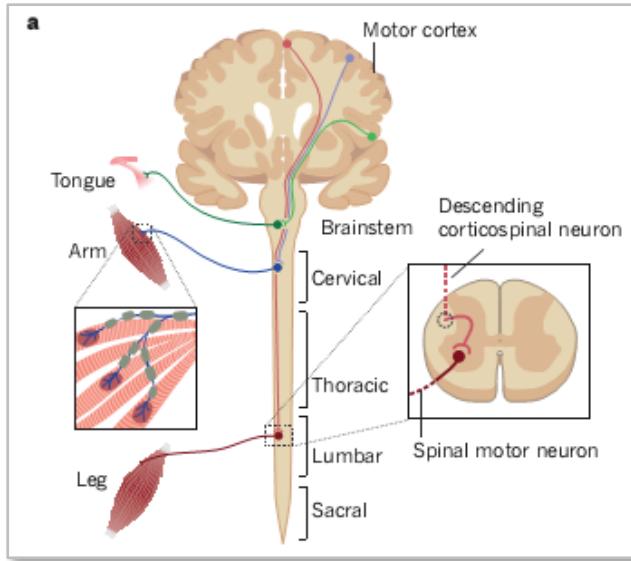
Lou Gehrig



Stephen Hawking

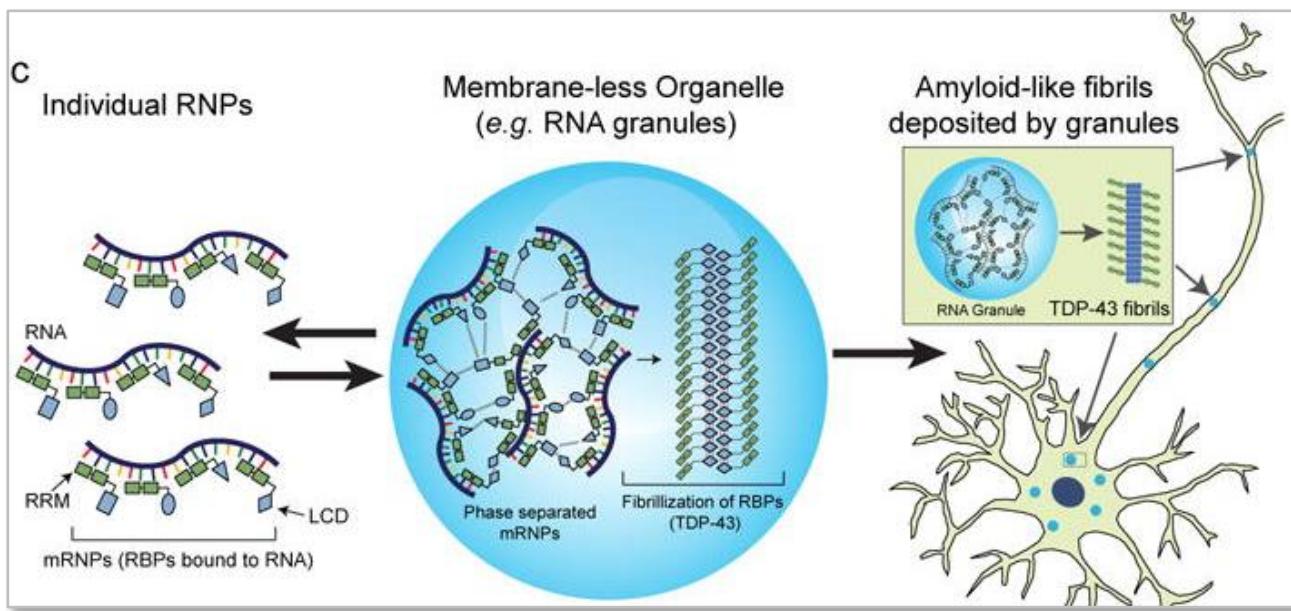
- Progressive loss (atrophy) of muscles
- Survival after first symptoms: 2 - 5y
- No cure

ALS is motor neuron disease



LLPS may lead to disease

- stress granules in ALS (?) -

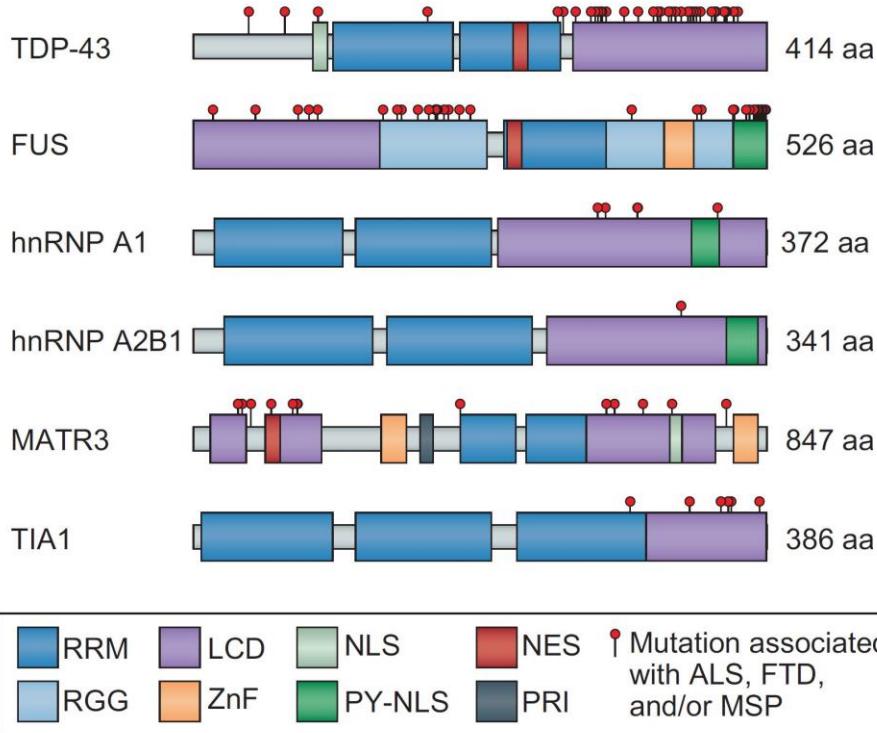


Many neurodegenerative proteins phase separate

Table 1. Summary of Neurodegenerative Diseases and LLPS/MLOs

Protein	Associated diseases	Evidence for disease association	Granule types	Evidence for <i>in vitro</i> LLPS	Evidence for <i>in vivo</i> LLPS or granules
FUS	ALS/FTD	Point mutations (NLS and others)	Stress	[18,44,89]	[44,89]
TDP-43	ALS/FTD, AD	Point mutations, truncations	Stress, transport	[19,90]	[47,48,90,105]
hnRNPA1	ALS/FTD/multisystem proteinopathy (MSP), MS	Point mutations in ALS/FTD/MSP, mislocalized in MS [64]	Stress, transport	[6]	[6]
hnRNPA2	ALS/FTD/MSP	Point mutations in ALS/FTD/MSP	Stress, transport	[17]	[106]
C9ORF72	ALS/FTD, SCA	G4C2 expansion in ALS/FTD, SCA [72]	Stress	[45,46,52]	[45,46,51,52]
UBQLN2	ALS/FTD	Point mutations in ALS/FTD	Stress	[54,107]	[54]
TIA1	ALS/FTD, AD, MS, SMA	Point mutations in ALS/FTD; mislocalized in AD [61], MS [64], and SMA [77]	Stress	[55,56]	[55,56]
Profilin	ALS/FTD, HD	Point mutations in ALS/FTD; modifies HD aggregation <i>in vitro</i> [66]	Stress		[57]
Ataxin-2	ALS/FTD, SCA	PolyQ expansion, 32 or more in SCA, 29–32 is a risk factor for ALS [73]	Stress, transport	[38]	[38]
Tau	AD	Mutated in AD	Stress	[29,60,108]	[29]
DJ-1	PD	Mutated in PD	Stress (P bodies)		[67]
Huntingtin	HD	PolyQ expansion in HD	Stress	[65,66]	[65]
Staufen-1	SCA	Modifies SCA; recruited to aggregates and increased expression [74]	Transport		[74]
SMN	SMA	Mutated/reduced levels in SMA	Stress, transport		[77,78]
FMRP	Fragile X, FXTAS	5'-UTR repeat expansion in fragile X, FXTAS	Stress, transport	[22,80]	[8,13]

