Role of phase transitions of fusion proteins in cancer

K 124670 PROJECT – FINAL REPORT

Project start: 2017. 09. 01.

Closing date: 2022. 08. 31.

Bioinformatics analyses:

Data integration and bioinformatics analyses on the oncogenic fusions of ALK and other receptor tyrosine kinases:

By screening anaplastic lymphoma kinase, ALK-specific literature and cancer databases/datasets, we have identified 67 hitherto described ALK oncogenic fusion partners totalling 75 different gene fusion arrangements occurring in 14 different cancer types. For each fusion, we have registered the corresponding literature references, cancer type, chromosomal translocations, partner gene, UniProt identifier and precise residue boundaries of the protein regions fused, as well as the subcellular localization of the fusion product. To our knowledge, this is the most comprehensive such collection.

Fused regions of the partner proteins have been thoroughly analyzed for diverse features: 1) we predicted their structural disorder, aggregation propensity and sequence complexity by dedicated bioinformatics methods. They did not show a significant deviation from randomly selected regions regarding these features. 2) We then compared them to experimentally validated human amyloidogenic protein regions from AmyPro (http://amypro.net/#/), and regions experimentally confirmed to drive liquid-liquid phase separation (LLPS) from PhasePro (see later). These analyses suggested that amyloid formation and LLPS do not play a role in oncogenic ALK fusions. 3) We then analyzed the annotated domains of the fused regions of the partner proteins to see if known dimerization domains occur. We first checked their UniProt annotations for known dimerization domains or coiled coil regions. If we did not find enough evidence for dimerization in UniProt, we also checked PDB structures of the component domains for dimers, and lastly the literature for any observation of dimerization of the given protein. Interestingly, in almost each case we found evidence for homodimerization of the fused region of the partners, which completely supports the assumption on constitutive ALK tyrosine kinase activation through fusion partner-mediated dimerization and cross-phosphorylation of the kinase domains. The medically most relevant ALK fusions have been experimentally studied, namely the NPM1-, TPM3-, STRN-, EML4-ALK fusions, which are responsible for at least 90% of the ALK fusion protein mediated cancers. In all cases, dimerization or discrete homo-oligomerization (for NPM1 (Bischof D et al. Mol Cell Biol. 1997; 17(4):2312-25, Fujimoto J et al. PNAS. 1996; 93(9):4181-6)) of the fusion partner was confirmed, leading to constitutive ALK activation and transformation potential. Since the other

identified partners also have coiled coils or other validated dimerization domains and most have been only described in one or maximum a few cancer patients, we agreed that their detailed experimental characterization would be rather superfluous, while not being medically very relevant either.

Our original hypothesis was that liquid-liquid phase separation or aggregation could contribute to the transformation potential of cancer fusion proteins. Since we could not detect such regions within the fusion partners of ALK, we have also checked the fusion partners of the other 57 human receptor tyrosine kinases (RTKs; Lemmon MA and Schlessinger J Cell. 2010; 141(7):1117-34) in the Atlas of Genetics and Cytogenetics in Oncology and Haematology. Other RTKs had no or considerably less known fusion partners than ALK and their fusion partners showed a large overlap with those of ALK. Therefore, again, the vast majority of the fused partner regions had coiled coils or other known dimerization domains and we did not find any overlap with LLPS or amyloid-forming proteins. Since we were not successful in identifying regions driving LLPS or aggregation within RTK fusion partners, we decided to take another approach and analyze the oncogenic fusions of all human LLPS drivers.

Data integration and database building of proteins capable to drive LLPS:

To be able to perform an impactful, comprehensive analysis of the oncogenic fusions of human LLPS driver proteins, one needs a reliable dataset on LLPS drivers. However, no dedicated database or any comprehensive collection of LLPS proteins was available in the literature at that time. Since the lack of a dedicated database seemed to be a major drawback of the whole LLPS field in general, we decided to build a dedicated LLPS database ourselves. We have manually collected and integrated all the available LLPS data from the literature with dr Rita Pancsa leading this project. We created a comprehensive database of more than 120 LLPS proteins, PhaSePro (https://phasepro.elte.hu/), that contains driver region and functional annotations, as well as information on the regulation (post-translational modifications, alternative splicing) and disease-associated mutations of the curated LLPS systems/proteins. Cross references to existing databases/ontologies and the web interface of PhaSePro was prepared in collaboration with the MTA-ELTE Momentum Bioinformatics Research Group led by Dr Zsuzsanna Dosztanyi (Mészáros B et al. Nucleic Acids Res. 2020; 48(D1):D360-D367; https://doi.org/10.1093/nar/gkz848). The up-to-date, high-quality, manually curated LLPS dataset underlying PhaSePro largely facilitated the comprehensive bioinformatics analysis of oncogenic fusions of LLPS driver regions detailed below.

Also, to ensure that LLPS data get integrated into the core data resources of the IDP field, we proposed that the DisProt database should accept LLPS as a function assigned to intrinsically disordered regions (IDRs) and we then annotated the identified LLPS driver proteins into DisProt as part of the latest releases: https://doi.org/10.1093/nar/gkz975 ; https://doi.org/10.1093/nar/gkab1082.

