Structural Characteristics and Properties of the RNA-Binding PROTEIN hnRNPK at Multiple Physical States

Quang D Le , Amanda Lewis , Alice D Matthews , Philippe Ringler , Anthony Duff , Andrew E. Whitten , Rob Atkin , Manuel Brunner , Diwei Ho , K Swaminathan Iyer , Andrew C. Marshall , Archa H. Fox , Charles S Bond *

Posted Date: 30 December 2024

doi: 10.20944/preprints202412.2428.v1

Keywords: hnRNPK protein; material state transitions; monomer; aggregate; phase separation; hydrogel; environmental factors; protein characteristics

 doi:10.20944/preprints202412.2428.v1

Article

Structural Characteristics and Properties of the RNA-Binding PROTEIN hnRNPK at Multiple Physical States

Quang D Le [1,2], Amanda Lewis [3], Alice D Matthews [1], Philippe Ringler [3], Anthony Duff [4], Andrew E. Whitten [4], Rob Atkin [1], Manuel Brunner [1], Diwei Ho [1], K. Swaminathan Iyer [1], Andrew C. Marshall [1], Archa H. Fox [1,5] and Charles S. Bond [1]

[1] School of Molecular Sciences, University of Western Australia, Crawley, Western Australia 6009, Australia
[2] Faculty of Biology, VNU University of Science, 334-Nguyen Trai Street, Ha Noi, Viet Nam
[3] Center for Cellular Imaging and NanoAnalytics (C-CINA), Biozentrum, University of Basel, Basel, Switzerland
[4] Australian Centre for Neutron Scattering, Australian Nuclear Science and Technology Organisation, New Illawarra Road, Lucas Heights, NSW 2234, Australia
[5] School of Human Sciences, University of Western Australia, Crawley, Western Australia 6009, Australia

Abstract: Heterogeneous nuclear ribonuclear protein K (hnRNPK) is an RNA-binding protein with low-complexity domains (LCDs) that regulate its behavior under stress conditions. This study demonstrates the ability to control hnRNPK's transitions into four distinct material states—monomer, soluble aggregate, liquid droplet, and fibrillar hydrogel—by modulating environmental factors such as temperature and protein concentration. Importantly, the phase-separated and hydrogel states are newly identified for hnRNPK, marking a significant advancement in understanding its material properties. A combination of biophysical techniques, including DLS and SEC-LS, was used to further characterize hnRNPK in monomeric and soluble aggregate states. Structural methods, such as SANS, SAXS, and TEM, revealed the elongated morphology of the hnRNPK monomer. Environmental perturbations, such as decreased temperature or crowding agents, drove hnRNPK into phase-separated or gel-like states, each with distinct biophysical characteristics. These novel states were further analyzed using SEM, X-ray diffraction, and fluorescence microscopy. Collectively, these results demonstrate the complex behaviors of hnRNPK under different conditions and illustrate the properties of the protein in each material state. Transitions of hnRNPK upon condition changes could potentially affect functions of hnRNPK, playing a significant role in regulation of hnRNPK-involved processes in the cell.

Keywords: hnRNPK protein; material state transitions; monomer; aggregate; phase separation; hydrogel; environmental factors; protein characteristics

1. Introduction

The intracellular environment is densely packed with biological materials [1], leading to macromolecular crowding effects that influence protein structure, function, behavior, and stability [2]. Proteins exhibit dynamic physical states that can change under varying conditions [3], allowing functional flexibility [4]. These transitions, which depend heavily on environmental factors [5,6], enable proteins to shift from their native state to various forms such as solid-like aggregates, gels, or liquid-like droplets [7–9]. This adaptability helps protect proteins from degradation [10] and allows them to function in membraneless compartments or organelles [11–13]. Despite these transitions, proteins are prone to aggregation under stress. To counter this, cells employ molecular chaperones to assist in protein folding and maintain their native states [14,15]. Additionally, low-complexity

 doi:10.20944/preprints202412.2428.v1

domains (LCDs) play a key role in promoting transitions that protect proteins under changing conditions such as temperature, concentration, or pH [16,17].

Heterogeneous nuclear ribonucleoprotein K (hnRNPK), a DNA and RNA-binding protein [18,19], abundant in the nucleus, cytoplasm, and mitochondria [20]. hnRNPK contains low complexity domains (LCDs) identified using the SEG algorithm [21] from PlaToLoCo [22]. As a central signaling hub, hnRNPK regulates various cellular processes through its multiple domains, including three K-homologous (KH) domains, an RGG domain, and a K-interactive (KI) region [23]. hnRNPK has been observed in three material states: monomer, soluble aggregate (uncharacterized), and KH3 domain crystals [24,25].

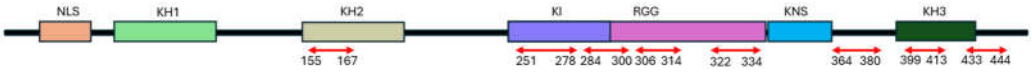


Figure 1. Domain map of full-length hnRNPK used in this study NLS (amino acids 21-37), KH1 (amino acids 46-98), KH2 (amino acids 149-197), KI (amino acids 236-273), RGG (amino acids 236-335), KNS (amino acids 338-361), KH3 (amino acids 391-439); and sequences of hnRNPK low complexity domains in red arrows (bottom of Figure 1).

Environmental stressors, such as heat shock and mutations, can cause proteins to adopt abnormal conformations, exposing hydrophobic regions and leading to irreversible aggregation [26]. The aggregation propensity varies significantly among different proteins. RNA-binding proteins with low complexity domains (LCDs), such as FUS and TDP-43, are particularly prone to aggregation and are linked to neuromuscular disorders like amyotrophic lateral sclerosis (ALS) [27].

The low complexity domains LCDs are regions in proteins composed of a limited subset of amino acids, often resulting in intrinsically disordered structures [28]. These domains can initiate protein aggregation in vitro, as observed with the LCD fragments of U1-70K [29]. In addition to aggregation, LCDs can facilitate liquid-liquid phase separation, underscoring their role in the phase behavior of RNA-binding proteins, which is crucial for the formation of membraneless organelles [9,30]. Moreover, LCDs alone can drive the formation of protein hydrogels, as demonstrated by the prion-like LCDs of FUS and RBM14 forming amyloid-like fibrils under specific conditions [17,31].

This study investigates the transitions of hnRNPK, a protein containing both nucleic acid-binding domains and low-complexity domains (LCDs) that are prone to aggregation, phase separation, and gelation [29–33]. Previous research focused on stabilizing hnRNPK to prevent aggregation, but provided limited insights into its properties and transition from monomer to aggregate [24]. Herein, we employed dynamic light scattering (DLS) to measure changes in hnRNPK's hydrodynamic diameter, revealing that aggregation is a dynamic process influenced by temperature and protein concentration. Light scattering and TEM further characterized the distinct monomeric and aggregated states of hnRNPK, illustrating the measured stoichiometry of novel structures of hnRNPK monomer and aggregate. Additional small-angle X-ray scattering (SAXS) and small-angle neutron scattering (SANS) experiments provided structural details, showing an elongated monomer shape.

Notably, at low temperatures and high concentrations, GFP-hnRNPK formed a hydrogel with an amyloid-like fibril structure, confirmed by scanning electron microscope (SEM) and X-ray diffraction. In the presence of PEG 8000, GFP-hnRNPK phase separated into liquid droplets, observed via fluorescence microscopy. Further experiments demonstrated the vital role of PEG 8000 and salt concentrations on phase behavior of hnRNPK.

Collectively, these results demonstrated the complicated behaviors of hnRNPK, resulting in four distinct material states: disperse monomer, soluble aggregate, liquid condensate, and hydrogel. The emergence of two novel states, hydrogel and liquid condensate, together with structural characterization of monomeric and aggregate states of hnRNPK, suggests the exciting potential of hnRNPK transitions as a cellular mechanism to regulate the protein's functions in the cell.

 doi:10.20944/preprints202412.2428.v1

2. Results

2.1. The Unstable Nature of hnRNPK

Prior study has reported the instability of the hnRNPK protein and its tendency to aggregate in solution [24]. The same aggregation phenomenon was observed during hnRNPK purification in this study. To identify what factors inducing hnRNPK aggregation, DLS was employed to investigate the solution properties of hnRNPK over time and in varying conditions by monitoring the change in the hydrodynamic diameter of the hnRNPK protein. hnRNPK is unstable and quickly aggregates immediately after secondary purification using SEC, therefore freshly purified hnRNPK protein was used to minimize the impact of swift aggregation. To avoid variation between hnRNPK purification, all hnRNPK samples used for DLS experiments came from the same batch of purification.