And carry disease-specific mutations

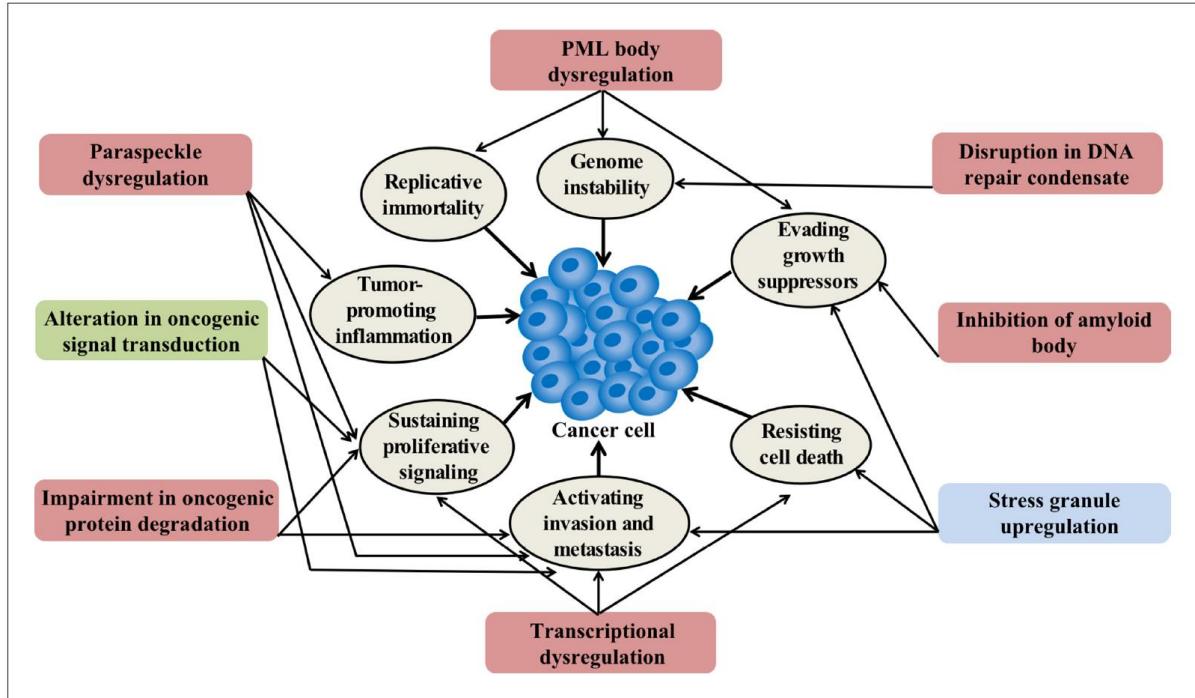


Break (?)



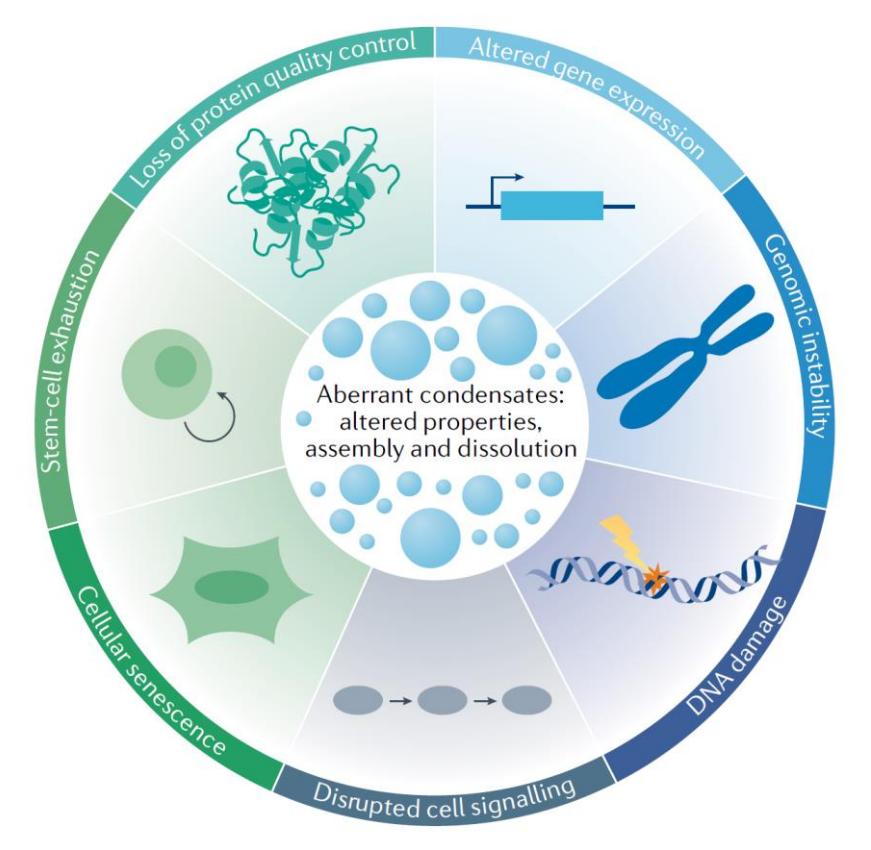
LLPS might also be involved in cancer

- “hallmarks of cancer” -



LLPS might also be involved in ageing (!)

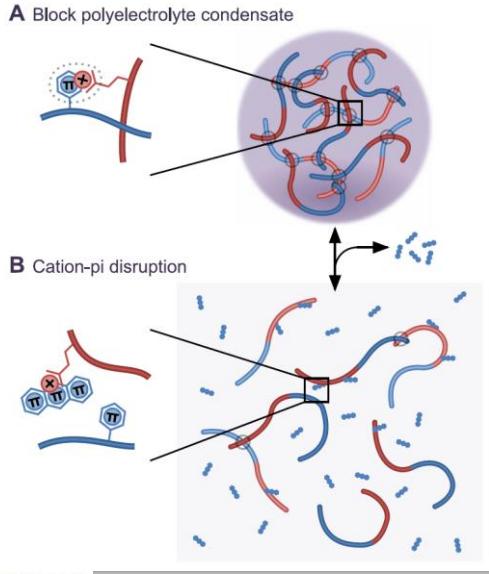
- “hallmarks of ageing” associated with ageing -



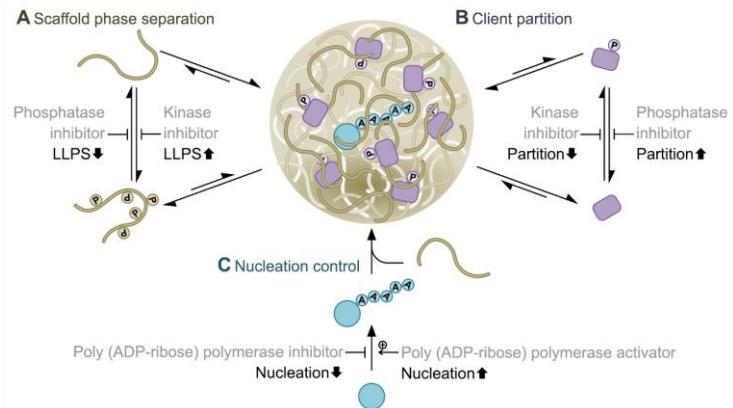
Alberti, Hyman (2021) *Nat. Rev. MCB*

Targeting LLPS?

