Tyrosine Residues Increase the Propensity for Liquid-Liquid Phase Separation of the hnRNPA1 Low Complexity Domain

Leanne Kuwahara, Ivan Peran, and Tanja Mittag

Department of Structural Biology

Pediatric Oncology Education Program

St. Jude Children's Research Hospital, Memphis, TN 38105, USA

03 August 2018

ABSTRACT

Liquid-liquid phase separation (LLPS) is a reversible process in which a protein solution de-mixes into liquid protein-dense and protein-light phases. This process is known to be involved with the formation of membrane-less organelles, including stress granules which form in the cell when it undergoes some external stressor. One of the components of stress granules is the RNA-binding protein hnRNPA1, which is known to be recruited into stress granules by undergoing liquid-liquid phase separation. This protein contains a surprisingly compact intrinsically disordered low complexity domain (LCD) which can undergo LLPS on its own. Positively charged and aromatic "sticker" residues within the LCD are thought to mediate this process by forming cation-π and π-stacking interactions. This study demonstrates that there was an increased propensity for phase separation when all of the phenylalanine residues in the LCD are converted to tyrosine, suggesting that the structural differences between aromatic residues impact phase separation.

INTRODUCTION

A fundamental aspect of cellular systems is the cell's ability to organize the multitude of biochemical reactions required for function (Shin & Brangwynne, 2017). Intracellular components are compartmentalized into various canonical membrane-bound organelles, as well as membrane-less organelles, such as the nucleolus, P-bodies, nuclear speckles, Cajal bodies, or stress granules (Shin & Brangwynne, 2017). Membrane-less organelles are known to play a critical role in intracellular organization by sequestering factors not needed by the cell, concentrating reactants, and controlling a dynamic exchange between reactants and products (Shin & Brangwynne, 2017). For instance, cytoplasmic stress granules have been seen to store stalled translational complexes when the cell is under stress, and nuclear membrane-less organelles, such as the nucleolus, have been shown to be important in organizing the genome and ribosome formation. (Shin & Brangwynne, 2017).

These membrane-less organelles are held together by weak, transient multivalent interactions, and it has been observed that these membrane-less organelles form by undergoing liquid-liquid phase separation (Shin & Brangwynne, 2017; Li et. al., 2012). This is a reversible process in which a single-phase protein solution (i.e. dissolved cytoplasm or nucleoplasm) de-mixes into a two-phase regime of liquid protein-rich droplets and the surrounding liquid light phase (Martin et. al., 2018). LLPS has been observed to occur in proteins implicated in protein aggregation diseases, such as Amyotrophic Lateral Sclerosis (ALS) and Frontotemporal Dementia (FTD), and is hypothesized to promote protein fibrillization (Martin et. al., 2018). However, LLPS and fibrillization are two distinct processes, as fibrillization is not required for the formation of membrane-less organelles (Molliex et. al., 2015).

Heterogenous nuclear ribonucleoprotein A1 (hnRNPA1) is a component of stress granules, which are membrane-less organelles that form by undergoing phase separation when the cell undergoes external or environmental stress. hnRNPA1 consists of two RNA recognition motifs (RRMs) that bind to RNA molecules at its N-terminus and an intrinsically disordered low complexity domain (LCD) at its C-terminus (Lin et. al., 2015). For this paper, intrinsically disordered will be defined as a region of a protein that has no unique three-dimensional structure, and low complexity will be defined as a protein region composed of a small subset of amino acids.

hnRNPA1 has also been implicated in neurodegenerative, protein aggregation diseases and is known to be intrinsically prone to form amyloid-like fibrils, which is enhanced in the two-phase regime (Molliex et. al., 2015). When missense mutations occur in the low complexity domain (LCD), the dynamic properties of stress granules are impaired leading to the formation of persistent stress granules, which have been observed in protein aggregation diseases (Ryan et. al, 2018).