Several LLPS-related data resources have been published at the same time, therefore, during 2020 we integrated the data on LLPS-associated proteins from 3 dedicated databases, PhaSePro, DrLLPS and LLPSDB. We could identify 57 proteins annotated as drivers by all

and another 59 proteins that were annotated by two of them. However, an additional 135 proteins were only annotated in one of them. We could altogether find 117 high-confidence human LLPS driver proteins. See our publication here: https://doi.org/10.3390/ijms22063017. In the same study we found that the 117 human LLPS-associated genes are highly enriched among dosage sensitive genes and among genes showing a strong association with cancer, indicating that somatic mutations or chromosomal translocation-derived gene fusions could not only cause problems by changing the sequences of the affected proteins qualitatively, but also by changing their cellular levels quantitatively.

LLPS and cancer, the state of the art:

The intensive research efforts resulting in a flurry of publications reporting on new cases or characteristics of LLPS largely furthered our understanding of its physical, structural and functional features under physiological conditions, however, we know much less on how perturbations of phase separation and the associated condensates may contribute to the development of various diseases (Alberti S and Dormann D. Annu Rev Genet. 2019; 53:171-194). RNA-binding proteins, e.g. the FUS family proteins, are abundantly represented among LLPS drivers and are implicated in diverse diseases, such as neurodegenerative disorders, muscular atrophies and cancer. It is also well known that phase-separated liquid-like structures can transition into less dynamic hydrogels or protein aggregates containing amyloid-like filaments that are involved in neurodegenerative diseases, such as amyotrophic lateral sclerosis, frontotemporal dementia and Alzheimer's disease. While its link to neurodegeneration has been more extensively explored, the association of LLPS to cancer in the form of cancer-associated mutations/translocations directly affecting LLPS drivers and the corresponding condensates has only been demonstrated for some individual proteins, such as SPOP and SHP2 and some oncogenic fusions mainly of the FET family (Kwon I et al. Cell. 2013; 155(5):1049-1060) and of nucleoporins. Also, condensate dysregulation through changes in the availability (localization and/or expression level) of the drivers is a common feature of cancer cells (Boija A et al. Cancer Cell. 2021; 39(2):174-192). The role of LLPS in cancer was extensively reviewed recently (Nozawa et al. Cancer Sci. 2020; 111(9):3155-3163; Mehta S and Zhang J. Nat Rev Cancer. 2022; 22(4):239-252; Jiang S Elife. 2020; 9:e60264.) and a computational analysis showed that proteins implicated in diseases, including cancer, are enriched in phase separation propensity predicted by the PScore method (Tsang B. et al. Cell. 2020; 183(7):1742-1756) that solely relies on pi-pi inter-residue interactions. However, to our knowledge the relationship between the hitherto experimentally characterized LLPS drivers and known cancer drivers has not yet been thoroughly analyzed on statistical terms.

LLPS as an important feature in somatic cancers: data integration and bioinformatics analysis

To investigate the level of involvement of validated human LLPS drivers in different disease categories, we obtained the genes implicated in somatic cancers from the COSMIC database, and integrated the genes implicated in neurodegenerative diseases and hereditary cancers from ClinVar and HUMSAVAR. We then analyzed the enrichment of LLPS driver genes in these different disease-related gene sets. A large number of novel LLPS drivers have been published since the latest releases of the above integrated three LLPS resources, therefore we have

scanned through the novel bulk of LLPS literature and extended our dataset with further 54 manually curated proteins published in 2020, to ultimately gain a dataset of 141 high-confidence human LLPS drivers.