DLS analysis reveals the behavior of hnRNPK upon temperature change. At 4°C, a 0.5 mg/ml hnRNPK sample has a peak at 43.8 nm (Figure 2a-black line). The same sample incubated at room temperature for one, and three hours shows an increase in size of particles to 68.1 nm (Figure 2a-red line) and 91.3 nm (Figure 2a–blue line), respectively, with reduced peak height and increased peak width. These results show a clear increase in the diameter of hnRNPK particle upon raising the temperature.

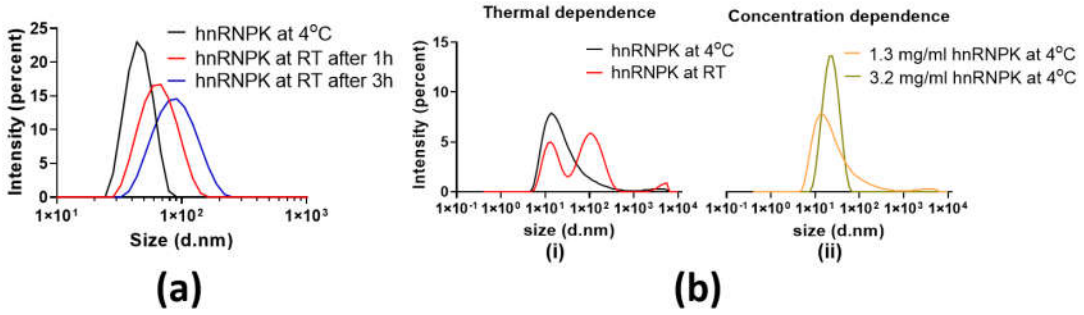


Figure 2. DLS measurement of the hydrodynamic diameter of hnRNPK derived from size distribution by intensity. **(a)** hnRNPK aggregation over time. DLS traces of 0.5 mg/ml hnRNPK at 4°C, at 25°C after 1 hour, and at 25°C after 3 hours. **(b)** Impact of temperature and concentration on hnRNPK aggregation. (i) DLS traces of 1.3 mg/ml hnRNPK at 4°C, and at 25°C after 2.5 hours; (ii) DLS traces of 1.3 mg/ml and 3.2 mg/ml hnRNPK at 4°C.

Having observed the instability of hnRNPK at room temperature, we next used DLS to further quantify the impact of temperature and concentration on freshly purified hnRNPK protein in order to obtain better insight into behaviors of hnRNPK protein upon condition changes.

At 4°C, a freshly purified hnRNPK sample at 1.3 mg/ml had one single peak at 13.5 nm in diameter, followed by a long tail and a small peak of aggregate at 3580.0 nm. At 4°C, the freshly purified hnRNPK protein was substantially smaller (13.5 nm). Upon raising temperature to 25°C, the peak at 13.5 nm was reduced in height, indicating a reduction in the number of particles at this size. At the same time, a major second peak is observed at 105.7 nm, together with another small peak of aggregate at 5560.0 nm (Figure 2b(i)) with clear resolution between three peaks. Intriguingly, hnRNPK protein aggregates changes from a single peak to bimodal distribution in Figure 2b(i), while the shift is monomodal distribution in Figure 2a, suggesting hnRNPK protein might take different modes of self-association during aggregation process.This DLS data indicates the impact of temperature on hnRNPK, increasing the size of hnRNPK particles overtime.

The same protein sample was concentrated from 1.3 mg/ml to 3.2 mg/ml to assess the impact of protein concentration on hnRNPK aggregation. DLS measurements revealed a similar increase in size of hnRNPK upon rising protein concentration at 4°C. After being concentrated from 1.3 mg/ml to 3.2 mg/ml, the main peak moves from 13.5 nm to 21.0 nm with higher peak height and becomes symmetrical as the long tail from the 1.3 mg/ml hnRNPK sample disappears (Figure 2b(ii)). Both samples have small aggregate peaks at 3580 nm for 1.3 mg/ml sample and 5560 nm for 3.2 mg/ml

 doi:10.20944/preprints202412.2428.v1

sample. The substantial increase of the peak height at higher protein concentration indicates the existence of larger species, resulting in stronger scattering intensity. The DLS analysis unravels the significant impact of the temperature and concentration, as two key factors on hnRNPK's behavior, driving the protein into aggregation upon condition changes.

2.2. Monomeric State of hnRNPK

Based on our finding that the temperature and concentration can drive hnRNPK into aggregation, we therefore designed a series of experiments using low concentrations of hnRNPK at low temperature to characterize hnRNPK protein in its monomeric state.

We first used SEC-LS at room temperature to determine the molar mass (MM) and homogeneity of hnRNPK particles when the hnRNPK sample was kept at 4°C. The theoretical MM of the full-length wild-type hnRNPK monomer with an additional glycine residue at the N-terminus, used in this study, is 51033 Da. SEC-LS analysis can be used to compare the measured MM of hnRNPK to the theoretical MM of the measured samples and reveal the stoichiometry of the hnRNPK sample kept at 4°C. The elution profile displays a small shoulder, followed by a major peak with good resolution (Figure 3a). The small shoulder (Peak 1) is eluted at 12.4 ml with an estimated MM of 101.32 kDa (Table 1). The MM distribution shows a wide range (49.0-152.9 kDa), most probably representing a small number of protein stoichiometries (1-3 molecules). The major peak (peak 2) elutes at 13.9 ml, exhibiting a MM value of 59.1 kDa (Table 1), suggesting the majority of particles in peak 2 are hnRNPK monomers. The molar mass points, distributed linearly across the major peak, range from 54.0 kDa to 73.0 kDa. This MM distribution indicates homogeneity of hnRNPK as the protein was kept at low temperature and low concentration. A slight slope of MM distribution across peak 2, indicates that at this condition, there is still a small amount of hnRNPK aggregate coexisting together with the majority of hnRNPK monomer in solution. In combination with DLS, SEC-LS showed that 4°C and low protein concentration was sufficient to maintain the majority of hnRNPK in the monomeric state. Based on these results, we designed SAXS and negative stain transmission electron microscopy (TEM) experiments in an optimal condition to study the structure of hnRNPK monomer.

Table 1. Parameters of SEC-LS for hnRNPK at room temperature and 4°C.

	Peak 1				Peak 2			
	Elution volume (ml)	**Measured MW (kDa)**	**MW distribution (kDa)**	**Stoichiometry**	**Elution volume (ml)**	**Measured MW (kDa)**	**MW distribution (kDa)**	**Stoichiometry**
hnRNPK at 4°C	12.4	101.3	49.0-152.9	1.98:1	14.0	59.1	54.0-73.0	1.16:1
hnRNPK at RT	12.5	146.9	114.8-282.4	2.88:1	13.6	78.0	51.0-138.0	1.53:1

Using SEC-sy-SAXS (Table 2), we observed that while the protein sample was kept at 4°C, it displayed heterogeneity as both aggregate and monomer were present in the sample. The chromatography trace following $I(0)$ for the separation of 6.2 mg/ml hnRNPK at 4°C shows two well separated peaks (Figure 3b). The first peak, with high intensity, is due to large polydisperse aggregated protein as no *Rg* could be established from the scattering data. The second peak, on the other hand, contains scattering data which can be used for further analysis, when background subtracted using buffer-only frames from an appropriate part of the experiment. The MM distribution across peak 2 varies from 40 to 210 kDa, suggesting the presence of multiple oligomers along with monomeric hnRNPK (S1—Supplementary data). The MM variation of hnRNPK at 4°C from SAXS resembles such values observed in SEC-LS, ranging from 49.0–152.9 kDa (Table 2). SAXS samples were subjected for sample delivery process to the Australian Synchrotron. The higher MM of hnRNPK measured from SAXS is perhaps due to the effect of freeze-thaw cycles in comparison to freshly purified hnRNPK immediately used after purification for SEC-LS experiments at 4°C. The

 doi:10.20944/preprints202412.2428.v1

MM distribution of hnRNPK at 4°C measured by SAXS and SEC-LS indicates hetero-oligomerisation of hnRNPK, while multiple oligomeric states of hnRNPK coexist with monomeric hnRNPK in solution.