It is known that the LCD of hnRNPA1 alone is capable of inducing phase separation, however the RRMs are known to enhance LLPS in the presence of RNA (Molliex et. al., 2015). Previous studies have shown that tyrosine and phenylalanine, as well as positively charged residues, arginine and lysine, play a critical role in the LCD's ability to undergo phase separation by acting at "stickers," whilst polar residues, such as glycine and serine, act as molecular "spacers" that govern the material properties of dense phase droplets (Wang et. al, 2018). Wang et. al. has shown that cation-π interactions between aromatic and positively charged residues are one of the key driving forces for phase separation of the LCD. Furthermore, it was observed that the transient interactions between tyrosine and arginine were significantly stronger than tyrosine-

lysine, phenylalanine-arginine, and phenylalanine-lysine interactions, suggesting that the specific chemical structure of these side chains are important to the LCD's ability to phase separate (Wang et. al, 2018) (Figure 1A and B).

The Mittag lab has studied the role of aromatic residues in liquid-liquid phase separation. A plot of the sequence separation within a protein versus the distance provide information about how compact a chain is, in the way of a scaling exponent. Theoretical scaling exponents have been obtained for an expanded self-avoiding random coil (v=0.59), theta-chain (v=0.50), and globular (v=0.33) conformations (Figure 2A). Small X-ray scattering (SAXS) is a technique that allows one to determine the overall size and shape of biomolecules such as proteins and nucleic acids. Scattering data can provide information about the scaling properties of a chain, and the Mittag lab has used this method to examine the overall compactness of the LCD. Mathematical models (Molecular Form Factors) that fit SAXS data give a scaling exponent for the LCD of hnRNPA1 under denaturing and native conditions (Figure 2B). Figure 2B demonstrates that the LCD of hnRNPA1 is surprisingly compact with a scaling exponent of v=0.45. A graph of contact order vs. residue illustrates that aromatic residues make the most intramolecular contacts within the LCD of hnRNPA1, showing that this compaction is likely due to interactions occurring between aromatic residues (Figure 2C). Intermolecular interactions between aromatic residues has also been observed. NMR Nuclear Overhauser Effects (NOEs) between two residues are observed when they approach each other within several Angstroms. NOEs have been seen between phenylalanine and tyrosine residues, which suggest a transient clustering of these aromatic residues. Furthermore, Figure 2D demonstrates that these intermolecular interactions increase with decreasing temperatures.

When the concentration of protein in solution is high enough, intermolecular interactions between aromatic residues may induce phase separation. A SAXS-Kratky plot illustrates that the scaling exponent increases (indicated by the slope at $q*R_g$ between 2 and 5) as aromatic residues are removed (Figure 2E). Figure 2F illustrates a phase diagram of wildtype protein and a variant in which ⅓ of the aromatic residues have been removed (aro1). The left arm indicates the concentration of protein at which phase separation can occur at a certain temperature. Figure 2F shows that a higher concentration of protein is needed to induce phase separation when a portion of the aromatic residues have been removed. These findings indicate that removing aromatic residues from the LCD lead to a more expanded protein structure, as well as a decreased propensity to undergo phase separation (Figure 2E and F). However, it remains unknown how important the structural differences between tyrosine and phenylalanine are in inducing phase separation. Therefore, this study was conducted to determine what role tyrosine residues have in driving liquid-liquid phase separation of the hnRNPA1 low complexity domain compared to phenylalanine residues.

METHODS

Gateway Cloning and Transformation of Y-only mutant into BL21-DE3 Gold Cells

The gene coding for the Y-only variant of the hnRNPA1-LCD (in which all native phenylalanine residues in the LCD of hnRNPA1 were converted to tyrosine) `with AttL sites was ordered from Thermo Fisher. The coding region was incorporated into the expression vector using the Gateway Cloning reaction, then transformed into BL21-DE3 Gold cells.

Protein Expression and Purification from Inclusion Bodies

The Y-only variant was expressed in auto-induction media at 37°C overnight. The cells were lysed using a sonicator, and the inclusion bodies were isolated by centrifugation of the cell

lysate where the pellet was solubilized in the denaturant Guanidine Hydrochloride (GmdHCl) for two days. The solubilized protein was then spun down and the supernatant was filtered through a 0.8µm filter, then vacuum filtered through 0.4µm filter paper. All purification steps were performed under denaturing conditions. The filtered solution was first loaded onto two 5ml Histrap nickel columns and eluted with a 1M imidazole gradient, then dialyzed into 2M urea with TEV enzyme to cleave the histag. The cleaved protein was then loaded onto two 5ml SP ion-exchange columns and eluted with a 1M NaCl gradient. Lastly, the protein solution was run over an S75 size exclusion column with 4M GmdHCl and concentrated to ~1000µM. The protein can be stored in 4M GmdHCl at 4°C for extended periods of time.