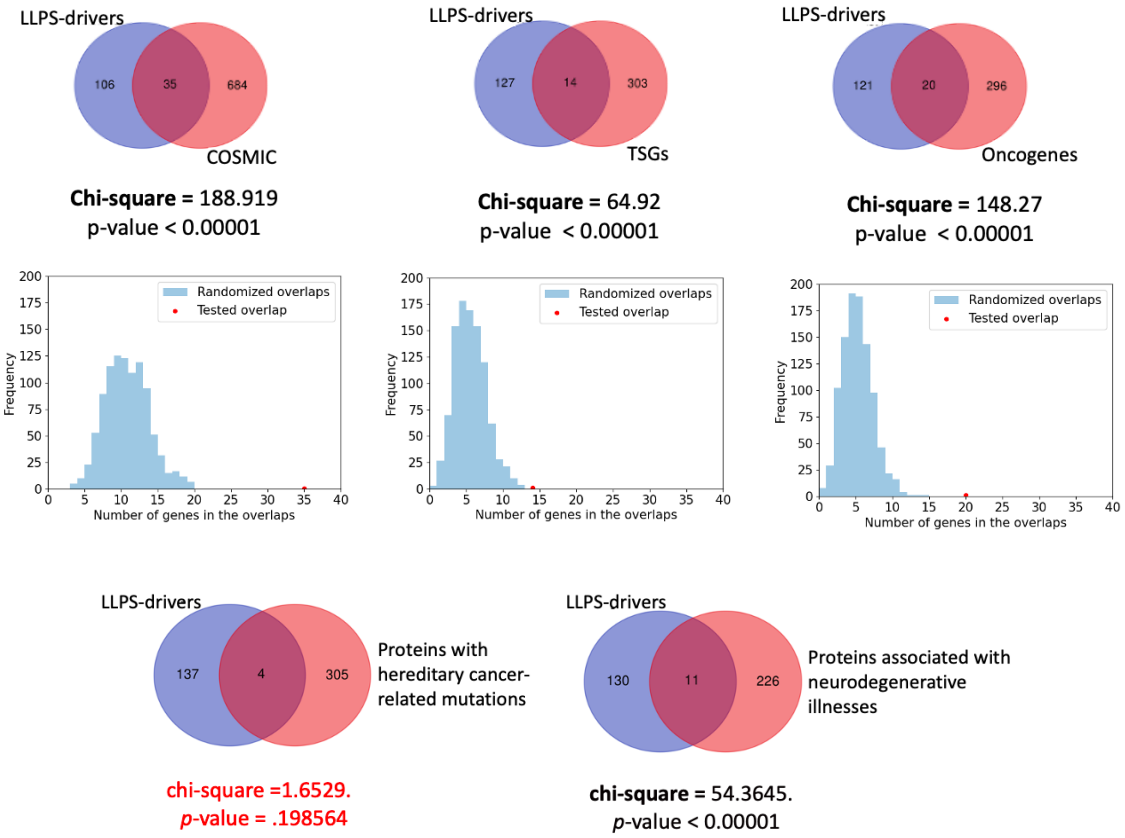

Figure 1: LLPS drivers are enriched in diseases including neurodegenerative diseases and even more in somatic cancers (both among tumor suppressor genes (TSGs) and oncogenes), but not in germline cancers. Enrichment of these in protein sets implicated in different disease types was analyzed by Chi-square tests. Statistic analyses involving randomized selections brought similar results (see histograms). Also, LLPS drivers were more enriched in somatic cancers than amyloid-forming proteins from AmyPro (data not shown).

We found that validated LLPS drivers are enriched in genes associated with neurodegenerative diseases, and even more in genes associated with somatic cancers, but not in those associated with germline cancers. Also, they are enriched in both oncogenes and tumor suppressors, but more in oncogenes (Figure 1).

By analyzing the GO molecular functions and biological processes associated to proteins implicated in somatic cancers (COSMIC consensus proteins) in general and cancer-associated LLPS drivers in particular, we found that cancer-associated LLPS drivers are enriched in a specific set of functions/processes (mainly gene expression, epigenetics, regulation of protein maturation, protein localization to organelles, regulation of protein stability, protein acylation,

chromosome segregation, DNA replication and DNA recombination) compared to all cancer-associated proteins (data not shown).

We then decided to analyze the features of the 35 LLPS drivers heavily implicated in somatic cancers by comparing them to the whole set of cancer consensus genes. Luckily for our analysis, COSMIC contains a number of very important features assigned to all the census proteins that we could use for our analysis.

We found that cancer-associated LLPS drivers are mostly oncogenes with dominant genetics and low actionability, which means that LLPS is mainly linked to non-treatable forms of cancer with strong phenotypes (Figure 2A). Furthermore, gene fusion was found to be the dominant mutation type for three-fifths of cancer-associated LLPS drivers, while regulators are mainly affected by missense mutations, and clients by large deletions (Figure 2B). We also found that cancer-associated LLPS drivers are heavily associated with almost all known cancer hallmarks, while regulators show weaker associations with fewer hallmarks and clients show now significant associations (Figure 2C).