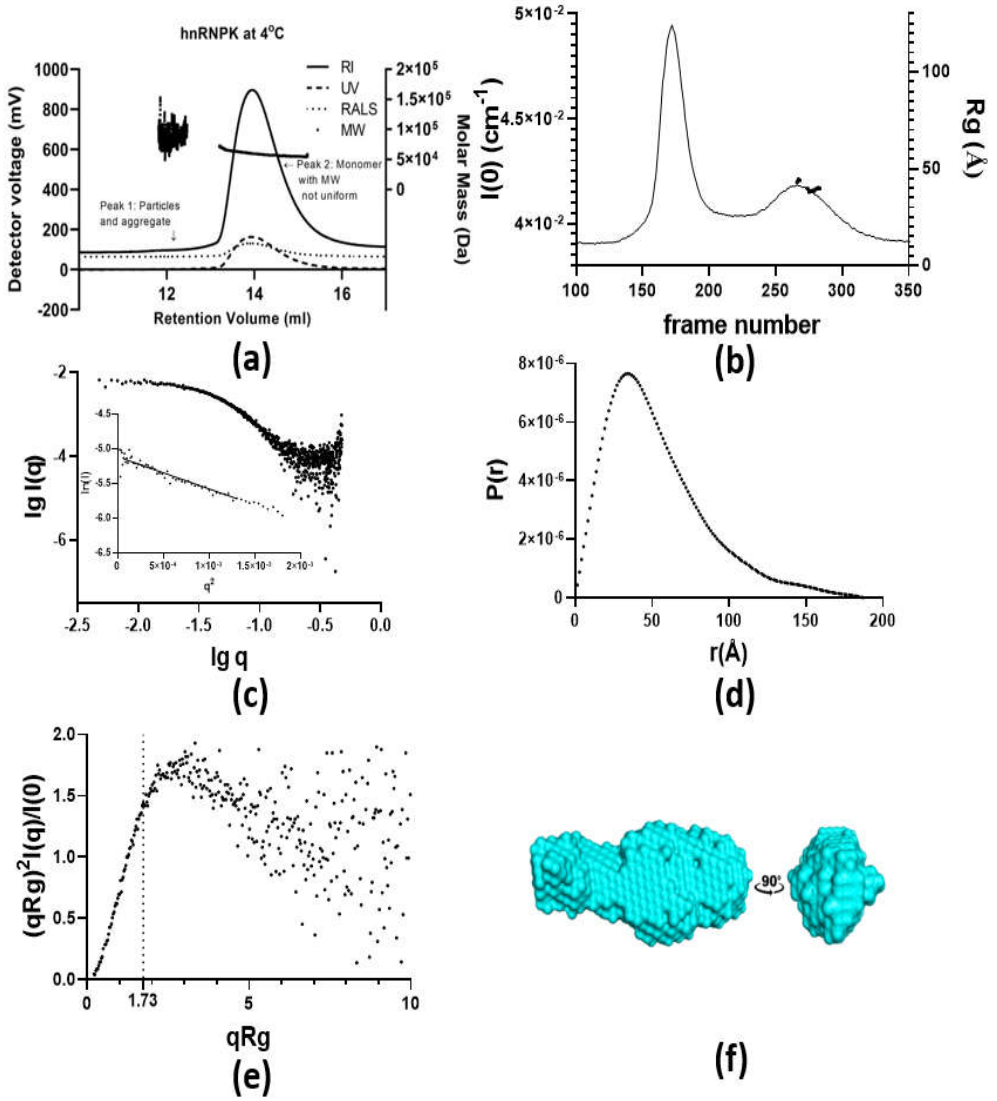


Figure 3. Stoichiometry and structure of hnRNPK at 4°C. **(a)** SEC-LS of hnRNPK at 4°C with RI, UV, RALS traces, and molecular mass distribution. **(b)** Plot of *I*(0) trace for the SEC-SAXS of hnRNPK at 4°C. **(c)** log I(q) vs log q plot and Guinier plot. **(d)** The *P(r)* profile. **(e)** Dimensionless Kratky plot of the scattering data. **(f)** Ab initio model of hnRNPK at 4°C.

SEC-SAXS can resolve heterogeneous sample into multiple populations of elution. The scattering data, from frames 360–379 (buffer), and 273–283 (protein sample), were averaged to study hnRNPK monomer at 4°C. By studying this trait of SEC-SAXS, solution structure of interested homogeneous population can be determined [34]. Due to complex behaviors of hnRNPK sample at 4°C, selecting appropriate scattering frames is a decisive step to analyze solution structure of monomeric hnRNPK, eliminating out any sudden increase in MM value to ensure the homogeneity of hnRNPK monomers [35].

The log I(q) vs log q plot shows the X-ray scattering of hnRNPK particles in the averaged frames from the second peak. The plot is linear and approaching the y-axis horizontally at low angles (Figure 3c). In addition, the Guinier plot of the scattering data displays linear behavior (inset-Figure 3c), despite a minimal up-tick at the lowest resolution, most probably due to the presence of some large particles.

 doi:10.20944/preprints202412.2428.v1

The MM, estimated by Bayesian probability calculation, is one of the primary parameters derived from SAXS analysis to demonstrate the solution state of protein sample [35]. The estimated MM of hnRNPK sample is 54.4 kDa (Table 2) with 6.7% deviation from the theoretical MM of hnRNPK monomer, confirming this set of scattering data derived from hnRNPK particles at their monomeric state.

The *P(r)* distribution shows a smooth, concave curve with a peak at approximately 4.0 nm, and an extended tail that approaches zero at D_{max} of 18.7 nm (Figure 3d). This distribution indicates that the hnRNPK monomer in solution adopts a globular, elongated shape. The interpretation of the extended tail depends on whether higher oligomeric species are contaminating the hnRNPK monomer. In the absence of contamination, given that the selected scattering frames have estimated molecular masses closely aligned with the theoretical molecular mass of the hnRNPK monomer (51 kDa), the globular, compact particles would exhibit a bell-shaped *P(r)* distribution, while the extended tail could be attributed to the flexibility of unfolded particles. Despite careful selection of scattering frames to focus on hnRNPK monomers, the low resolution of size exclusion chromatography implies a potential contamination by higher oligomeric hnRNPK species, with estimated molecular masses ranging from 40 to 210 kDa (Figure S1). The presence of a few hnRNPK oligomers could also explain the extended tail from 11 to 18 nm in the *P(r)* distribution.

The dimensionless Kratky plot initially rises sharply to a peak at low qR_g (ca. 1.8), and then gradually drops down toward the x-axis, suggesting that the hnRNPK monomer possesses an ordered structure with a degree of flexibility (Figure 3e). The plot becomes too noisy to interpret above the qR_g value of 5.

From the scattering data, 20 independent bead models were generated using DAMMIF program [36]. The final normalized spatial discrepancy (NSD) is 0.651 (Table 2), which suggests that most of the models are very similar in shape. The 20 models were averaged using DAMAVER [37], producing a representative ab initio model of hnRNPK monomer with elongated shape as shown in Figure 3f. Moreover, the χ^2 of 0.26 (Table 2), indicates an extremely good fit between the model and scattering data (the error model employed at the SAXSWAXS beamline of the Australian Synchrotron results in the χ^2 of 0.25 representing ideal fit). The ab initio model in Figure 3f is the most accurate solution structure representing either flexible hnRNPK monomers or hnRNPK monomers with mixed scattering signals from a few higher MM hnRNPK oligomers.

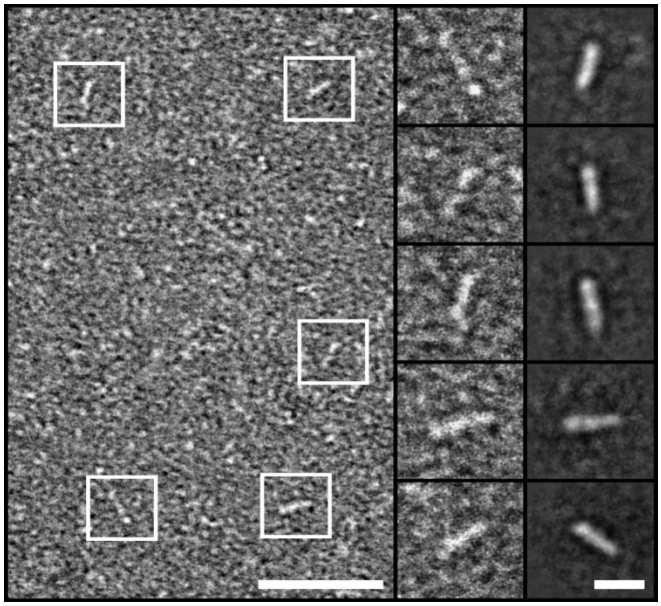
Table 2. SANS and SEC-SAXS derived structural parameters of hnRNPK.