Constructing the Left Arm of a Phase Diagram

The Y-only variant was exchanged into a 20mM HEPES buffer at pH 7. 3M NaCl was added to 500ul of the purified Y-only variant until the final salt concentration was 150mM, in order to induce phase separation. The turbid solution was filtered through a 0.2µm filter and aliquoted in 20µl fractions. Three fractions were then incubated at each of eleven different temperatures (4°C, 6°C, 8°C, 10°C, 12°C, 14°C, 16°C, 8°C, 20°C, 22°C, 24°C), then centrifuged for five minutes at the given temperature to pellet the dense phase. 7µl was then removed from the light phase and three absorbance readings were measured at 280nm using a nanodrop and averaged. Beer's Law was used to calculate the concentration (µM) of protein in the light phase (ε_{280} = 28310 $M^{-1}cm^{-1}$). The incubation temperature was then plotted against the protein concentration using Excel. Error bars were calculated as the standard deviation between the three samples measured at a certain temperature.

RESULTS

Y-only Variant Undergoes Phase Separation at Lower Concentrations Compared to Wildtype

A phase diagram was constructed using Microsoft Excel software to illustrate the temperature and protein concentration thresholds required for the hnRNPA1 LCD to undergo phase separation. Concentrations for the wildtype and aro1 mutant of hnRNPA1 were obtained through a previous study and plotted with the Y-only variant for comparison purposes. Figure 3 demonstrates that the Y-only variant required lower protein concentrations to phase separate compared to the wildtype and aro1 mutant at the same temperatures.

DISCUSSION

The results of this study demonstrated that the Y-only variant has an increased propensity to phase separate compared to the wildtype, illustrating that tyrosine residues in the low complexity domain of hnRNPA1 play a critical role in driving phase separation. This data suggests that tyrosine residues form stronger interactions compared to phenylalanine residues. This data demonstrates that the structural properties of these aromatic residues are important for inducing phase separation. This study shows that phase separation is likely induced through π-π interactions, and the study by Wang et. al. demonstrates that cation-π interactions between tyrosine and arginine residues could also be a contributing factor.

One hypothesis for this result is that the hydroxyl group in the tyrosine side chain is forming hydrogen bonds, which could be leading to stronger interactions between tyrosine and surrounding "sticker" residues. The hydroxyl group of tyrosine has the ability to act as H-bond donor which could increase the binding ability of phenolic ring (Sivasakthi et. al., 2013). Benzene rings are also known to have strong quadrupole moments (Figure 4) therefore, interactions occur between aromatic residues when the π-electron cloud of one residue interacts with the σ-system of another (Figure 4) (Knowles, 2005). These interactions may be enhanced by

resonance structures formed by the hydroxyl group on tyrosine residues compared to phenylalanine, thus giving tyrosine a greater quadrupole moment.

This study will be continued to measure the concentration of the dense phase pellet of the Y-only variant, and polymer theories could be used to fit a model to our data, providing thermodynamic information. We will also construct a phase diagram for an F-only mutant to obtain a more reliable comparison on the impact of the structural differences between aromatic residues.

ACKNOWLEDGEMENTS

Thank you to the Dr. Tanja Mittag, Dr. Ivan Peran, and the entire Mittag lab for your encouragement, guidance, and patience this summer, as well as Dr. Suzanne Gronemeyer, James Marmion, Molly Martin, and Academic Programs for running the POE program.