A) general features B) mutation types C) hallmarks

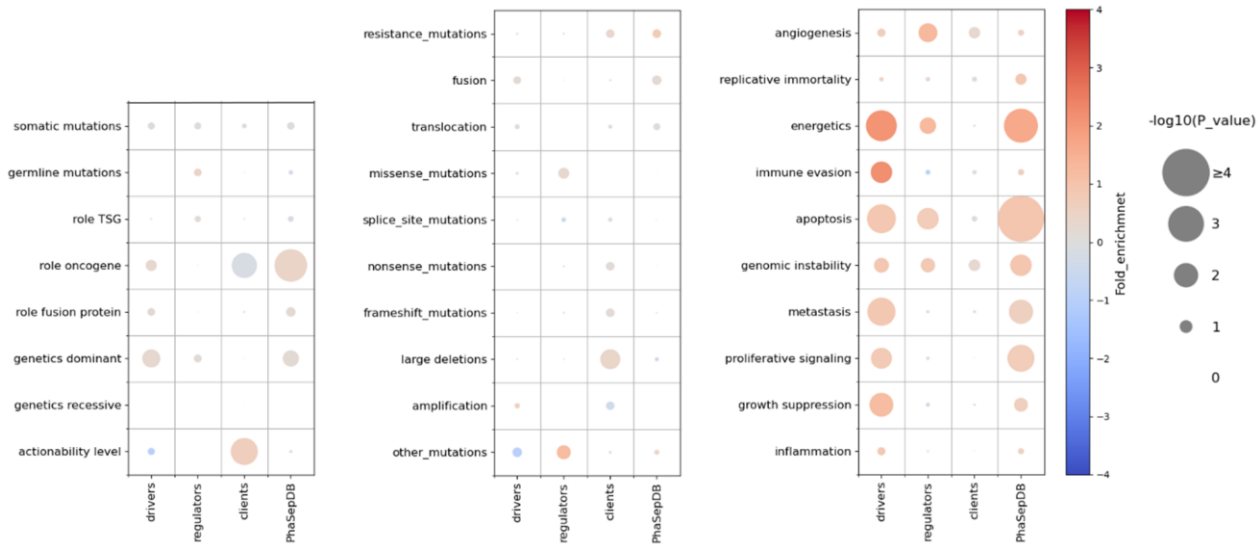

Figure 2: Cancer-associated LLPS drivers tend to be oncogenes with dominant genetics and low actionability, that are heavily associated with most cancer hallmarks and frequently form oncogenic fusion proteins through gene fusions. Strengths of associations between the different protein groups and different cancer features/hallmarks are indicated by patches, where the color of the patches represents the fold enrichments, while the size of the patches the p-values of the calculated enrichments. In this case PhaSepDB (fourth category) was used as an alternative dataset of LLPS drivers, that is although larger than our confident set of LLPS drivers (first category), but to some extent contaminated by proteins that are not true drivers of LLPS just associated to condensates (regulators or clients). Regulators indicated in the second category were obtained from the DrLLPS database. Clients were integrated from diverse resources storing gene lists of different condensates.

A functional dissection of oncogenic fusion proteins:

Since three-fifths (21 of the 35) of cancer-associated LLPS drivers have gene fusions indicated as the dominant mutation type, we decided to analyze the gene fusions in more detail. To this end, we first collected all the fusions of the 729 cosmic consensus genes based on their fusion partners listed by either COSMIC or UniProt. We only accepted a single fusion arrangement for each fusion gene pair into our dataset, possibly the one that occurred most frequently in studied cancer samples. Also, we only accepted oncogenic fusion proteins that had a clear connection to a particular cancer type (or a well-defined set of cancers) and represent the in-frame fusions of two different coding regions. This means that for example cases where a gene inherited the gene regulatory region of another gene by fusion were not collected. Therefore, the 308 oncogenic fusion proteins of our final dataset contain only chimera proteins that incorporate (in-frame translated) protein regions of both fusion partners, meaning that they potentially inherited functional protein regions/modules from both fusion partners and thus represent new combinations of those. This dataset represents a comprehensive collection of fusion proteins, where fusion boundaries are defined on the protein level (with UniProt residue boundaries). Therefore, the inclusion/exclusion of particular domains/modules of the fusion partners into the fusions are clearly defined.

To be able to discover typical combinations of molecular functions that occur in oncogenic fusion proteins, but do not occur in wild-type proteins (and are therefore likely to be responsible for oncogenicity), we first assigned module-level functional information to the sequences of the wild type proteins forming the fusions and to the sequences of the fusions themselves. This is not an easy task because for example GO molecular functions are usually assigned to full proteins and not to protein domains/motifs/regions. Due to this reason, we used a combination of InterPro2GO domain-level GO annotations, published domain lists associated with certain molecular functions and UniProt annotations to produce a wealthy module-level annotation for the protein sequences. Attaching GO terms/functions to protein modules allowed us to do a systematic analysis of the associations between functions. The above described dataset, i.e. the annotated fusion proteins with precise protein boundaries and with molecular functions assigned to their modules will be published as a stand-alone online resource in 2023.

This richly annotated dataset allowed us to calculate associations between the different molecular functions for the wild-type proteins and for the fusions using overlap coefficients. The strength of our data and analysis approach is well demonstrated by the fact that it readily highlights the frequent association between dimerization/ homo-oligomerization domains and tyrosine kinase domains in fusions (an overlap coefficient of 0.62), an already known oncogenic combination of molecular functions that is typical in the fusion proteins of receptor tyrosine kinases, but is not seen in wild-type proteins. Encouraged by recovering this already known fusion-specific association of functions in a statistically meaningful manner, we also looked for novel fusion-specific associations of functions in our dataset, especially focusing on those that involved the ability to drive LLPS as one of the functions.

Interestingly, we obtained some statistically significant novel associations. As expected based on our preliminary data previously described in our yearly reports, we found that LLPS driver

regions are frequently associated with intact DNA-binding domains of transcription factors (TFs) in fusions. Functional association tables are two large to be shown here, but overlap coefficients range between 0 and 1, where values around and above 0.2 could be considered significant and the value for the association between LLPS driver regions and DNA-binding was 0.46). Our results therefore highlight that phase-separation-prone oncogenic fusion proteins tend to incorporate a specific combination of domains/functions.