Sample	hnRNPK at 25° C	hnRNPK at 4° C	Hydrogenated hnRNPK monomer	Hydrogenated hnRNPK aggregate
Data Collection Parameters				
Instrument	Australian Synchrotron SAXS/WAXS beamline		QUOKKA, SANS beamline, ANSTO	
q range (Å^{-1})	0.0031-0.0751	0.0080-0.4800	0.0060-0.42	0.0060-0.0824
Concentration (mg/ml)	0.82	6.20	1.1 (unmatched); 26.6 (total)	28.9 (unmatched)
Sample temperature (°C)	22° C	4°C	20° C	20° C
Structural Parameters				
I(0), cm^{-1}, from Guinier	0.1036±0.001000	0.0062±0.000075	0.0692 ± 0.001500	34.1000 ± 0.000880
R_g, Å, from Guinier	106.0±1.40	39.4±2.96	28.8 ± 0.90	88.48 ± 0.01
I(0), cm^{-1}, from *P(r)*	0.109300±0.0010280	0.006347±0.0002068	0.074000 ± 0.0011000	34.120000 ± 0.0008103
R_g, Å, from *P(r)*	113.9±1.710000	44.0±2.695000	32.9 ± 0.600000	88.71 ±0.004509
D_{max}, nm	42.5	18.7	11.2	28.6

 doi:10.20944/preprints202412.2428.v1

P(r) Quality estimate	0.6733	0.6411	0.5840	0.6500
Molecular Mass Estimation by Bayesian Probability				
Estimated Molecular mass, M_r(kDa)	873.1	54.4	67.5	1267.9
MW Probability, %	32.6	21.5		
Credibility Interval (kDa)	614.5, 1013.1	46.2, 63.1		
Credibility Interval Probability, %	93.2	93.1		
Calculated $M_r$from sequence (kDa) †	51.0	51.0		
Estimated Ratio/state†	Aggregate	Monomer	Monomer	Aggregate
DAMMIF (default parameters, 20 calculations)				
Symmetry, anisotropy assumptions	P1, none	P1, none	P1, none	P1, none
χ^2	0.2892	0.2594	0.6888	1454
NSD (standard deviation)	1.094 (0.078)	0.651 (0.037)	0.893 (0.048)	0.708 (0.056)
Resolution (from SASRES) (Å)	97 ± 7	39 ± 3	40 ± 3	85 ± 6

To further characterize the morphology of hnRNPK monomer, we carried out TEM experiments beginning with a low concentration of hnRNPK protein. The preparation of TEM grids with hnRNPK protein was performed on ice to minimize hnRNPK aggregation. TEM images show that at low concentration of 2.48 μM (ca. 0.13 mg/ml), hnRNPK protein particles were well distributed over the grids. The overview panel of Figure 4 reveals raw hnRNPK particles which are highlighted in white boxes. The contrast of images is sufficient to identify and select raw hnRNPK particles from the background as shown in the first column of Figure 4. From the raw particles, 2D averaging was performed to produce five distinct 2D class averages with elongated shapes, each was averaged from around 20 raw particles. The length of these five 2D averages ranges from 10.8 nm to 13.5 nm. The images suggest a structure with 2-3 regions of higher electron density, which may reflect the domain organization of hnRNPK. The overall size and shape of the hnRNPK particles imaged by TEM is in agreement with the solution structure of the hnRNPK monomer SAXS profile.



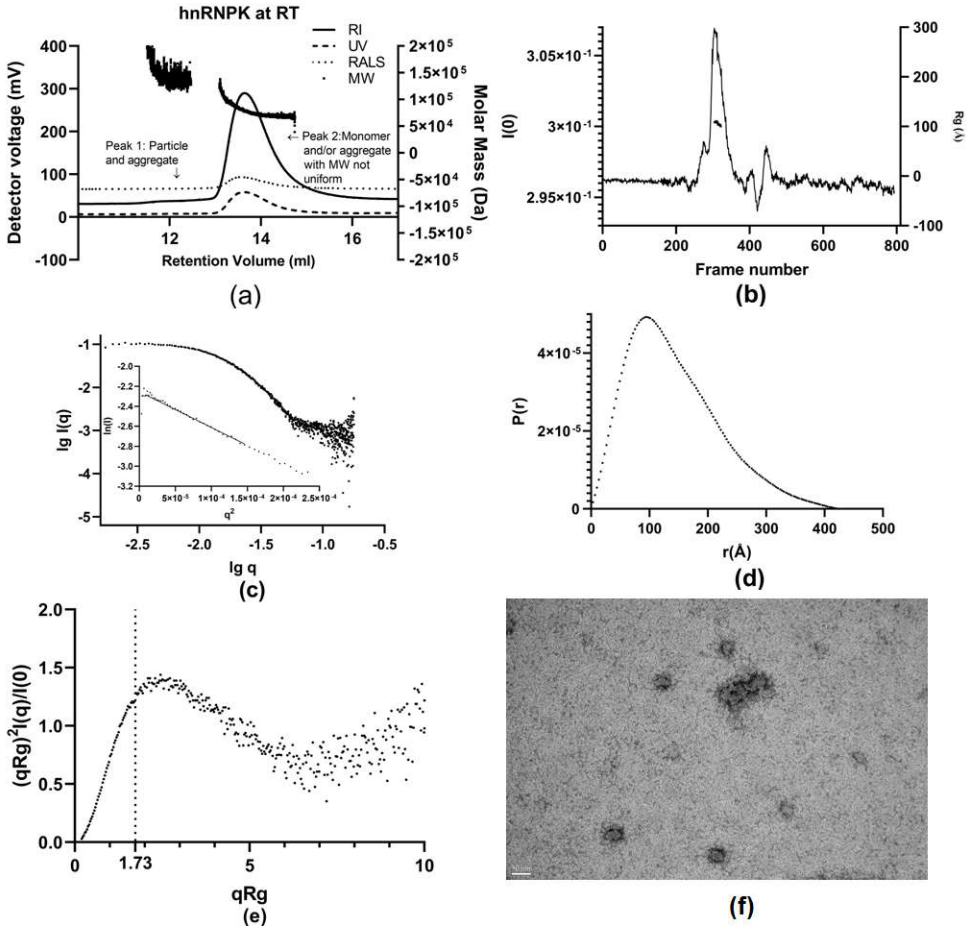
 doi:10.20944/preprints202412.2428.v1

Figure 4. Structural analyses of hnRNPK by negative stain transmission electron microscopy. The left, overview image showing well distributed particles, 5 particles are highlighted (white boxes), scale bar: 50 nm. Insets: First column, 2x enlarged views of the 5 representative particles highlighted in the corresponding overview. Second column, 5 2D class averages, each obtained by averaging about 20 particles. Scale bar: 10 nm, all inset boxes are 25 by 25 nm.

2.3. Aggregation State of hnRNPK

Having characterised the properties and structure of the hnRNPK monomer, we next set out to characterize hnRNPK aggregates induced by incubation at room temperature. SEC-LS was again used to investigate the stoichiometry and homogeneity of hnRNPK, while TEM and SAXS experiments were carried out to study the structure of hnRNPK.

The SEC-LS elution profile of hnRNPK at room temperature indicates an increase of protein aggregates compared to the hnRNPK sample at 4°C (Figure 3a), as larger species can be observed from the chromatogram. The MM distributions of peak 1 and peak 2 overlap each other with poor resolution between two peaks (Figure 5a). Peak 1 (12.5 ml) had a MM of 146.9 kDa, with MM points varying from 114.8 kDa to 282.4 kDa, strongly suggesting polydisperse aggregation. Similarly, the majority of protein particles in peak 2 (13.6 ml) increases in size in comparison to the sample at 4°C (13.9 ml) from 59.1 kDa to 78.0 kDa (Table 1). The MM distribution of peak 2 is markedly uneven as the MM values vary from 138.0 kDa to 51.0 kDa, indicating heterogeneity of sample across peak 2 with higher oligomeric states of hnRNPK formed due to the effect of room temperature on the sample. Together with DLS, SEC-LS analysis illustrates the impact of temperature on hnRNPK aggregation, establishing the specific condition for us to further study the aggregation process of hnRNPK protein.



 doi:10.20944/preprints202412.2428.v1

Figure 5. Stoichimetry and solution structure of hnRNPK at room temperature. **(a)** SEC-LS of hnRNPK at room temperature with RI, UV, RALS traces, and molecular mass distribution. **(b)** Plot of *I*(0) trace for the SEC-SAXS of hnRNPK at room temperature. **(c)** log I(q) vs log q plot and Guinier plot. **(d)** The *P(r)* profile. **(e)** Dimensionless Kratky plot of the scattering data. **(f)** TEM of hnRNPK's aggregate at room temperature, scale bar: 50 nm.