FIGURES

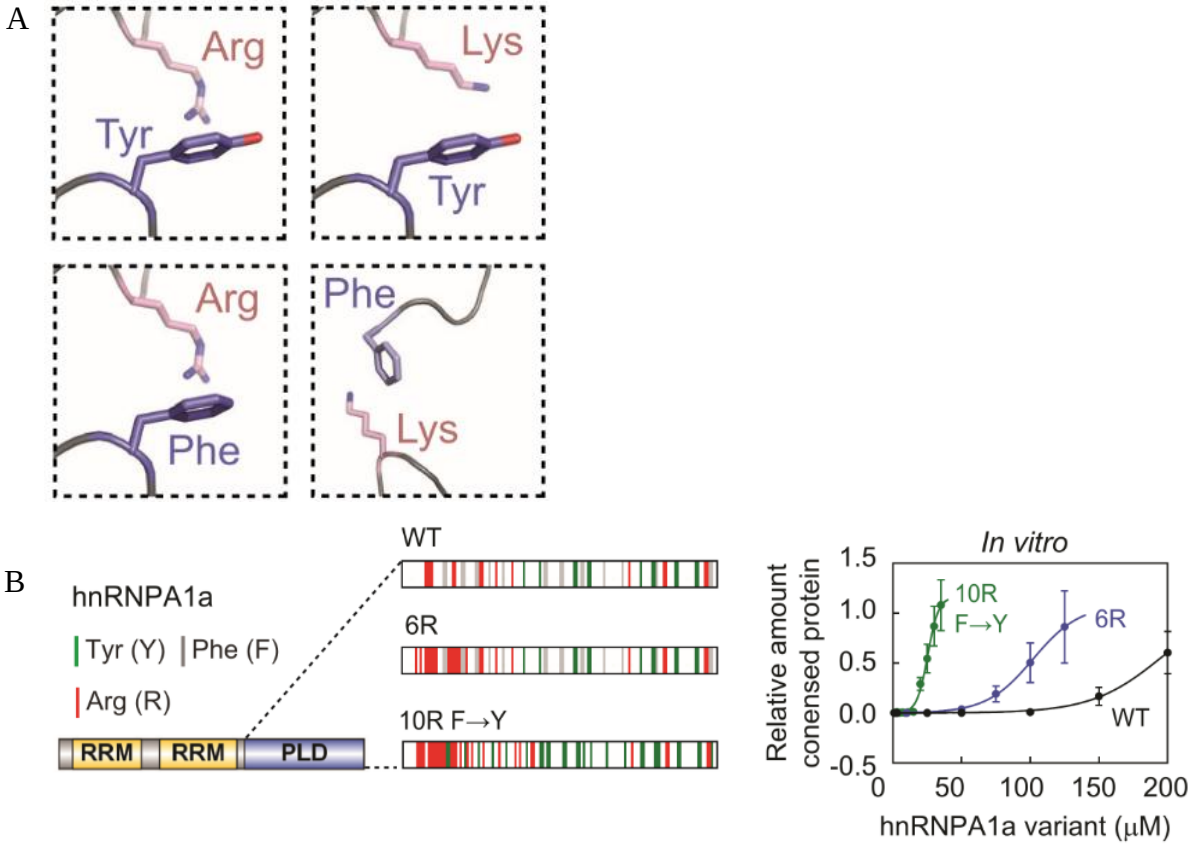

Figure 1. Aromatic and positively charged residues drive phase separation of the hnRNPA1-LCD

(A) Types of cation-pi interactions thought to govern phase separation of low complexity domains.

(B) hnRNPA1 mutants in which either 6 arginine residues were added or 10 phenylalanine residues were converted to tyrosine in the LCD.

Figure 1 was taken from Wang et. al., 2018.