Although many transcription factors (TFs) were demonstrated to undergo phase separation in vitro, on their own they could only phase separate at very high concentrations (Boija A et al. Cell. 2018; 175(7):1842-1855.e16), that are not compatible with their otherwise notoriously low cellular levels. For phase separation to occur at physiological-like concentrations TFs require a coactivator and a DNA segment that carries multiple copies of their respective recognition elements (Shrinivas K et al. Mol Cell. 2019; 75(3):549-561.e7.). Therefore, transcription factors are not self-sufficient, but largely context-dependent LLPS drivers. However, in the most dreadful phase-separation-prone oncogenic fusion proteins the N-terminal transactivation domains of transcription factors are exchanged to self-sufficient (i.e. context-independent) LLPS driver regions mainly of nucleoporins (e.g. NUP98) or FUS family proteins (EWSR1, FUS, TAF15) that can by themselves drive LLPS at physiological concentrations without a need for any partners (Wang J et al. Cell. 2018; 174(3):688-699.e16; Schmidt HB et al. Elife. 2015; 4:e04251). Such a combination suggests that the context-dependence of TFs may be lost/resolved in the fusions, so they could undergo LLPS without the requirement for partners (their specific recognition elements or coactivators). This is supported by CHIP-seq experiments done by HOXA9 and its fusion protein NUP98-HOXA9 that show that the fusion product binds to distinct sites in the genome than the original transcription factor (Rio-Machin A et al. Leukemia. 2017; 31(9):2000-2005). Also, the EWS-FLI fusion was reported to bind GGAA-containing elements (microsatellites) that are distinct from the regions targeted by FLI1 (Gangwal K et al. Proc Natl Acad Sci U S A. 2008; 105(29):10149-54).

Our dataset also highlighted that the LLPS drivers mostly reside on the N-terminal side of the fusion proteins. Therefore, the cellular level/copy number of the fusion product may be drastically higher than that of the original transcription factor in healthy cells, because the fusion product inherits the gene regulatory regions of the N-terminally fused gene (in this case the LLPS driver), which will probably confer stronger expression that is less restricted in space (tissues) and time (developmental stages, phases of the cell cycle, etc.) than that of the original TF (TFs are lowly expressed genes with highly restricted/tightly regulated expression patterns).

Taken together, the results of our bioinformatics analyses propose a common molecular pathomechanism for the most dreadful phase-separation-prone oncogenic fusion proteins where the original transactivation domains of TFs that are relatively weak, context-dependent LLPS drivers are exchanged to much stronger, self-sufficient LLPS driver regions, which is also coupled with a marked increase in the spatio-temporal range and level of their expression. This leads to out-of-context, abnormal phase-separation events along the DNA that deregulates the transcription of a high number of genes through divergent chromatin remodeling

mechanisms that include the perturbation of both histone modification patterns and epigenetic dynamics of target genes.

The above results on the role of LLPS in somatic cancers and on the functional dissection of oncogenic fusion proteins constitute part of a manuscript that is considered as one of the major deliverables of the project and will be soon submitted to the journal Molecular BioSystems.

LLPS and fusion proteins in cancer

In the experimental part of the project we studied Nup98 and developed a new method to study phase separation of aggregation-prone proteins that is more descriptive of their native behavior than the commonly used techniques. We demonstrated that at carefully selected pH values proteins such as the low-complexity domain of hnRNPA2, TDP-43, and NUP98, or the stress protein ERD14, can be kept in solution and their LLPS can then be induced by a jump to native pH.
We also attempted to characterize the interaction between the GLEBS domain of Nup98 and its partner proteins such as POL II CTD, WDR5 and WDR82. We could not detect interaction with POL II CTD under the circumstances we used and the interaction studies with WDR5 and WDR82 were inconclusive due to solubility issues of the WDR proteins.

We also studied androgen receptor (AR), a nuclear hormone receptor that regulates the transcription of genes involved in the development of testis, prostate and the nervous system. Misregulation of AR is a major driver of prostate cancer (PC) and frequently gives rise to the generation of fusion proteins. As it was demonstrated that full-length AR can undergo LLPS in a cellular model of PC, we have analyzed which AR region is responsible for LLPS. We found that its DNA-binding domain (DBD) is capable of RNA binding and undergoes RNA-dependent LLPS. As RNA binds DBD weaker than DNA, while both RNA and DNA localize into AR droplets, its LLPS depends on the relative concentration of the two nucleic acids. Importantly, the oncogenic splice variant of AR, AR-v7 lacks the ligand binding domain is incapable of LLPS. According to our findings, the functional part of the long N-terminal disordered transactivation domain termed activation function 1 inhibits AR-v7 phase separation. This diminished LLPS may contribute to the misregulation of the transcription function of AR in prostate cancer.

Another important oncogenic protein that goes through LLPS is Heterogeneous nuclear ribonucleoprotein A2 (hnRNPA2). We showed that the low complexity domain (LCD) of hnRNPA2 interacts with RNA via an embedded Tyr/Gly-rich region. We identified this segment as a disordered RNA-binding motif. Because RNA binding is maintained upon mutating tyrosine residues to phenylalanines, but not by mutating to alanines, we named the RNA-binding region 'F/YGG motif'. The F/YGG motif can bind a broad range of structured (e.g. tRNA) and disordered (e.g. polyA) RNAs, but not rRNA. As the F/YGG motif can also interact with DNA, we can consider it a general nucleic acid-binding motif. Similar sequence motifs can be found in other proteins as well, such as the LCD of TDP43 and Nup proteins (including Nup98), suggesting that the F/YGG motif is a general nucleic acid-interaction motif.

hnRNPA2 can go through LLPS: at physiological pH, protein-rich droplets quickly appear in the solution of hnRNPA2 LCD. Over time, these droplets mature and turn into aggregates. When adding polyU RNA to hnRNPA2 LCD droplets, they co-phase separate, showing that hnRNPA2 LCD not only binds RNA, but also recruits the RNA into phase-separated droplets, but higher RNA concentrations lead to decreased droplet size and the abrogation of aggregation. Our results showed that the effect of increased RNA concentration is the increasement of the protein concentration in the dilute phase.
Our results highlighted an important feature of the hnRNPA2 LCD in LLPS and the description of the F/YGG motif can have further implications in the study of protein-nucleic acid interactions.