Having observed the aggregation of hnRNPK induced by room temperature, we next used SEC-sy-SAXS to investigate the solution structure of hnRNPK aggregate. The frames used to generate data were 51- 102 for buffer, and 305–321 for protein sample. The *I*(0) trace of hnRNPK at room temperature displays one large peak along three small peaks and a noisy background, (Figure 5b), suggesting heterogeneity for hnRNPK sample at room temperature. The noisy background is perhaps due to a minimal degree of air in the SAXS system causing the issue.

The log I(q) vs log q plot, the primary data of hnRNPK sample at room temperature (Figure 5c), is in linear form and horizontal at low log q values. The scattering data shows linear behavior in the Guinier region (inset-Figure 5c).

The *P(r)* distribution shows a maximum peak at 10.0-11.0 nm and approaches the x-axis smoothly at D_{max} at 41.0 nm (Figure 5d), indicating a substantial increase in size in comparison to the protein sample at 4°C. Despite the fact that the hnRNPK concentration of the room temperature sample was 0.82 mg/ml compared to 6.2 mg/ml of the 4°C sample, the maximum dimension of the *P(r)* from the room temperature sample (Figure 5d) is much higher than the value of the 4°C sample (Figure 3d). The increase in dimension and area under the *P(r)* curve even at lower protein concentration is clear evidence of hnRNPK aggregation at room temperature.

The Kratky plot shows a maximum value of 1.4 at qR_g value of 2.5, dropping down toward the x-axis to form a partly bell shape before rising at qR_g value of 7 (Figure 5e). Given the large size of hnRNPK aggregate, the Kratky plot most probably represents a globular oligomeric structure with some degree of polydispersity.

The established molecular mass of hnRNPK particles at room temperature is 873.1 kDa (Table 2), evidently showing aggregate of hnRNPK formed in the solution. An attempt has been made to generate an ab initio model of hnRNPK aggregate, however the final NSD is 1.1 suggesting a wide variety of classes from the calculations (Table 2). This result indicates the high polydispersity of hnRNPK aggregates, which makes the bead modeling an inappropriate approach to illustrate the structure of hnRNPK aggregate in solution.

In order to further characterize the morphology of the hnRNPK aggregate, TEM was performed, using TEM grids of hnRNPK prepared at room temperature, to visualize the effect of temperature on the overall shape of hnRNPK protein. The TEM images display hnRNPK aggregates as several separated clumps and a big cluster with diameters varying from 22 nm to 200 nm (Figure 5f), which is in agreement with SAXS analysis of hnRNPK aggregate showing a highly polydisperse protein sample in solution. This is consistent with the highly uneven molar mass distribution of hnRNPK recorded by the SEC-LS result, showing the coexistence of multiple higher order hnRNPK species in solution at room temperature.

2.4. Characterization of hnRNPK Monomer in Mixture with Aggregates

Having established the structures of hnRNPK monomer and aggregate, next we performed a SANS experiment to further characterize the structure of monomeric hnRNPK within heterogeneous sample. SANS relies on contrast variation mixing hydrogenated and deuterated hnRNPK (H-hnRNPK:D-hnRNPK) at a certain ratio (H:D), enabling us to determine the structure of hydrogenated components within a multi-component complex. Given the negative impact of high temperature and concentration to hnRNPK monomer, hydrogenated hnRNPK was purified at 4°C and maintained at a concentration of 1.1 mg/ml to minimize aggregation. The optimal H:D ratio was 1:25 with 1.1 mg/ml of hydrogenated hnRNPK mixed with deuterated hnRNPK to achieve a total concentration of 26.6 mg/ml. While the low concentration of 1.1 mg/ml minimized aggregation of H-hnRNPK, it was sufficiently strong to extract the scattering signal of the hydrogenated monomer

 doi:10.20944/preprints202412.2428.v1

from the mixture. At the same time, our previous DLS analysis at room temperature suggests that at 20°C, it is highly possible the 26.6 mg/ml H-hnRNPK:D-hnRNPK mixture will form aggregates due to the impact of high concentration and temperature.

SANS analysis in Figure 6a shows the background subtracted SANS data of hnRNPK from the 1:25 (H:D) mass ratio sample. To prevent scattering signals from protein aggregates overlapping the signals from the hydrogenous hnRNPK sample, the bulk of the protein present was deuterated and contrast matched in buffers containing ~100% D_2O. Hydrogenous protein was visible in D_2O solutions at concentration of 1.1 mg/ml, thus the resulting scattering was from the individual hydrogenous hnRNPK particles distributed randomly throughout the deuterated hnRNPK bulk condensate. In addition, using unmatched buffers, we were able to measure the scattering of the whole aggregate.

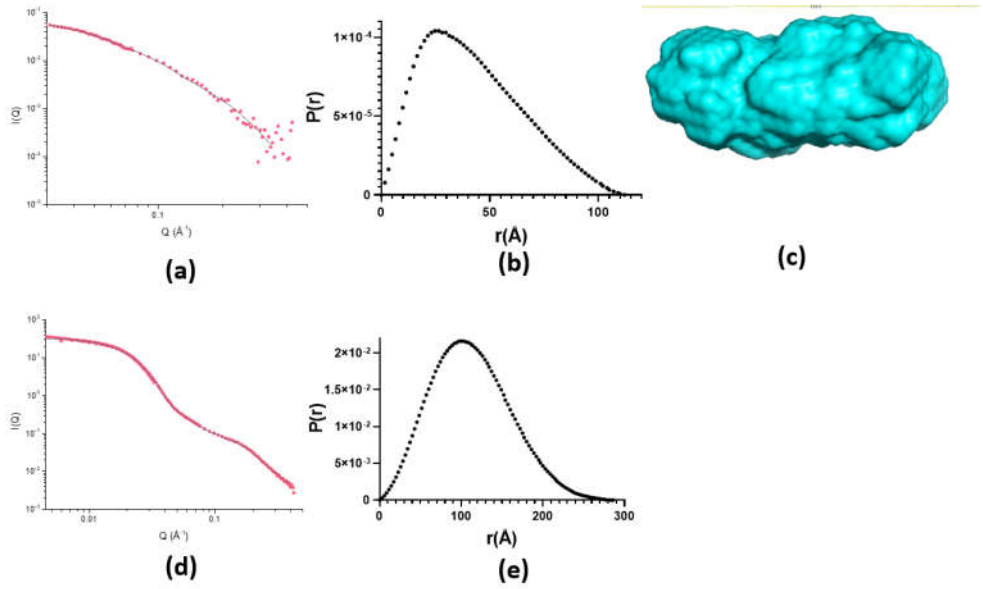


Figure 6. Solution structure of hnRNPK derived from SANS **(a)** Background subtracted SANS data for hydrogenous and deuterated hnRNPK (H:D = 1:25) in deuterated buffer with 1.1 mg/ml of unmatched hnRNPK in a total concentration of 26.6 mg/ml **(b)** The *P(r)* profile of 1:25 (H:D) hnRNPK sample. **(c)** Ab initio model of hnRNPK monomer from SANS scattering profile. **(d)** Background subtracted SANS data for hydrogenous hnRNPK in deuterated buffer at a concentration of 28.9 mg/mL **(e)** The *P(r)* profile of hnRNPK aggregates.

The scattering profile of the 1:25 (H:D) sample in the low q region under 0.02 $Å^{-1}$ is relatively linear. The estimated molecular mass (MM) of protein measured from SANS data is approximately 67.5 kDa (Table 2) [38,39]. The measured MM of the hydrogenous hnRNPK particles is close to the theoretical MM of hnRNPK monomer at 51033 Da, confirming that the H-hnRNPK maintained at low temperature and low concentration was indeed hnRNPK monomer. The *P(r)* distribution of H-hnRNPK has a concave curve shape, with a peak at 25-30 Å and a long smooth tail approaching the x-axis (Figure 6b). The *P(r)* profile of the H-hnRNPK suggests an elongated shape for H-hnRNPK protein, which is slightly more compact compared to hnRNPK monomer measured from SAXS experiment with D_{max} from SANS and SAXS are 11.2 nm and 18.7 nm respectively (Table 2).

The program DAMMIF [36] was used to generate 20 independent ab initio models from the scattering data, revealing an uniform elongated shape of hnRNPK monomers for 20 models. The final normalized spatial discrepancy (NSD) is 0.893 which is relatively high (Table 2), indicating good similarity in shape between models which were then averaged using DAMAVER [37].

SANS analysis of hnRNPK aggregate was carried using 28.9 mg/ml hydrogenated hnRNPK protein. Figure 6c shows the background subtracted SANS data of hnRNPK in deuterated buffer solution. The scattering data have been reduced to log[$I(q)$] vs log(q), descending gradually at low q region. Such scattering profile suggests aggregate formation in hnRNPK sample [40].In addition, the estimated molecular mass (MM) of the 28.9 mg/ml hnRNPK sample measured from SANS data is

 doi:10.20944/preprints202412.2428.v1

approximately 1267.9 kDa (Table 2) [38,39]. The high measured MM of the sample confirms aggregation of the sample at high concentration.