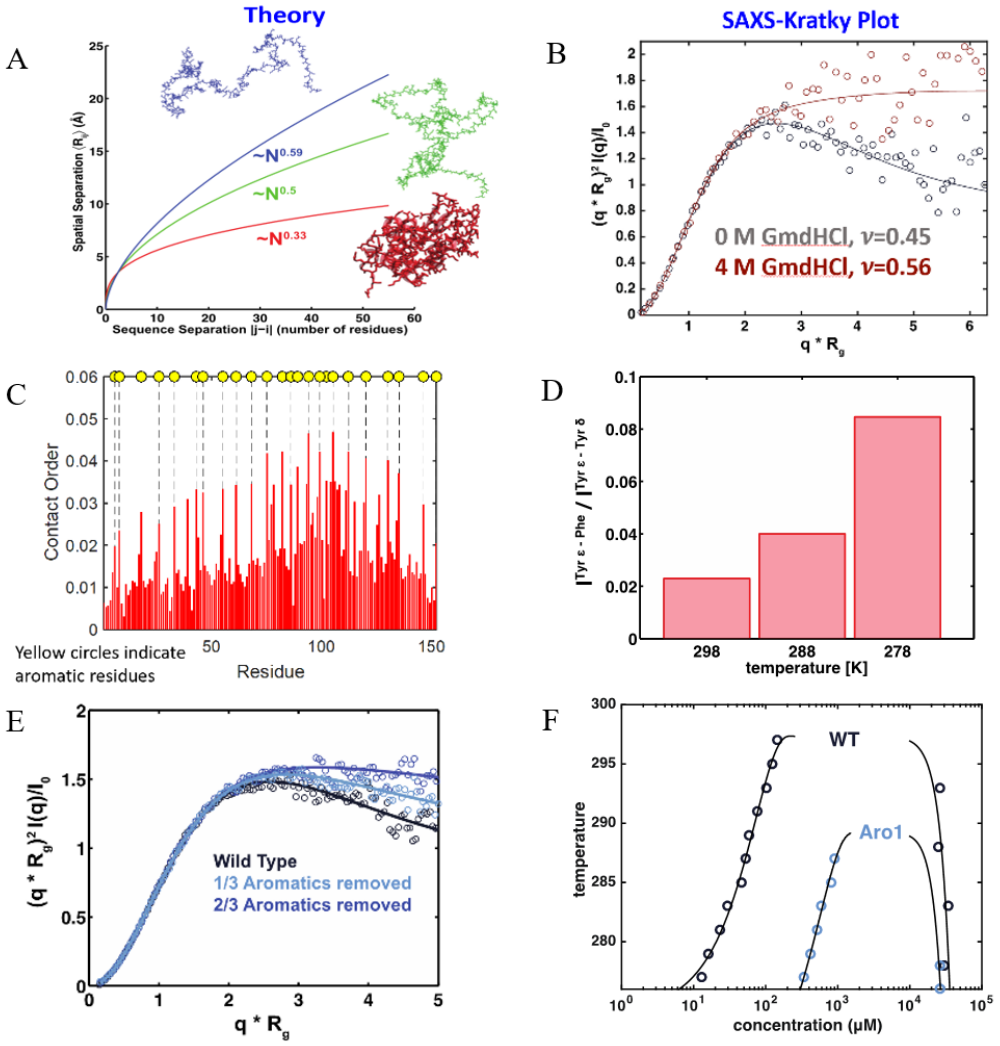

Figure 2. Aromatic residues in the hnRNPA1 low complexity domain make it surprisingly compact

(A) Theoretical graph of spatial separation vs. sequence separation illustrates a scaling exponent for an expanded chain (v=0.59), theta chain (v=0.5), and globular protein (v=0.33).

(B) SAXS data demonstrates that under denaturing conditions the hnRNPA1 LCD has a scaling exponent of 0.56, indicating that the protein conformation is more expanded, while the hnRNPA1 LCD is more compact under native conditions (v=0.45).

(C) Contact order vs. residue graph demonstrates that the aromatic residues make the most intramolecular contacts in the hnRNPA1 LCD.

(D) NMR Nuclear Overhauser Effects (NOEs) are seen between Phenylalanine and Tyrosine residues, indicating transient clustering of aromatics. This clustering is seen to increase with decreasing temperature.

(E) SAXS data shows that removal of aromatic residues within the hnRNPA1 LCD lead to a more expanded conformation, indicated by the change in slope between $q*R_g=2$ and $q*R_g=5$, which is related to the scaling exponent.

(F) Temperature vs. concentration phase diagram illustrates that the aro1 mutant (⅓ of aromatic residues removed) requires a higher protein concentration to undergo phase separation at the same temperature as the wildtype.