Targeting LLPS and aggregation



Targeting LLPS via regulation



“Competitive landscape”



Tony Hyman, Rick Young
\$ 60M Series A, \$ 77M Series B
Merck: \$ 305M (HIV)
Bayer: \$ 77M (cardiovascular, gynecology)
Pfizer: \$ 239M (myotonic dystrophy type 1)
oncology, neuro, metabolic, immunology



Cliff Brangwynne
\$50M Series A
neurodegeneration, cancer



Michael Rosen
\$81M Series A
ALS



David Weitz, Tuomas Knowles, Peter St George-Hyslop
\$9M Seed



PhasAGE

4) Experimental techniques to study LLPS





A User's Guide for Phase Separation Assays with Purified Proteins

Alberti... (2018) *J. Mol. Biol.* 430: 4806

Simon Alberti[†], Shambaditya Saha[†], Jeffrey B. Woodruff, Titus M. Franzmann, Jie Wang and Anthony A. Hyman



Methods for Physical Characterization of Phase-Separated Bodies and Membrane-less Organelles

Mitre... (2018) *J. Mol. Biol.* 430: 4773

Diana M. Mitrea¹, Bappaditya Chandra^{1,†}, Mylene C. Ferrolino^{1,†}, Eric B. Gibbs^{1,†}, Michele Tolbert^{1,†}, Michael R. White^{1,†} and Richard W. Kriwacki^{1,2}

Leading Edge
Primer

Cell

Considerations and Challenges in Studying Liquid-Liquid Phase Separation and Biomolecular Condensates

Simon Alberti,^{1,2,*} Amy Gladfelter,^{3,4,*} and Tanja Mittag^{5,*}

Alberti... (2019) *Cell* 176: 419



PhasAGE

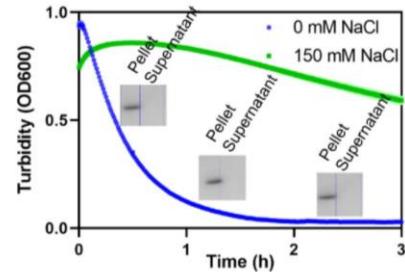
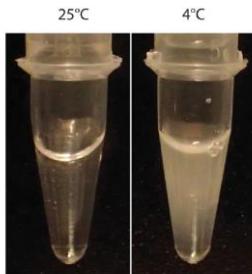
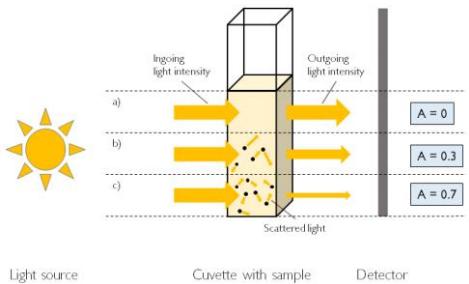
Protein purification

LLPS proteins are often aggregation prone, frequently found in inclusions in neurodegenerative diseases

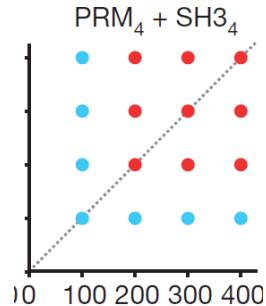
- 1) Often require a trick to keep them in solution: (i) solubility tag (MBP, GST, GFP), (ii) non-native/denaturing conditions (e.g. urea, pH away from pI), (iii) additives disfavoring LLPS (high salt, high Arg)
- 2) Purification not only from *E. coli* (insect cell, mammalian cells)
- 3) Advised to be aliquoted, only used once (no repeated freeze-thaw cycles)

Turbidity

- OD₃₄₀, OD₄₀₀, OD₆₀₀ -



hnRNP A2 LCD LLPS



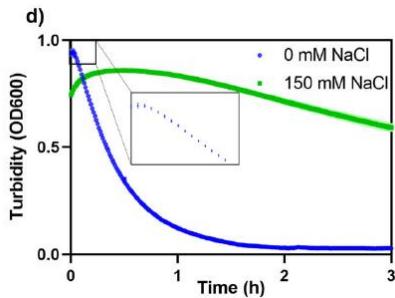
Neph. - Nck - WASP (Rosen 2012)

- 1) fast
- 2) easy control (of conditions, T)
- 3) appropriate for determining C_{sat}
- 4) can be made highly parallel (phase diagram)
- 5) but: convolution of number and size of scattering particles

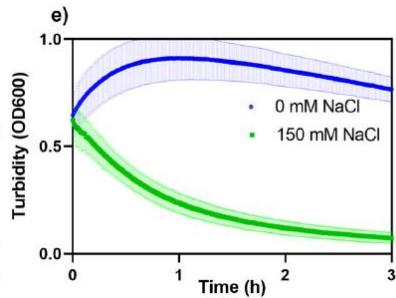
Turbidity

- hnRNPA2 LCD LLPS initiated by: -

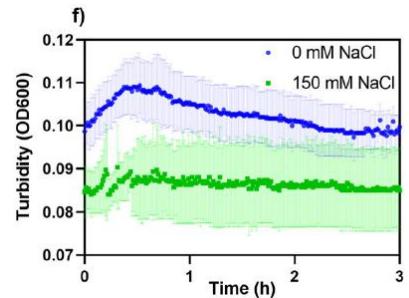
pH jump: 11.0 → 7.5



Urea: 8M → 80mM



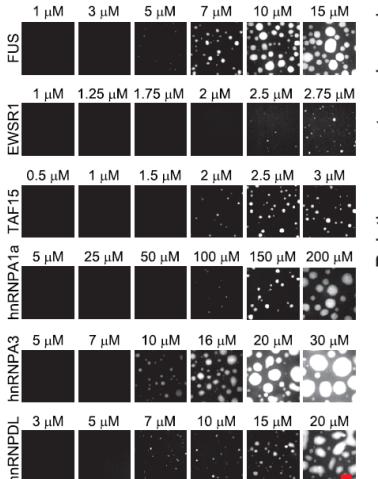
Cleavage of MBP tag



Microscopy

A Molecular Grammar Governing the Driving Forces for Phase Separation of Prion-like RNA Binding Proteins

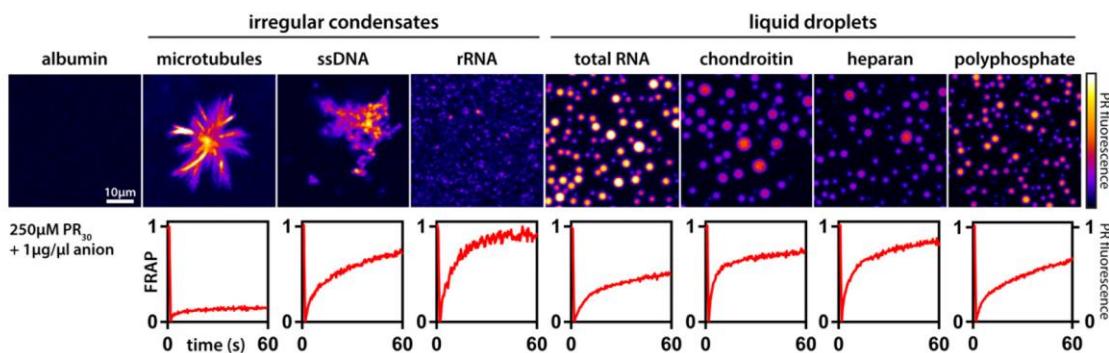
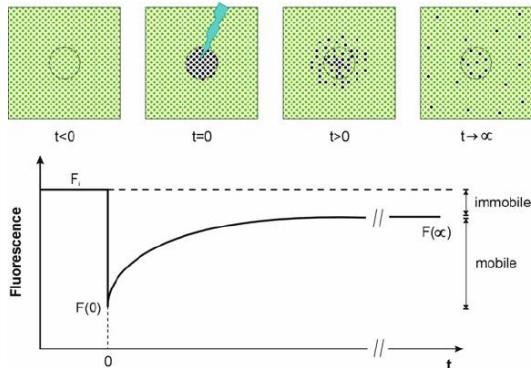
Jie Wang,¹ Jeong-Mo Choi,² Alex S. Holehouse,² Hyun O. Lee,¹ Xiaojie Zhang,¹ Marcus Jahnel,¹ Shovamayee Maharana,¹ Régis Lemaitre,¹ Andrei Pozniakovsky,¹ David Drechsel,³ Ina Poser,¹ Rohit V. Pappu,² Simon Alberti,^{1,*} and Anthony A. Hyman^{1,4,*}