Publications

1. Bratek-Skicki, A. ✉ ; Van, Nerom M. ; Maes, D. ; Tompa, P.
Biological colloids: Unique properties of membraneless organelles in the cell/
ADVANCES IN COLLOID AND INTERFACE SCIENCE 310 Paper: 102777 (2022)

2. Guharoy, Mainak ; Lazar, Tamas ; Macossay-Castillo, Mauricio ; Tompa, Peter
Degron masking outlines degronons, co-degrading functional modules in the proteome
COMMUNICATIONS BIOLOGY 5 : 1 Paper: 445 (2022)

3. Lazar, Tamas ; Tantos, Agnes ; Tompa, Peter ; Schad, Eva
Intrinsic protein disorder uncouples affinity from binding specificity
PROTEIN SCIENCE 31 : 11 Paper: e4455 (2022)

4. Quaglia, Federica ; Mészáros, Bálint ; Salladini, Edoardo ; Hatos, András ; Pancsa, Rita ; Chemes, Lucía B ; Pajkos, Mátyás ; Lazar, Tamas ; Peña-Díaz, Samuel ; Santos, Jaime et al.
DisProt in 2022: improved quality and accessibility of protein intrinsic disorder annotation
NUCLEIC ACIDS RESEARCH 50 : D1 pp. D480-D487. (2022)

5. Van Lindt, Joris ; Lazar, Tamas ; Pakravan, Donya ; Demulder, Manon ; Meszaros, Attila ; Van Den Bosch, Ludo ; Maes, Dominique ; Tompa, Peter
F/YGG-motif is an intrinsically disordered nucleic-acid binding motif
RNA BIOLOGY 19 : 1 pp. 622-635. , 14 p. (2022)

6. Ahmed, Junaid ; Meszaros, Attila ; Lazar, Tamas ; Tompa, Peter
DNA -binding domain as the minimal region driving RNA -dependent liquid–liquid phase separation of androgen receptor
PROTEIN SCIENCE 30 : 7 pp. 1380-1392. , 13 p. (2021)

7. Farahi, N. ; Lazar, T. ; Wodak, S.J. ; Tompa, P. ✉ ; Pancsa, R. ✉

Integration of data from liquid–liquid phase separation databases highlights concentration and dosage sensitivity of llps drivers
INTERNATIONAL JOURNAL OF MOLECULAR SCIENCES 22 : 6 Paper: 3017 , 26 p. (2021)

8. Korkmazhan, E. ; Tompa, P. ; Dunn, A.R. ✉
The role of ordered cooperative assembly in biomolecular condensates
NATURE REVIEWS MOLECULAR CELL BIOLOGY 22 : 10 pp. 647-648. , 2 p. (2021)

9. Lazar, Tamas ; Martínez-Pérez, Elizabeth ; Quaglia, Federica ; Hatos, András ; Chemes, Lucía B ; Iserte, Javier A ; Méndez, Nicolás A ; Garrone, Nicolás A ; Saldaño, Tadeo E ; Marchetti, Julia et al.
PED in 2021: a major update of the protein ensemble database for intrinsically disordered proteins.
NUCLEIC ACIDS RESEARCH 49 : D1 pp. D404-D411. (2021)

10. Murvai, N. ; Kalmar, L. ; Szabo, B. ; Schad, E. ; Micsonai, A. ; Kardos, J. ; Buday, L. ; Han, K.-H. ; Tompa, P. ; Tantos, A.
Cellular chaperone function of intrinsically disordered dehydrin erd14
INTERNATIONAL JOURNAL OF MOLECULAR SCIENCES 22 : 12 Paper: 6190 (2021)

11. Necci, M. ; Piovesan, D. ; Hoque, M.T. ; Walsh, I. ; Iqbal, S. ; Vendruscolo, M. ; Sormanni, P. ; Wang, C. ; Raimondi, D. ; Sharma, R. et al.
Critical assessment of protein intrinsic disorder prediction
NATURE METHODS 18 : 5 pp. 472-481. , 10 p. (2021)

12. Pakravan, D. ; Michiels, E. ; Bratek-Skicki, A. ; De, Decker M. ; Van, Lindt J. ; Alsteens, D. ; Derclaye, S. ; Van, Damme P. ; Schymkowitz, J. ; Rousseau, F. et al.
Liquid–liquid phase separation enhances tdp-43 lcd aggregation but delays seeded aggregation
BIOMOLECULES 11 : 4 Paper: 548 (2021)

13. Parisi, G. ✉ ; Palopoli, N. ; Tosatto, S.C.E. ; Fornasari, M.S. ; Tompa, P. ✉
“Protein” no longer means what it used to
CURRENT RESEARCH IN STRUCTURAL BIOLOGY 3 pp. 146-152. , 7 p. (2021)