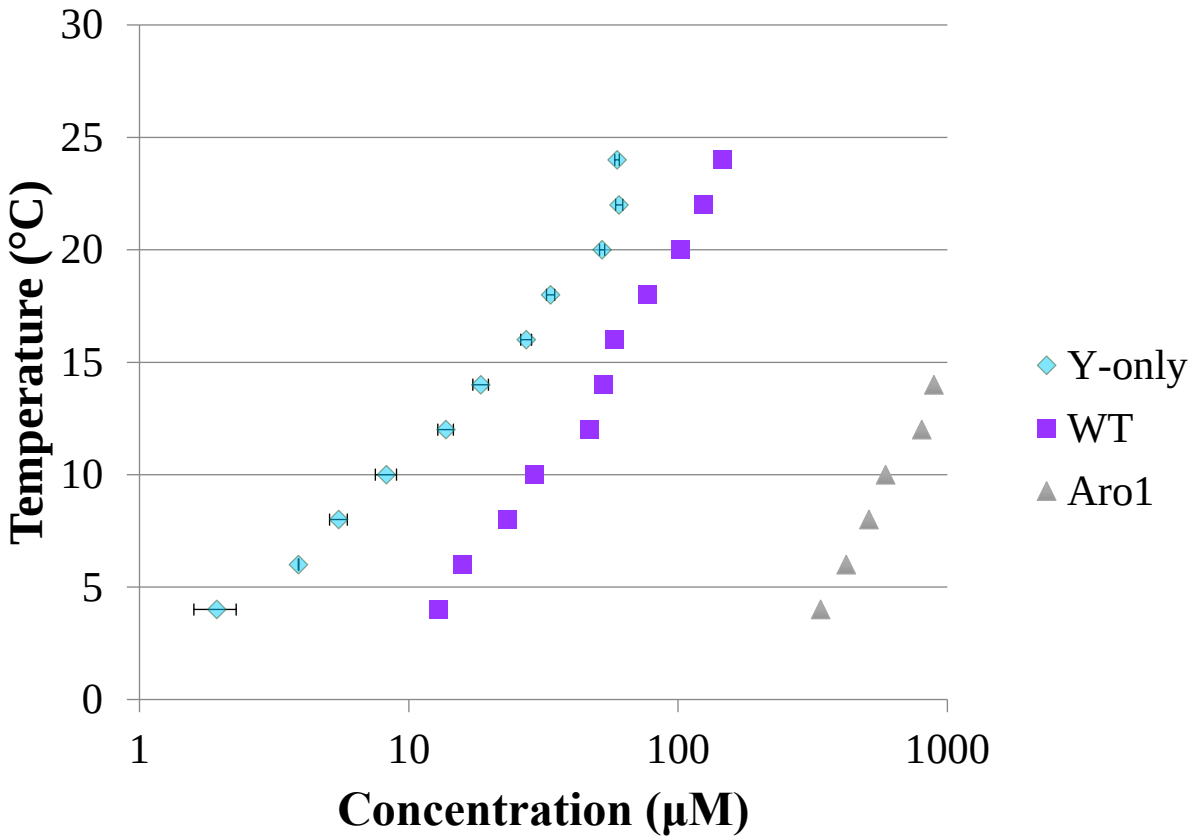

Figure 3. Temperature vs. protein concentration phase diagram of the hnRNPA1 wildtype, aro1 variant, and Y-only variant
The Y-only variant undergoes phase separation at significantly lower concentrations compared to the wildtype and aro1 mutants at the same temperature.