The *P(r)* profile of hnRNPK aggregate has a relatively symmetrical shape, with a short tail approaching x-axis at 28-29 nm (Figure 6e). The *P(r)* distribution of hnRNPK aggregate is in agreement with SASView analysis, suggesting an overall ellipsoidal shape of hnRNPK aggregate with a D_{max} value of 28.6 nm (Table 2).

The SANS analysis is in agreement with the DLS data showing the negative impact of high temperature and concentration on hnRNPK aggregation overtime. Collectively, our DLS and SANS data illustrate the molecular basis of hnRNPK transition from monomer to aggregate. Specifically, SANS data suggests that H-hnRNPK monomer retains its elongated structure when forming larger aggregates due to the impact of the 20°C temperature and the high concentration of 26.6 mg/ml of the 1:25 (H:D) mixture. Although the SANS result suggest a similar elongated shape of hnRNPK monomer, the D_{max}value recorded by SANS is only 11.2 nm, substantially shorter than such values recorded by TEM (up to 13.5 nm) and SAXS (18.7 nm) (Table 2). This shorter D_{max} of SANS may be the result of a large compaction from aggregates on hnRNPK monomers, reducing the length of the monomers. Meanwhile, SANS data of hnRNPK aggregate shows an averaged structure with D_{max} value of 28.6 nm. In consistency with SAXS, and TEM results of hnRNPK aggregate, the SANS model again demonstrates hnRNPK aggregation as an ever-changing process, resulting in a various sizes of hnRNPK aggregates.

2.5. Hydrogel of hnRNPK

hnRNPK contains several low complexity domains (LCDs) within its sequence as demonstrated in Figure 1. Given the ability of LCDs in driving hydrogel formation in protein such as FUS and hnRNPKA2 in vitro [41,42], hnRNPK was tested whether it can form hydrogel as other proteins possessing prion-like LCDs. After being purified, concentrated and dialyzed in gel filtration buffer overnight at 4°C, gelation of eGFP-hnRNPK occurred as shown in Figure 7a. Beside two common material states, monomer and aggregate, hydrogel is a novel material state of hnRNPK observed in in vitro assays. Further experiment to study the biophysical properties of eGFP-hnRNPK hydrogel was performed using X-ray diffraction which reveals 2 diffraction rings at 4.7 Å and 10 Å (Figure 7b). These two diffraction patterns are characteristic signals of the cross-β structure existing in amyloid fibrils [43]. In addition, scanning EM also displays a fibrillar mesh surface of eGFP-hnRNPK (Figure 7c). Together, this data reveals for the first time the ability of eGFP-hnRNPK to form a fibrillar hydrogel in vitro, with amyloid-like structure.

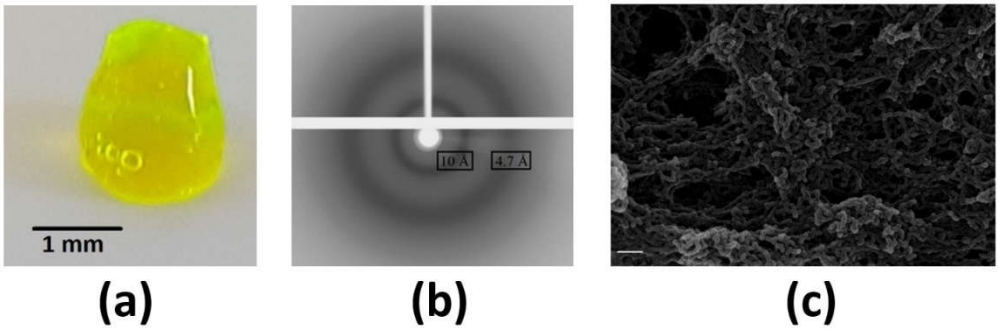


Figure 7. Gelation of eGFP-hnRNPK induced at low temperature and high protein concentration. **(a)** image of eGFP-hnRNPK at hydrogel form. Bar 1 mm**(b)** X-ray diffraction of eGFP-hnRNPK hydrogel with typical amyloid rings at 4.7 Å and 10 Å. **(c)** SEM image of eGFP-hnRNPK hydrogel showing fibrillar structure. Bar 200 nm.

2.6. Phase Separation of hnRNPK

LCDs have been linked to liquid-liquid phase separation (LLPS) of proteins, where the protein can partition into a condensed phase forming liquid-like droplets. LLPS is commonly observed in

 doi:10.20944/preprints202412.2428.v1

nuclear RNA binding proteins, and is implicated in the formation of membraneless organelle structures in the cell nucleus including nucleoli and paraspeckles, of which hnRNPK is a component. To explore the potential of hnRNPK to phase separate, we used a crowding agent, PEG 8000, to induce phase separation of eGFP-hnRNPK. In experiments carried out on ice, dilution of eGFP-hnRNPK into gel filtration buffer at various KCl concentrations produced clear solutions. In contrast, the addition of 10% PEG 8000 in gel filtration buffer turned eGFP-hnRNPK solution turbid, indicating the protein had precipitated in some form (Figure 8a). Fluorescence microscopy of eGFP-hnRNPK confirms the phenomenon as LLPS, with the formation of spherical droplets suspended in the solution (Figure 8b).

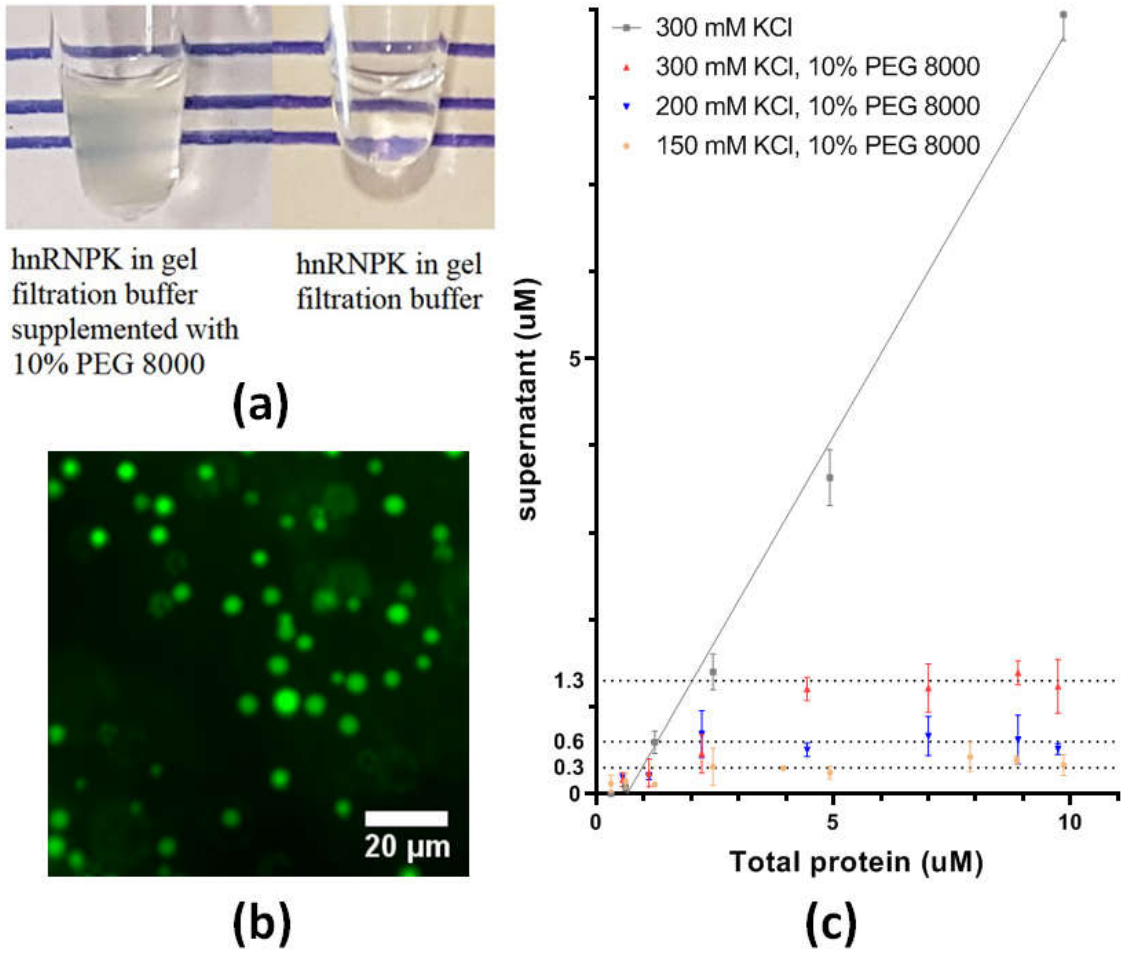


Figure 8. Liquid phase separation of eGFP-hnRNPK. **(a)** Addition of 10% PEG 8000 induces phase separation of hnRNPK. **(b)** Fluorescence microscopy of eGFP-hnRNPK droplets induced by 10% PEG 8000 reveals liquid droplets. **(c)** Measurement of the saturation concentration of eGFP-hnRNPK phase separation induced at three different salt concentratio.ns.