14. Quaglia, F. ; Lazar, T. ; Hatos, A. ; Tompa, P. ; Piovesan, D. ✉ ; Tosatto, S.C.E. ✉
Exploring Curated Conformational Ensembles of Intrinsically Disordered Proteins in the Protein Ensemble Database
Current Protocols 1 : 7 Paper: 192 (2021)

15. Van, Lindt J. ; Bratek-Skicki, A. ✉ ; Nguyen, P.N. ; Pakravan, D. ; Durán-Armenta, L.F. ; Tantos, A. ; Pancsa, R. ; Van, Den Bosch L. ; Maes, D. ; Tompa, P. ✉

A generic approach to study the kinetics of liquid–liquid phase separation under near-native conditions
COMMUNICATIONS BIOLOGY 4 : 1 Paper: 77 (2021)

16. Bokor, M. ; Tantos, Á. ; Mészáros, A. ; Jenei, B. ; Haminda, R. ; Tompa, P. ; Tompa, K.
Molecular Motions and Interactions in Aqueous Solutions of Thymosin-β4, Stabilin CTD and Their 1 : 1 Complex, Studied by 1H-NMR Spectroscopy
CHEMPHYSCHEM: A EUROPEAN JOURNAL OF CHEMICAL PHYSICS AND PHYSICAL CHEMISTRY 21 : 13 pp. 1420-1428. , 9 p. (2020)

17. Bokor, M. ; Tantos, Á. ; Tompa, P. ; Han, K.-H. ; Tompa, K.
WT and a53t α-synuclein systems: Melting diagram and its new interpretation
INTERNATIONAL JOURNAL OF MOLECULAR SCIENCES 21 : 11 Paper: 3997 , 16 p. (2020)

18. Bratek-Skicki, Anna ; Pancsa, Rita ; Meszaros, Balint ; Van Lindt, Joris ; Tompa, Peter
A Guide to Regulation of the Formation of Biomolecular Condensates.
FEBS JOURNAL 287 : 10 pp. 1924-1935. , 12 p. (2020)

19. Csermely, P. ✉ ; Kunsic, N. ; Mendik, P. ; Kerestély, M. ; Faragó, T. ; Veres, D.V. ; Tompa, P.
Learning of Signaling Networks: Molecular Mechanisms
TRENDS IN BIOCHEMICAL SCIENCES 45 : 4 pp. 284-294. , 11 p. (2020)

20. Hatos, András ; Hajdu-Soltész, Borbála ; Monzon, Alexander M ; Palopoli, Nicolas ; Álvarez, Lucía ; Aykac-Fas, Burcu ; Bassot, Claudio ; Benítez, Guillermo I ; Bevilacqua, Martina ; Chasapi, Anastasia et al.
DisProt: intrinsic protein disorder annotation in 2020
NUCLEIC ACIDS RESEARCH 48 : D1 pp. D269-D276. (2020)

21. Iserte, J.A. ; Lazar, T. ; Tosatto, S.C.E. ; Tompa, P. ; Marino-Buslje, C.
Chasing coevolutionary signals in intrinsically disordered proteins complexes
SCIENTIFIC REPORTS 10 : 1 Paper: 17962 (2020)

22. Kosol, S. ; Contreras-Martos, S. ; Piai, A. ; Varadi, M. ; Lazar, T. ; Bekesi, A. ; Lebrun, P. ; Felli, I.C. ; Pierattelli, R. ; Tompa, P. ✉
Interaction between the scaffold proteins CBP by IQGAP1 provides an interface between gene expression and cytoskeletal activity
SCIENTIFIC REPORTS 10 : 1 Paper: 5753 (2020)

23. Lazar, T. ; Guharoy, M. ; Vranken, W. ; Rauscher, S. ; Wodak, S.J. ✉ ; Tompa, P.

Distance-Based Metrics for Comparing Conformational Ensembles of Intrinsically Disordered Proteins
BIOPHYSICAL JOURNAL 118 : 12 pp. 2952-2965. , 14 p. (2020)

24. Mészáros, Bálint ; Erdős, Gábor ; Szabó, Beáta ; Schád, Éva ; Tantos, Ágnes ; Abukhairan, Rawan ; Horváth, Tamás ; Murvai, Nikoletta ; Kovács, Orsolya P ; Kovács, Márton et al.
PhaSePro: the database of proteins driving liquid-liquid phase separation.
NUCLEIC ACIDS RESEARCH 48 : D1 pp. D360-D367. (2020)

25. Murvai, Nikoletta ; Kalmar, Lajos ; Szalaine Agoston, Bianka ; Szabo, Beata ; Tantos, Agnes ; Csikos, Gyorgy ; Micsonai, András ; Kardos, József ; Vertommen, Didier ; Nguyen, Phuong N. et al.
Interplay of Structural Disorder and Short Binding Elements in the Cellular Chaperone Function of Plant Dehydrin ERD14
CELLS 9 : 8 Paper: 1856 (2020)

26. Nguyen, H.H. ; Ábrányi-Balogh, P. ; Petri, L. ; Mészáros, A. ; Pauwels, K. ; Vandenbussche, G. ; Keserű, G.M. ; Tompa, P.
Targeting an Intrinsically Disordered Protein by Covalent Modification
METHODS IN MOLECULAR BIOLOGY 2141 pp. 835-854. , 20 p. (2020)