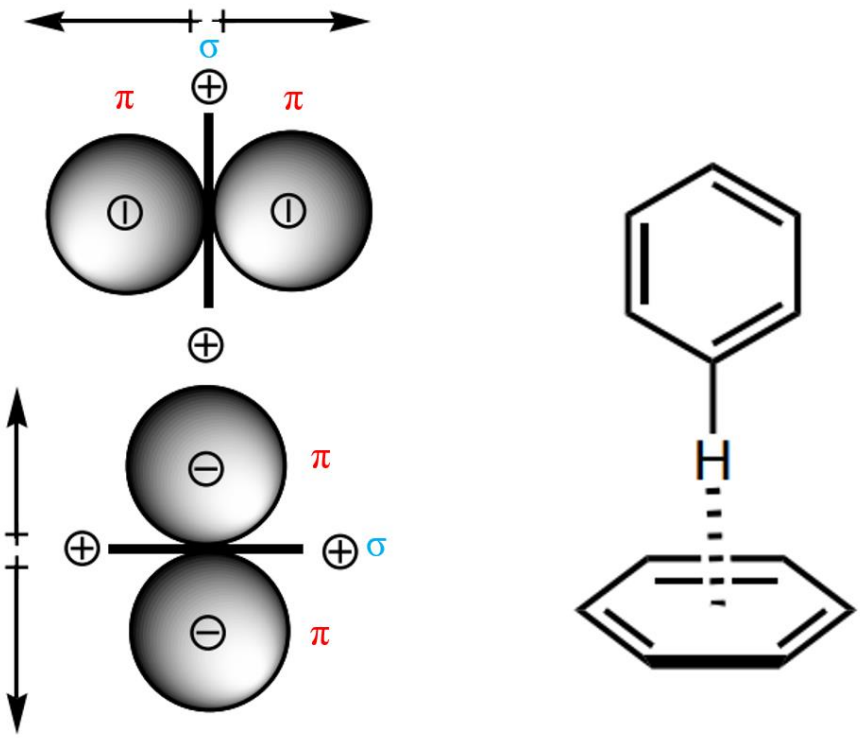

Figure 4. Benzene rings have strong quadrupole moments
A quadrupole moment occurs when two dipoles are aligned such that there is no net dipole moment. π-stacking may occur in a "T-shaped" formation between the π-electron cloud of one aromatic residue and the σ-system of another form electrostatic attractions.
This figure was taken from Knowles, 2005.

REFERENCES

Knowles, R. (2005). *Aromatic Interactions* [PowerPoint slides]. Retrieved from: https://www.princeton.edu/chemistry/macmillan/group-meetings/cation%20pi.pdf

Li, P., Banjade, S., Cheng, H. C., Kim, S., Chen, B., Guo, L., ... & Russo, P. S. (2012). Phase transitions in the assembly of multivalent signalling proteins. *Nature*, *483*(7389), 336.

Lin, Y., Protter, D. S., Rosen, M. K., & Parker, R. (2015). Formation and maturation of phase-separated liquid droplets by RNA-binding proteins. *Molecular cell*, *60*(2), 208-219.

Martin, E. W., Peran, I., & Mittag, T. (2018). The Collapsed Conformational Landscape of the Hnrnpa1 Low Complexity Region Revealed by SAXS, NMR and Simulation. *Biophysical Journal*, *114*(3), 367a-368a.

Molliex, A., Temirov, J., Lee, J., Coughlin, M., Kanagaraj, A. P., Kim, H. J., Mittag, T., & Taylor, J. P. (2015). Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. *Cell*, *163*(1), 123-133.

Ryan, V. H., Dignon, G. L., Zerze, G. H., Chabata, C. V., Silva, R., Conicella, A. E., ... & Fawzi, N. L. (2018). Mechanistic view of hnRNPA2 low-complexity domain structure, interactions, and phase separation altered by mutation and arginine methylation. *Molecular cell*, *69*(3), 465-479.

Shin, Y., & Brangwynne, C. P. (2017). Liquid phase condensation in cell physiology and disease. *Science*, *357*(6357), eaaf4382.

Sivasakthi, V., Anitha, P., Kumar, K. M., Bag, S., Senthilvel, P., Lavanya, P., ... & Ramaiah, S. (2013). Aromatic-aromatic interactions: analysis of π-π interactions in interleukins and TNF proteins. *Bioinformation*, *9*(8), 432.

Wang, J., Choi, J. M., Holehouse, A. S., Lee, H. O., Zhang, X., Jahnel, M., ... & Poser, I. (2018). A Molecular Grammar Governing the Driving Forces for Phase Separation of Prion-like RNA Binding Proteins. *Cell*.