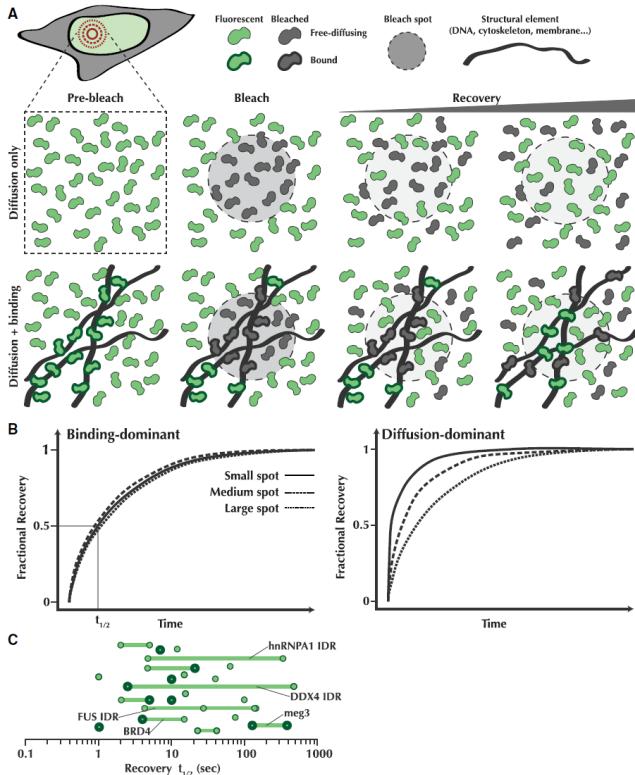
Here: FUS family of proteins

- 1) direct visualization
- 2) confirmation of spherical shape
- 3) shape change, fusion, colocalization
- 4) targeted FRAP
- 5) cellular applications

Fluorescence recovery after photobleaching (FRAP)

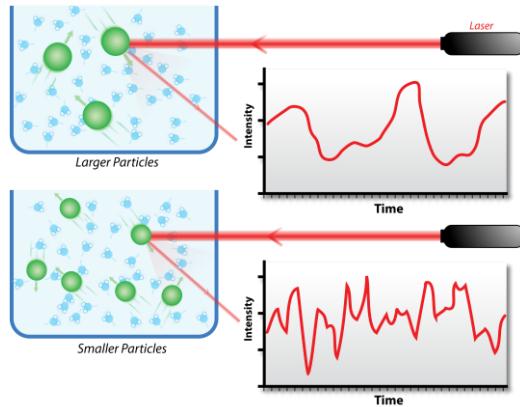


Interpretation of FRAP: model dependent

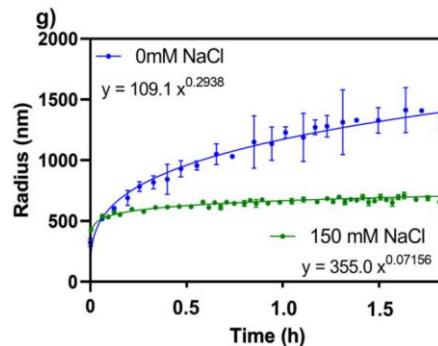
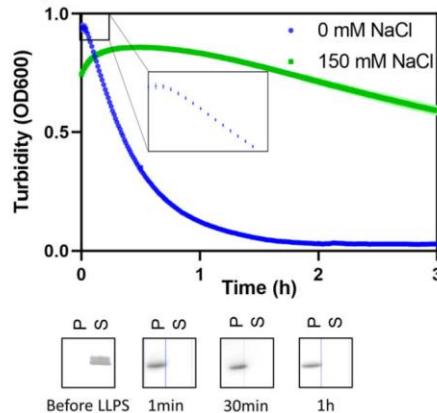


Dynamic light scattering (DLS)

- measuring size distribution profile of droplets -



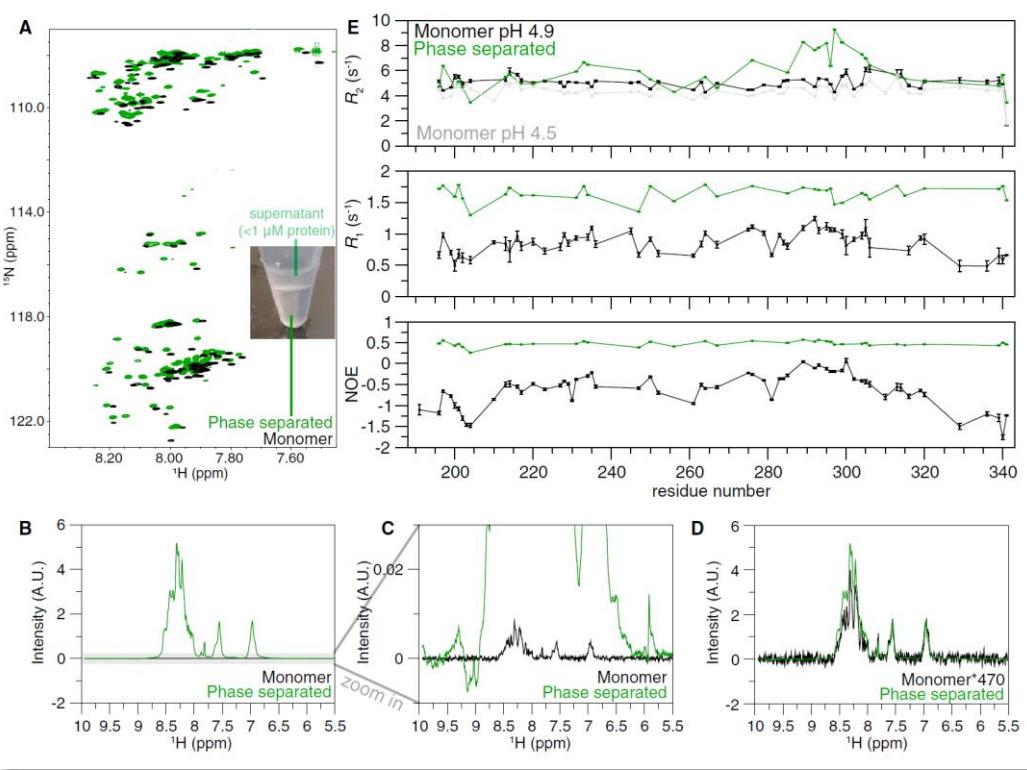
Analyzing temporal fluctuations
(autocorrelation) of light intensity