27. Nguyen, Phuong N ; Tossounian, Maria-Armineh ; Kovacs, Denes S ; Thu, Tran T ; Stijlemans, Benoit ; Vertommen, Didier ; Pauwels, Jarne ; Gevaert, Kris ; Angenon, Geert ; Messens, Joris et al.
Dehydrin ERD14 activates glutathione transferase Phi9 in Arabidopsis thaliana under osmotic stress.
BIOCHIMICA ET BIOPHYSICA ACTA-GENERAL SUBJECTS 1864 : 3 Paper: 129506 (2020)

28. Tsytlonok, M. ; Hemmen, K. ; Hamilton, G. ; Kolimi, N. ; Felekyan, S. ; Seidel, C.A.M. ; Tompa, P. ; Sanabria, H. ✉
Specific Conformational Dynamics and Expansion Underpin a Multi-Step Mechanism for Specific Binding of p27 with Cdk2/Cyclin A
JOURNAL OF MOLECULAR BIOLOGY 432 : 9 pp. 2998-3017. , 20 p. (2020)

29. Boeynaems, Steven ; Holehouse, Alex S ; Weinhardt, Venera ; Kovacs, Denes ; Van Lindt, Joris ; Larabell, Carolyn ; Van Den Bosch, Ludo ; Das, Rhiju ; Tompa, Peter S ; Pappu, Rohit V et al.
Spontaneous driving forces give rise to protein-RNA condensates with coexisting phases and complex material properties.
PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA 116 : 16 pp. 7889-7898. , 10 p. (2019)

30. Davey, N.E. ✉ ; Babu, M.M. ✉ ; Blackledge, M. ✉ ; Bridge, A. ; Capella-Gutierrez, S. ; Dosztanyi, Z. ; Drysdale, R. ; Edwards, R.J. ✉ ; Elofsson, A. ; Felli, I.C. ✉ et al.
An intrinsically disordered proteins community for elixir [version 1; peer review: 2 approved]
F1000RESEARCH 8 : ELIXIR Paper: 1753 (2019)

31. Deryusheva, Evgenia ; Nemashkalova, Ekaterina ; Galloux, Marie ; Richard, Charels-Adrien ; Eleouet, Jean-Francois ; Kovacs, Denis ; Van, Belle Karo ; Tompa, Peter ; Uversky, Vladimir ; Permyakov, Sergei
Does Intrinsic Disorder in Proteins Favor Their Interaction with Lipids?
PROTEOMICS 19 : 6 Paper: 1800098 (2019)

32. Gelin, Pierre ; Van, Lindt Joris ; Bratek-Skicki, Anna ; Stroobants, Sander ; Krzek, Marzena ; Ziemecka, Iwona ; Tompa, Peter ; De, Malsche Wim ; Maes, Dominique
Focusing of Microcrystals and Liquid Condensates in Acoustofluidics
CRYSTALS 9 : 3 Paper: 120 (2019)

33. Macossay-Castillo, Mauricio ; Marvelli, Giulio ; Guharoy, Mainak ; Jain, Aashish ; Kihara, Daisuke ; Tompa, Peter ; Wodak, Shoshana J
The Balancing Act of Intrinsically Disordered Proteins: Enabling Functional Diversity while Minimizing Promiscuity.
JOURNAL OF MOLECULAR BIOLOGY 431 : 8 pp. 1650-1670. , 21 p. (2019)

34. Nguyen, Hung Huy ; Tompa, Peter ; Pauwels, Kris
Calpain Purification Through Calpastatin and Calcium: Strategy and Procedures.
METHODS IN MOLECULAR BIOLOGY 1929 pp. 233-244. , 12 p. (2019)

35. Pancsa, Rita ; Schad, Eva ; Tantos, Agnes ; Tompa, Peter
Emergent functions of proteins in non-stoichiometric supramolecular assemblies
BIOCHIMICA ET BIOPHYSICA ACTA-PROTEINS AND PROTEOMICS 1867 : 10 pp. 970-979. , 10 p. (2019)

36. Pancsa, Rita ; Kovacs, Denes ; Tompa, Peter
Misprediction of Structural Disorder in Halophiles.
MOLECULES 24 : 3 Paper: E479 , 15 p. (2019)

37. Szabo, Beata ; Horvath, Tamas ; Schad, Eva ; Murvai, Nikoletta ; Tantos, Agnes ; Kalmar, Lajos ; Chemes, Lucía Beatriz ; Han, Kyou-Hoon ; Tompa, Peter
Intrinsically Disordered Linkers Impart Processivity on Enzymes by Spatial Confinement of Binding Domains.
INTERNATIONAL JOURNAL OF MOLECULAR SCIENCES 20 : 9 Paper: E2119 (2019)

38. Tsytlonok, Maksym ; Sanabria, Hugo ; Wang, Yuefeng ; Felekyan, Suren ; Hemmen, Katherina ; Phillips, Aaron H ; Yun, Mi-Kyung ; Waddell, M Brett ; Park, Cheon-Gil ; Vaithiyalingam, Sivaraja et al.
Dynamic anticipation by Cdk2/Cyclin A-bound p27 mediates signal integration in cell cycle regulation.
NATURE COMMUNICATIONS 10 : 1 p. 1676 (2019)