Salt concentration has a crucial role in protein phase separation, influencing phase behaviors of proteins as the salt concentration varies in the solution. To further quantify the effect of PEG 8000 and KCl concentration on eGFP-hnRNPK phase separation, we carried out a centrifuge assay to assess the phase separation induced by PEG 8000 at different KCl concentrations. Upon phase separation, eGFP-hnRNPK co-exists in two separated phases, dilute phase and dense phase, in the solution. Centrifugation sedimented the dense phase from the protein depleted-dilute phase. We measured the concentration of eGFP-hnRNPK in the dilute phase, also known as the saturation concentration (C_{sat}), to evaluate the phase separation of eGFP-hnRNPK at different conditions. In the presence of PEG 8000, eGFP-hnRNPK in all three conditions phase separated to different degrees

 doi:10.20944/preprints202412.2428.v1

(Figure 8c). Specifically, eGFP-hnRNPK, diluted in gel filtration buffer alone (■), shows no sign of phase separation with the measured protein concentration of ca. 10 μM (Figure 8c). In the presence of 10% PEG 8000 (▲), substantial reduction of hnRNPK concentration in the supernatant to 1.3 μM indicates a high level of phase separation induced by PEG 8000. Strikingly, the decrease of the KCl concentration to 200 mM (▼) and 150 mM (●), reduces the C_{sat} of eGFP-hnRNPK in the two conditions to 0.6 μM and 0.3 μM respectively (Figure 8c). The variety of C_{sat} at different KCl concentrations indicates the impact of the salt concentration in phase separation of eGFP-hnRNPK as the lowest KCl concentration at 150 mM depleted the most eGFP-hnRNPK protein out the dilute phase into the dense phase. In comparison, higher KCl concentrations in the other two conditions, 200 mM and 300 mM KCl, allow more eGFP-hnRNPK protein to remain in the dilute phase. Various phase behaviors of eGFP-hnRNPK at different salt concentrations have proven the significance of a stable physiological salt concentration in the cells. Stress may change this balance in the cell which in turn depletes eGFP-hnRNPK from the cellular processes.

3. Discussion

Proteins have been found to exist in multiple material states. The transitions of protein monomer to other physical states, such as protein aggregates or liquid-liquid droplets, play significant roles in the pathogenesis of several neurodegenerative conditions [44,45]. Prior study has demonstrated hnRNPK aggregation in vitro, and successfully used stabilizing additives to reverse and protect hnRNPK monomers from aggregation, with the stoichiometry of hnRNPK monomers confirmed by AUC [24]. In this study, we have demonstrated the substantial impact of temperature and concentration on hnRNPK aggregation. By controlling these two factors, we were able to achieve an optimized condition maintaining freshly purified hnRNPK monomer at 4°C and low concentration to partially prevent hnRNPK aggregate. The beneficial effect of 4°C and low protein concentration enabled us to go beyond stoichiometry value, determining the structure of hnRNPK monomer out of highly heterogeneous hnRNPK sample. The measured stoichiometry data of hnRNPK by SEC-LS at 4°C confirmed the monomeric state of hnRNPK, while TEM and SAXS revealed the novel structure of full-length hnRNPK monomers as elongated shape particles.

Monomer of hnRNPK has been investigated by three distinct techniques in our study, TEM, SAXS, and SANS. Consistently, data from all three approaches reported a similar morphology of hnRNPK proteins as elongated-shape particles. Despite the consistent morphology, the maximum dimensions of hnRNPK measured by TEM, SAXS, and SANS, vary considerably ranging from 11.2 nm (SANS), 18.7 nm (SAXS), and 10.8 nm–13.5 nm (TEM). The variation of D_{max} might be due to different conditions of hnRNPK, causing different extent of interparticle interaction between hnRNPK proteins. In addition, during TEM grid preparation, dehydration of samples can lead to protein shrinkage and subsequently reduce protein length in comparison to D_{max}value measured by SAXS [46]. The beneficial impact of low temperature and concentration has sustained a part of highly heterogeneous hnRNPK sample at its monomeric form, which is the small peak following a strong aggregation peak of hnRNPK in SAXS scattering profile of hnRNPK sample at 4°C (Figure 3b). The heterogeneity of hnRNPK sample at 4°C, as shown in the aggregation peak in SAXS, together with the aforementioned flexibility of hnRNPK possibly caused by LCDs represent a major hurdle for crystallization of full-length hnRNPK protein. Instead, crystallization of hnRNPK truncates containing individual domains, namely KH1, KH2, NLS, KNS domains, might be a more reliable alternative to avoid the issue of full-length hnRNPK's instability. Another approach to solve full-length hnRNPK's structure is cryo-EM, now that we have successfully obtained the low resolution structure of hnRNPK protein using TEM. Addition effort with cryo-EM technique might allow us to make a step further achieving atomic structure of full-length hnRNPK protein.

Despite the beneficial impact of low temperature (4°C) and low protein concentration, hnRNPK aggregate was detected simultaneously with hnRNPK monomer (Figure 3b). Room temperature and high protein concentration, inducing hnRNPK aggregation, allow us to study the volatile changes of hnRNPK protein during aggregation. High temperature and concentration has a significant impact

 doi:10.20944/preprints202412.2428.v1

on hnRNPK, increasing both the hydrodynamic size and molecular mass of hnRNPK overtime. Additionally, a combination of SAXS and TEM showed large structures of hnRNPK aggregates at various sizes, ranging from 22 to 200 nm. As hnRNPK aggregation was irreversible and completely changed the stoichiometry and morphology of hnRNPK sample compared to hnRNPK monomer, we therefore define hnRNPK soluble aggregate as the second material state of hnRNPK protein.

Phase separation of protein is a crucial process to form membraneless compartments or organelles in the cell including stress granules and paraspeckles [47,48]. Many proteins containing low complexity domains (LCDs) have been found in these membraneless organelles [16,29,41]. These proteins can undergo phase separation in vitro upon changes in condition such as temperature, salt concentration or crowding agents [33,49,50], for example FUS, tau, or α-synuclein [51,52]. hnRNPK is an RNA-binding protein with several LCDs in its sequence. This protein has proven to be one of six proteins essential for paraspeckle formation. Indeed, the introduction of PEG 8000 supplemented with protein in solution has resulted in phase separation of hnRNPK, a novel state of this protein up to now. PEG 8000 has proven its effect on affecting the phase behavior of FUS, facilitating condensation of the full-length FUS protein [53]. The crowding agent PEG 8000 is able to induce phase separation of hnRN PK, reorganizing hnRNPK protein into dilute phase and dense phase with substantial protein concentration difference. Additionally, fluorescence microscopy revealed spherical droplets suspending in solution. The round shape, caused by surface tension, indicates the liquid-like property of the droplets [9]. Salt concentration has a significant impact on phase behaviors of many proteins, inducing phase separation on FUS [54,55], TDP-43 [56,57], and BRD4 proteins [58,59]. Despite the fact that low-salt conditions alone does not lead to hnRNPK phase separation, the protein phase separates in a salt concentration-dependent manner in the presence of PEG 8000, reaffirming the importance of salt to hnRNPK phase separation.

The LCDs of RNA-binding proteins have been linked to protein gelation, driving proteins into a hydrogel-like state. Two essential paraspeckle proteins, FUS and RBM14, are among these proteins, forming hydrogel at high concentration and low temperature [31,60]. hnRNPK is another essential paraspeckle protein, containing both LCDs and RNA-binding domains. In our in vitro experiments, hnRNPK has been observed to adopt hydrogel form. Additional studies using X-ray diffraction and SEM revealed the amyloid characteristics of hnRNPK hydrogel which are similar to the hydrogels of FUS and RBM14.