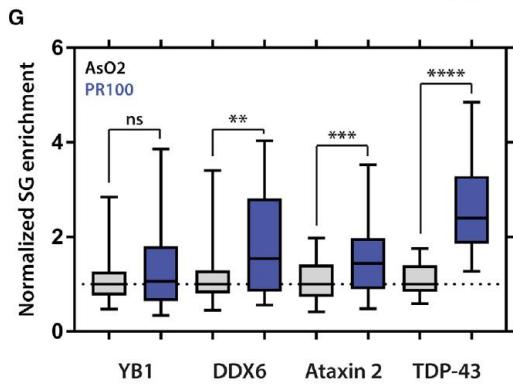
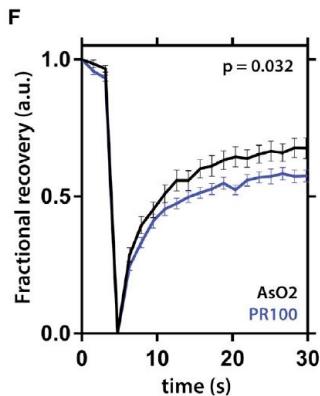
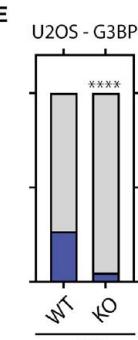
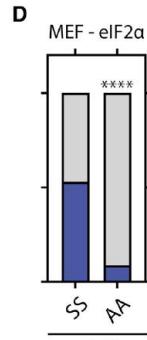
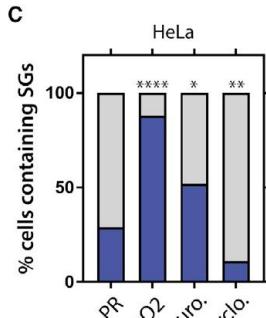
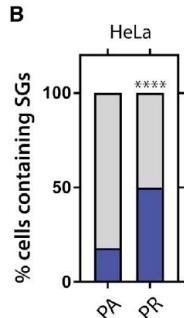
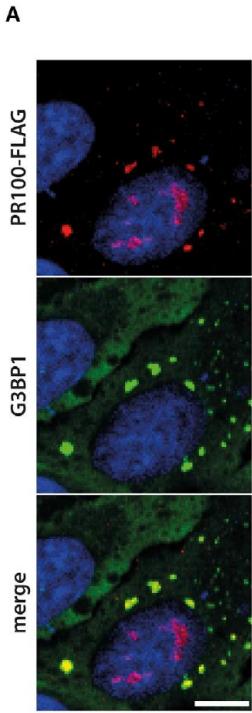
hnRNPA2 LCD LLPS

Nuclear magnetic resonance (NMR)

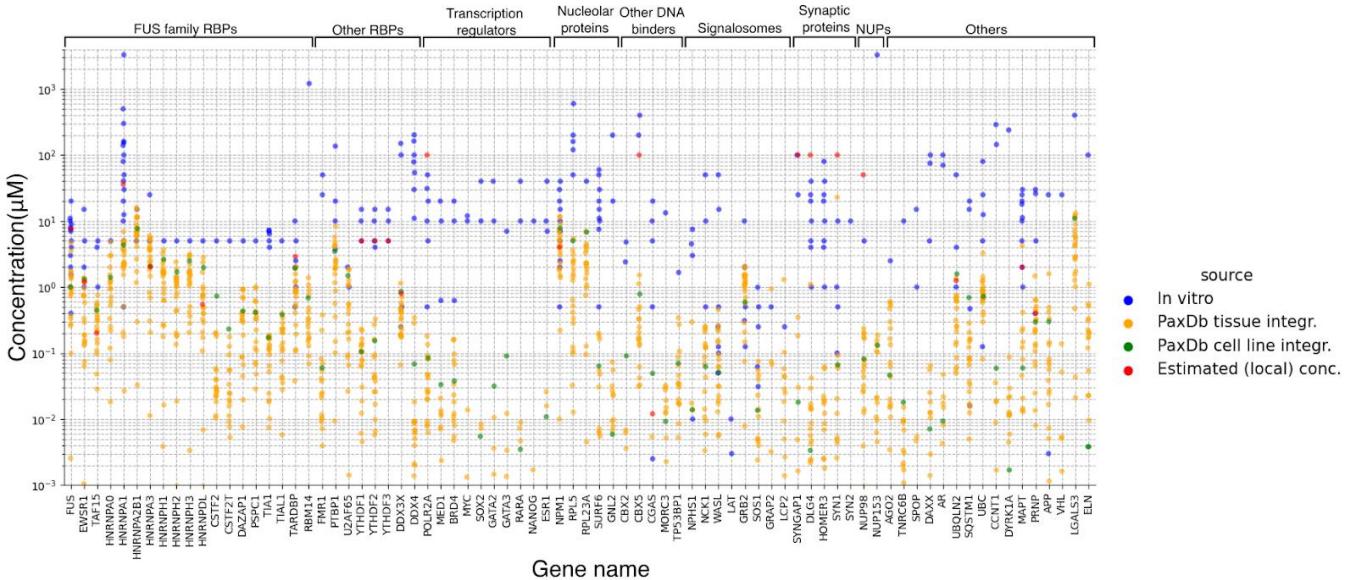
- hnRNPA2 LCD remains disordered in the droplet -



Cellular: overexpression - better: driver/modifier -



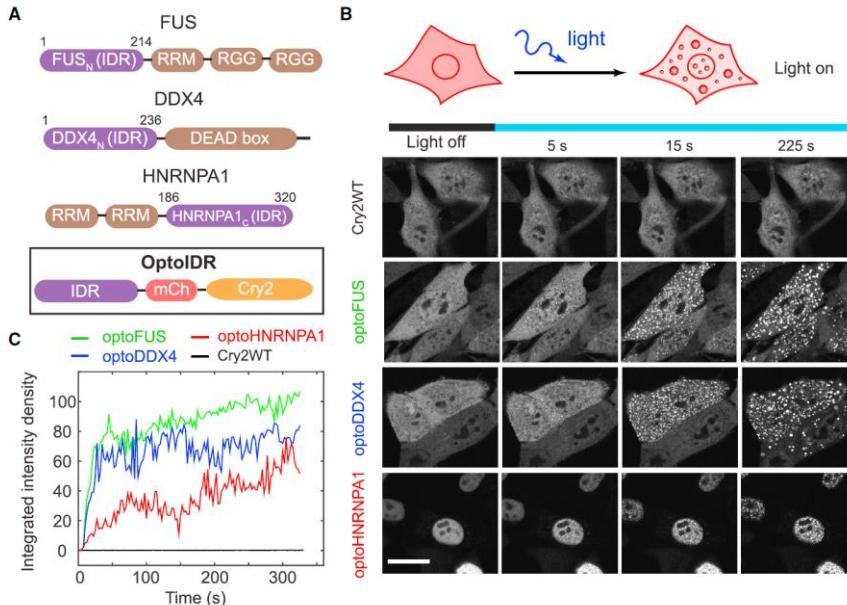
Issue with concentrations



Cellular: optogenetics

- Cry2 (cryptochrome) based -

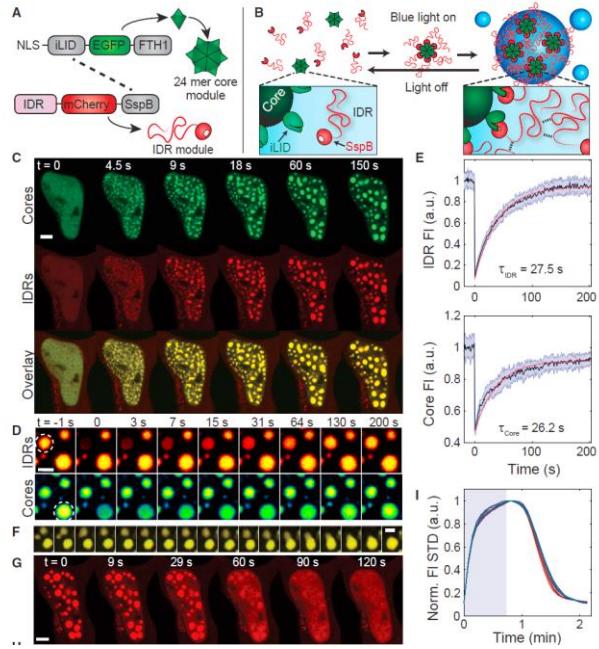
Photolyase homology region (PHR) of *A. thaliana* Cry2, a light-sensitive protein that dimerizes upon blue light exposure



Cellular: optogenetics

- Corelets -

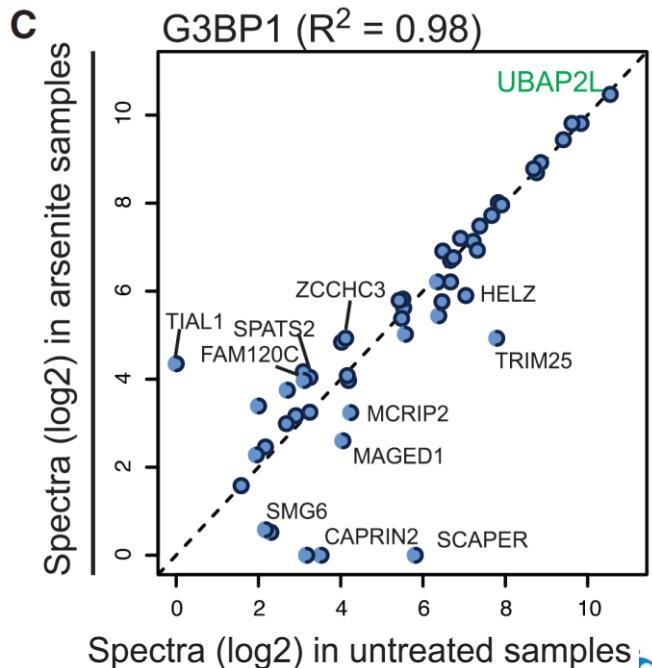
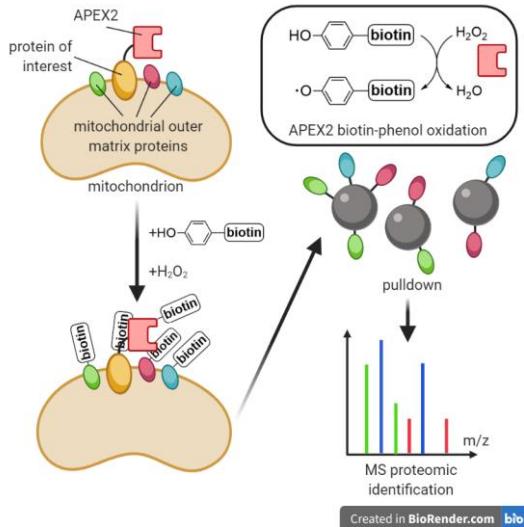
24-mer ferritin (FTH1) derivatized by photoactivatable iLID domain, plus its cognate partner, SspB



Shin... Brangwynne (2017) *Cell* 168: 159

Cellular: proximity tagging

Protein of interest carries an enzyme fused (biotin ligase, BirA* or ascorbate peroxidase, APEX) that can biotinylate nearby proteins in the cell.



5) Databases and bioinformatics tools



Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in
Structural Biology

First-generation predictors of biological protein phase separation

Robert M Vernon¹ and Julie D Forman-Kay^{1,2}



Vernon, Forman-Kay (2019) *COSB* 58: 88

Computational resources for identifying and describing proteins driving liquid–liquid phase separation

Rita Pancsa, Wim Vranken and Bálint Mészáros^{ID}

Pancsa... (2021) *Brief. Bioinfo.* 1-20

LLPS databases

- <https://phasepro.elte.hu> -

Home Browse/Search About/Help Candidates Statistics Download Annotate

Welcome to PhaSePro!

PhaSePro is the comprehensive database of proteins driving liquid-liquid phase separation (LLPS) in living cells. LLPS is a molecular process employed by all living organisms to form membraneless organelles, mediating crucial cellular functions. PhaSePro is manually curated, it is solely based on experimentally verified cases of LLPS, integrating a wide range of information on the biophysical driving forces, biological function and regulation of these molecular systems.

[Learn more »](#)



Search for gene names, full or partial common/UniProt protein names, or UniProt accessions.

[Example 1](#) [Example 2](#)

Getting started

To get an introduction to the structure and use of PhaSePro, you can visit the selected examples ([FUS](#) and [TDP-43](#)) by clicking the buttons above, or read the About/Help pages by clicking below.

[View details »](#)

Explore

You can start searching the database by entering keywords in the above field, or by browsing the available entries by clicking below.

[Browse/Search entries »](#)

Annotate

Help us expand the knowledge about proteins involved in liquid-liquid phase separation by submitting new entries into PhaSePro.

[View details »](#)



R. Pancsa et al.



Manually curated for LLPS drivers that has both *in vitro* and *in vivo* relevance

Farahi et al. (2021) *Int J Mol Sci.* 22: 3017



PhasAGE

LLPS databases

- <http://bio-comp.org.cn/llpsdb/home.aspx> -

The screenshot shows the homepage of the LLPSDB database. At the top, there's a navigation bar with links for Home, Browse, Search, Submit, Statistics, and Download. A decorative header image features a stylized protein structure and the text "LLPSDB A database of proteins undergoing liquid-liquid phase separation *in vitro*". The main content area has a teal background. It starts with a "Welcome to LLPSDB!" section containing a paragraph about LLPS and its biological significance. Below this is another paragraph describing the database's content and how it integrates experimental data from various sources. To the right, there are two boxes: one for "News" (mentioning a correction made and a PMID added) and one for "Database Linked" (listing Uniprot, PubMed, MobiDB, UniProt, OMIM, IDEAL, FuzzDB, AmylPro, and FuzzDB). Another box on the right is titled "Contact Us" and provides email information and the address: Beijing, P.R.China, 100049.

Welcome to LLPSDB !

Liquid-liquid phase separation (LLPS) of proteins has been discovered to underlie the compartmentalization of cells, through the formation of liquid biological condensates including membraneless organelles (MLOs), signaling puncta and so on. LLPS is associated with numerous biological processes such as RNA metabolism, gene regulation and signal transduction. However, the fundamental mechanism of protein LLPS still remains to be elucidated. It is important to systematically analyze all the available experimental data for a better understanding of LLPS. To this end, through extensive literature curation, we summarized the proteins and corresponding experimental conditions under which their phase separation tendencies have been detected *in vitro*, and deposited them in this database.

LLPSDB contains LLPS related proteins together with the corresponding phase separation conditions validated by experiments. For each protein, the database provides various information, including the protein sequence, modifications on specific amino acids, its ability of coalescing with nucleic acids, biological function etc., as well as specific experimental conditions such as temperature, salt concentration, pH, crowding agent, detected techniques, phase behavior and so on. In addition, several related databases are linked from LLPSDB including Uniprot, MobiDB, UniProt, OMIM, IDEAL, AmylPro, FuzzDB and PubMed. All the data summarized in LLPSDB are available for users.

A pie chart illustrating the distribution of LLPS experiments. The data is as follows:

Components	Count	Percentage
One component	536	46%
Two components	501	43%
More components	128	11%

In vitro LLPS experiments as entries, containing both natural designed proteins (the role of protein is not defined).



LLPS databases

The image shows the homepage of PhaSepDB. The header features the logo "PhaSepDB" on the left and navigation links "Home", "Data Sources", "Search", "Help", and "Download" on the right. The main title "PhaSepDB" is prominently displayed in large white letters. Below it is the subtitle "The database of phase-separation related proteins." A search bar at the bottom contains fields for "GeneSymbol", "Input to get results...", and "All", with a "Search" button. The background is dark with faint, glowing circular shapes resembling protein molecules.

What is PhaSepDB?

PhosSepDB is a novel database that provides a collection of phase separation related proteins manually curated from publications into public database.

As of October 2019, this database includes 2957 eligible proteins; 2303(77.9%) of the proteins were localized in different organelles.

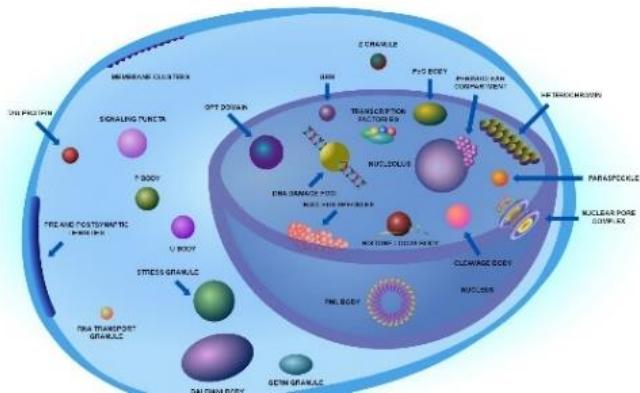
Clica a teclas ou faixas de quando em quando.

While 561(22%) of them were not associated with known parasites.

[Click here to bypass these greetings.](#)

In addition to curated proteins, PhoSepDB also provides the researchers with molecular signatures that may facilitate phase-separating related proteins identification for all human proteins.

[Click here to download human postures.](#)



Annotations based on MLO localization (literature evidence, UniProt annotations, and HTS localization experiment).



Farahi et al. (2021) *Int J Mol Sci.* 22: 3017



PhasAGE

LLPS databases

- <http://llps.biocuckoo.cn> -

The screenshot shows the homepage of the DrLLPS database. At the top, there's a navigation bar with links for HOME, BROWSE, SEARCH, DOWNLOAD, USER GUIDE, LINK, and CONTACT. To the left, there's a sidebar for 'PRODUCTS OF CUCKOO' with links for PTMS Predictor, Tools, and Databases. Below the sidebar is a world map with a grid overlay. In the center, there's a large text area about LLPS, followed by a detailed description of the database's scope and annotations. Below this are four cards: one showing a complex protein structure with the text '9,285 Curated Proteins'; another showing a cell with organelles and the text '40 Condensates'; a third showing a person and a cell with the text '164 Species'; and a fourth showing a database icon with the text '110 P'. At the bottom, there's a search bar labeled 'Simple Search'.

DrLLPS. Data resource of liquid-liquid phase separation Version 1.0

THE CUCKOO WORKGROUP

HOME BROWSE SEARCH DOWNLOAD USER GUIDE LINK CONTACT

PRODUCTS OF CUCKOO

- PTMS Predictor
- Tools
- Databases

Last update: Jun 13th, 2019

Phase separation, or liquid-liquid phase separation (**LLPS**), is a ubiquitous and important mechanism for the formation of membraneless condensates in cells. Living cells contain various types of membraneless condensates enriched with proteins and RNA molecules. These condensates, including stress granules (Millec et al., 2007), P granule (Brangwynne et al., 2009), spindle apparatus (Jiang et al., 2015) and nucleolus (Hut et al., 2017), play critical roles in regulating a variety of biological responses (Riback et al., 2017), RNA metabolism (Taylor et al., 2016), DNA damage response (Paioli et al., 2015) and signal transduction (Franzmann et al., 2016).

Here, we presented a data resource of LLPS (**DrLLPS**), a comprehensive data resource that contained **437,887** known and computationally detected LLPS-associated species. For LLPS-associated proteins in nine model organisms, we provided rich annotations by compiling and integrating the knowledge from additional **110** widely cover **16** aspects, including (i) Intrinsically disordered region; (ii) Domain annotation; (iii) Post-translational modification; (iv) Genetic variation; (v) Cancer mutation; (vi) Disease-associated information; (vii) Drug-target relation; (viii) Physicochemical property; (ix) Protein functional annotation; (x) Protein expression/proteomics; (xi) Subcellular localization; (xii) mRNA expression; (xiii) DNA & RNA element; (xiv) DNA methylation. The online service of DrLLPS was implemented in PHP + MySQL + and annotations are freely accessible for all users. We anticipate DrLLPS can serve as a helpful resource for further analysis of LLPS, and confirm that the database will be updated.

9,285 Curated Proteins

40 Condensates

164 Species

110 Pub

Simple Search

Please search the DrLLPS to find the information you need. Please input one keyword to find the related information:

MLO-associated proteins classified as scaffolds, regulators, and clients as assessed by automated text mining, followed by curator assessment. *In vitro*, *in vivo* and computational evidence is accepted (HTS, LTS), such as KO, silencing, overexpression.



What does “LLPS protein” mean?

PhaSePro: 120 proteins

The screenshot shows the PhagePro! website's search results page. At the top, there's a navigation bar with links for Home, Bioinformatics, About Us, Contact, Help, and Log In. Below the navigation is a large, bold title "Welcome to PhagePro!". To the right of the title is a circular graphic containing various colored, abstract shapes representing phage particles. The main content area has a light blue background and contains a search bar with the query "lambda". Below the search bar, there are two sections of search results. The first section is titled "Search results for lambda" and lists several entries: "lambda", "lambda bacteriophage", "lambda genome", "lambda protein", "lambda protein database", "lambda protein sequence", "lambda protein structure", "lambda protein function", "lambda protein localization", "lambda protein expression", "lambda protein regulation", and "lambda protein synthesis". Each result has a small thumbnail image of a phage and a "View details" button. The second section is titled "Search results for lambda" and lists several entries: "lambda", "lambda bacteriophage", "lambda genome", "lambda protein", "lambda protein database", "lambda protein sequence", "lambda protein structure", "lambda protein function", "lambda protein localization", "lambda protein expression", "lambda protein regulation", and "lambda protein synthesis". Each result has a small thumbnail image of a phage and a "View details" button.

LLPSDB: 1200

LLPSDB A database of proteins undergoing liquid-liquid phase separation in cells

[Home](#) [Browse](#) [Search](#) [Submit](#) [Statistics](#) [Download](#)

Welcome to LLPSDB !

Liquid-liquid phase separation (LLPS) of proteins has been discovered to underlie the compartmentalization of cells, through the formation of liquid-liquid (condensates), involving thousands of proteins. LLPSDB is a database that integrates experimental data from the biological processes such as the nucleolus, endoplasmic reticulum, and nuclear envelope, and the analysis of the available databases may be a better understanding of LLPS. In the report, numerous biomolecular systems, or biomolecules and corresponding nucleic acids, and various types of phase separation have been analyzed, and deposited relevant evidence.

LLPSDB contains LLPS related proteins together with the corresponding phase separation information gathered by researchers; for each protein, the detailed phase separation information, including the phase separation conditions, as well as specific phase events, as well as domain boundaries, are provided. The database also provides the detailed information about the phase separation conditions, including species, cell/tissue, method, phase behavior etc. In addition, several useful tools are provided, such as the search function, the visualization function, the comparison function, and the analysis function. All the information is freely available for research.

News
A new version of LLPSDB has been released.
and Protein condensates can now recruit RNA polymerase II on Day 19, 2019.

» [LLPSDB currently hosts 1177 entries](#)
» [Last updated: July 1, 2019](#)

Databases Related

[Lipid](#) [Protein](#) [Protein-DNA](#)
[Protein-RNA](#) [Protein-Protein](#)
[Protein-DNA](#) [Protein-RNA](#)

Contact Us

[Email: \[lyy@bnu.edu.cn\]\(mailto:lyy@bnu.edu.cn\)](#)
[Biology Department, Beijing Jiaotong University](#)
[Beijing, 100040, China](#)

PhaSepDB: 3000

DrLLPS: 9300



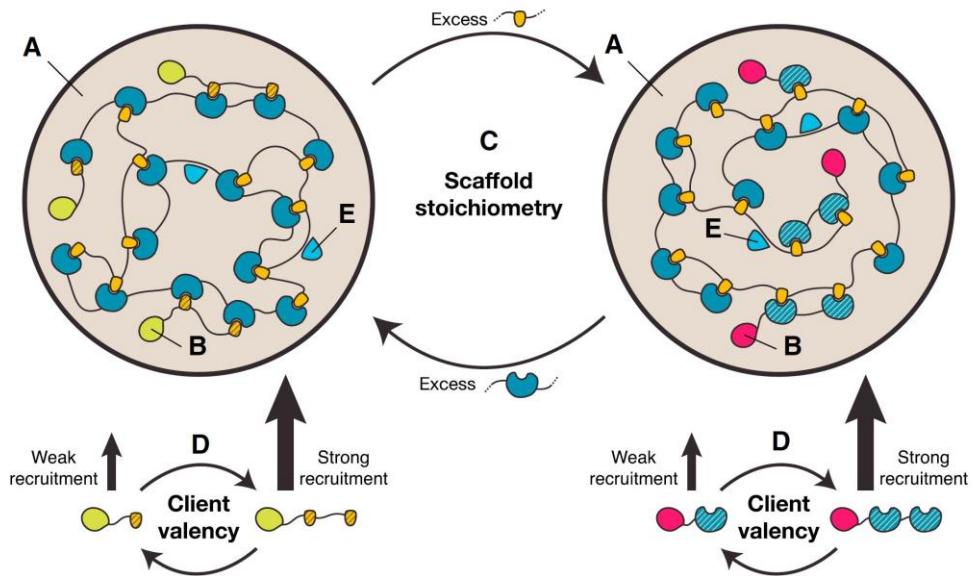
Farahi et al. (2021) *Int J Mol Sci.* 22: 3017

Why is not straightforward to call an LLPS protein?

- 1) the capacity to phase separate is not a binary classifier (not intrinsic but contextual property of the protein and its environment)
- 2) phase separation depends on the concentration of the protein (physiological?)
- 3) proteins have distinct roles in phase separation
- 4) LLPS is not equivalent to biomolecular condensation (which includes gelation, crystallization, clustering, pleiomorphic assembly, polymerization and amorphous or amyloid aggregation) or templated assembly.



Scaffolds and clients

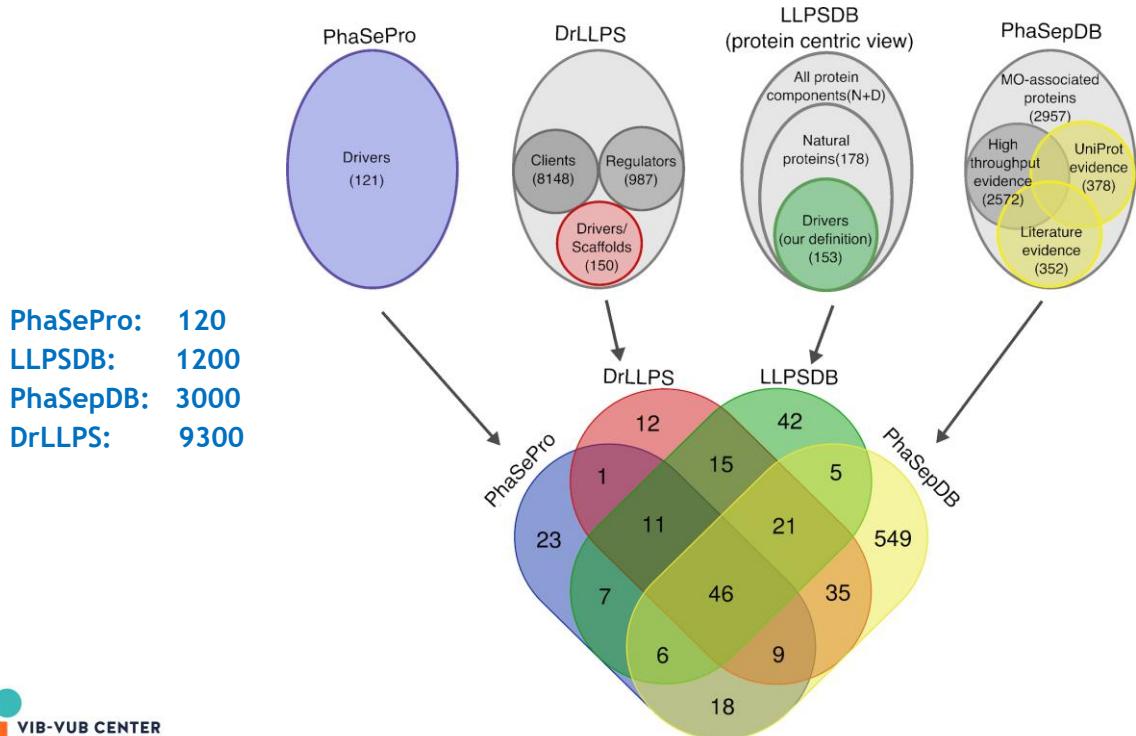


Different roles of proteins in LLPS.

- 1) **Driver (scaffold):** can phase separate on their own. If RNA is mandatory, we consider it as a “co-driver”. Small molecules (and crowder) are “condition”.
- 2) **Co-driver:** a macromolecule (protein, RNA or DNA) that strictly requires another macromolecule for phase separation (then both are “co-drivers”)
- 3) **Regulator:** its presence/activity is required for LLPS, but no part of condensate (modifying enzyme, transport protein, transcription factor, etc...)
- 4) **Client:** not required for and has no effect on LLPS, but localizes to the condensate formed (through interactions with driver/co-driver).



Different databases contain different type (and amount) of data



A primary goal: develop LLPS predictors

Table 1

Phase-separation prediction methods

Method name	Description	Residue # information	Availability	Ref.
PLAAC	Prion-like domain prediction by hidden Markov model	Yes	plaac.wi.mit.edu	[43,46]
LARK	Energetic compatibility with aromatic-rich kinked amyloid structures, by threading sequences and modeling with Rosetta	Specific regions	In supplement of original paper	[23**]
R + Y	Critical concentration prediction based on number of arginine and tyrosine residues, extrapolated from FET family proteins	No	In supplement of original paper	[50**]
DDX4-like	Sequence composition and residue spacing similarity to DDX4	Specific regions	From authors	[51]
CatGranule	Composition weighted by sequence length, R/G/F content, and amino acid propensity for nucleic acid binding and disorder	Yes	tartaglia.com	[52]
PScore	Prediction based on expected numbers of long-range planar sp2 pi-pi contacts	Yes	abragam.med.utoronto.ca/~JFKlab/crapome.org	[42**]
CRAPome	Empirical measurement related to non-specific interactions and <i>in vivo</i> concentration, taken from the frequency at which each human protein is identified by affinity purification mass spectrometry negative controls	No		[54]

PSPredictor machine learning based on LLPSDB database <http://www.pkuml.cn:8000/PSPredictor/>



VIB-VUB
CENTER
FOR STRUCTURAL
BIOLOGY

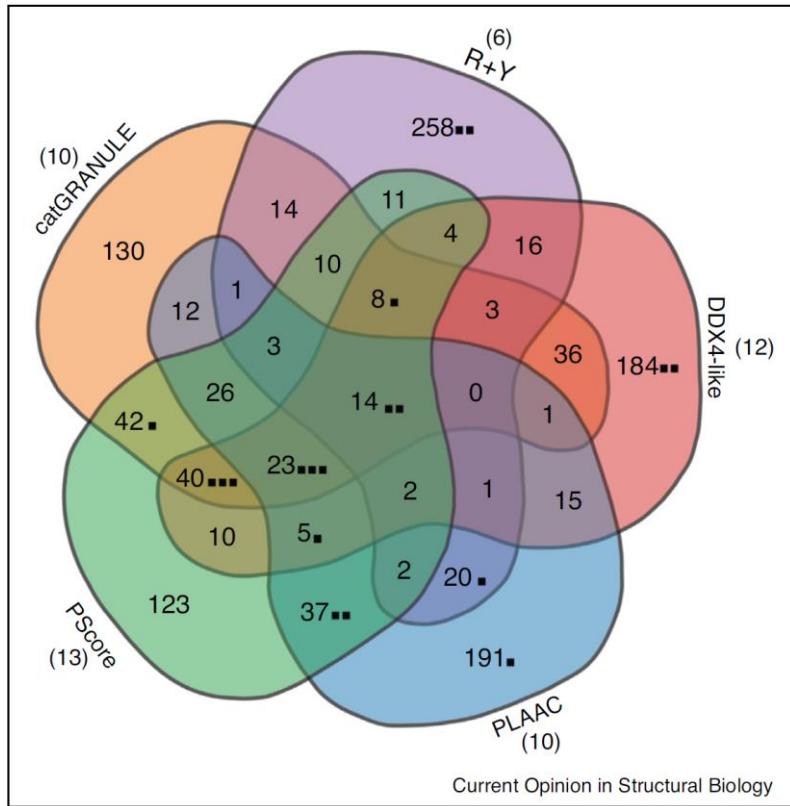


VRIJE
UNIVERSITEIT
BRUSSEL



PhasAGE

Prediction of LLPS proteins in human proteome



Thank you