Despite the beneficial effect of low temperature and low concentration, hnRNPK aggregation still occurs with high heterogeneity in solution. The heterogeneity of hnRNPK represents a major hurdle for functional and structural studies of the protein. Therefore, additional effort is required to establish an optimal condition favoring hnRNPK monomer over its aggregation form. Here, we also described two novel material states of hnRNPK, liquid-liquid phase separation, and hydrogel, as well as the properties of hnRNPK protein at each state. Phase separation has a critical role in mediating formation of membraneless organelles such as paraspeckle [47], while hydrogel formation of proteins is an essential cellular process, responsible for both normal physiology and pathology [61]. The ability of hnRNPK to transform from monomer to aggregates, liquid-like droplets, and hydrogel demonstrates the flexibility of hnRNPK to adapt to condition changes, enabling the protein to maintain its functions in a wide range of conditions. Here, we propose a hypothesis that the transitions of hnRNPK protein between different states are protective cellular mechanisms. These mechanisms are employed to regulate multiple hnRNPK-involved cellular processes during stress by transforming hnRNPK protein to appropriate physical state to maintain hnRNPK functions. Therefore, further functional studies of hnRNPK at different physical states, aggregation, phase-separated, hydrogel states, are required to test this theory.

 doi:10.20944/preprints202412.2428.v1

4. Materials and Methods

4.1. Protein Expression and Purification

The full-length wild-type hnRNPK protein with an additional glycine residue at the N-terminus was cloned into plasmid pETM-11, forming a 6xHis-eGFP-hnRNPK fusion protein. The 6xHis-eGFP-hnRNPK protein in hydrogenated media (H-protein) was overexpressed in *Escherichia coli* Rosetta2(DE3) cells in 500 mL of LB media supplemented with 50 μg/mL kanamycin, 34 μg/mL chloramphenicol, and 0.5 mM IPTG at 25°C for 16 hours.

The expression of the deuterated 6xHis-eGFP-hnRNPK protein (D-protein) was carried out in *E. coli* BL21(DE3) using a 2 L bioreactor (Real Time Engineering). For D-protein expression, expression media, with 90% D_2O and hydrogenated glycerol (40 g/L) as a carbon source was used to reach a deuteration level of 75%. The expression culture (LB media), supplemented with 50 μg/mL kanamycin, and 34 μg/mL chloramphenicol, was inoculated with 54 mL of starter culture and grown at 37°C until the optical density at 600 nm reached 37. The expression culture was then induced by 0.5 mM IPTG and incubated at 20°C for 16 hours.

For protein purification, harvested cells were resuspended in lysis buffer [50 mM HEPES, 2 M KCl, 20 mM imidazole, 100 mM dithiothreitol (DTT), EDTA-free protease inhibitor cocktail (Sigma Aldrich), pH 7.4] on ice. Suspended cells were lysed using a high-pressure homogenizer (Emulsiflex C5; Avestin). Lysate was clarified by centrifugation at 24,000 g at 4°C for 45 min. The supernatant was filtered through a 0.22 μm syringe filter (Merck Milipore) before being loaded to a 5 mL Hi-trap column charged with $NiCl_2$ (GE Healthcare). GFP-hnRNPK was eluted with hnRNPK elution buffer (50 mM HEPES, 2 M KCl, 500 mM imidazole, pH 7.4). Peak fractions were pooled, incubated with in-house produced His-tagged TEV protease, and dialyzed in hnRNPK gel filtration buffer (50 mM HEPES, 300 mM KCl, pH 7.4) supplemented with 1 mM DTT for 16 hours at 4°C. The digest was then passed through a 5 mL Hi-trap column to separate the cleaved hnRNPK from the mixture. HnRNPK was concentrated to 5mL [Amicon Ultra Centrifugal filter (EMD Milipore)] and applied to a HiLoad 16/60 Superdex 200 preparative-grade column (GE Healthcare) equilibrated with hnRNPK gel filtration buffer. Purified hnRNPK was concentrated to the desired concentrations for further experiments using a 30 kDaAmicon centrifugal concentrator (Merck Millipore). Purified protein was evaluated by SDS-PAGE stained with Coomassie blue. To achieve 98% deuteration, purified D-protein was dialysed for 6 hours into deuterated size exclusion buffer using a 10 000 MWCO Slide-A-Lyzer dialysis cassette (Thermo Scientific).

4.2. Gelation

GFP-tagged hnRNPK was concentrated to 11.5 mg/ml and dialyzed in buffer (50 mM HEPES, 300 mM KCl, pH 7.4) at 4°C overnight. Dialyzed protein was centrifuged at 3000 x g at 4°C for 10 min to separate soluble fraction and hydrogel. Hydrogel was kept at 4°C for 48 hours in a sealed tube to gelate into a more rigid form.

4.3. X-ray Diffraction

GFP-hnRNPK hydrogel was placed onto a nylon CryoLoop and mounted on an XtaLab Synergy-S X-ray generator (Rigaku Oxford Diffraction). X-ray diffraction imaging was conducted at room temperature with 10 minute exposure time using the *CrystalisPro* software version 1.171.40_64.14a (Rigaku Oxford Diffraction).

4.4. Scanning Electron Microscopy (SEM)

GFP-hnRNPK hydrogel was mounted onto a coverslip coated with poly-L-lysine. The hydrogel was fixed using 2.5% glutaraldehyde at 4°C for 2 hours, then washed with ethanol for dehydration. The dried sample was mounted on a SEM aluminum stub with double –sided carbon tape, and coated with gold. Imaging was carried out using a variable-pressure field-emission scanning electron

 doi:10.20944/preprints202412.2428.v1

microscope (1555, Carl Zeiss; Centre for Microscopy, Characterisation and Analysis, University of Western Australia) at 10 kV.

4.5. Fluorescence Microscopy

GFP-hnRNPK sample was prepared in a series of 20 µL sample solution in the presence of 10% PEG 8000 (50 mM HEPES, 300 mM KCl, pH 7.4, 10% PEG 8000) in a micro 384-well plate. The plate was imaged immediately using an ECLIPSE Ti2 Inverted Research Microscope (Nikon) at room temperature. The images obtained from the fluorescence microscope were processed using FIJI software [62].

4.6. Saturation Concentration Determination by Centrifugation

In order to evaluate the impact of PEG 8000 and salt concentrations to the phase separation of hnRNPK, three conditions containing 10% PEG 8000 in combination with different KCl concentrations (300 mM, 150 mM, or 100 mM KCl) were tested compared to a negative control using size exclusion buffer to dilute hnRNPK sample. After diluting, protein samples were centrifuged at 10 000 x g for 10 min at 4°C to separate the condensed phase from the dilute phase. The hnRNPK absorbance at 280nm of the dilute phase was measured using a NanoDrop Spectrophotometer (Thermo Scientific). The hnRNPK concentration was then calculated using the extinction coefficient at 280 nm of 41830 $M^{-1}cm^{-1}$ obtained from the *ExPasyProtparam* online tool [63]. Each measurement was carried out in triplicate.

4.7. Dynamic Light Scattering (DLS)

Protein and solvents were filtered through 0.22 µm syringe filters (Merck Milipore). Solvents was used later to rinse plastic disposable cuvettes before using, while protein subsequently was centrifuged at 16000 g at 4°C for 15 min to eliminate any insoluble aggregate or contaminants. Plastic disposable cuvettes were pre-cooled on ice to contain samples measured at 4°C. The cuvettes were sealed by parafilm after being filled with samples to avoid introduction of dust or contaminants into solution. To measure the change in size caused by thermal alteration, the same cuvettes were equilibrated at 25°C for at least 30 min before remeasurment. Each measurement was performed with at least three replicates. Dynamic light scattering (DLS) data were recorded using a Malvern ZetasizerNanoZS DLS (Malvern Instruments) and Zetasizer Software v7.11. Scattering light was recorded at 173° backscatter optic to minimize the effect of multiple scattering, dust, and supramolecular contamination on the measurement.

4.8. Size-Exclusion Chromatography Coupled to Light Scattering (SEC-LS)

SEC-LS measurements of hnRNPK were performed at room temperature. The SEC-LS system includes a Superdex 200 Increase 10/300 column (GE Healthcare) for sample separation, ViscotekGPCmax liquid chromatography system, a UV detector 2600 (recording at 280 nm), and a TDA 305 (Malvern Panalytical) (measuring RALS, and RI signals). The sample vial holder temperature was set at 4°C.

4.9. Small-Angle X-ray Scattering (SAXS)

SAXS data were collected at the SAXS/WAXS beamline of the Australian Synchrotron coupled to size exclusion chromatography (SEC-SY-SAXS) [64]. The setup includes an inline Superdex 200 Increase 5/150 column (GE Healthcare) attached to a Shimadzu HPLC system. Samples of hnRNPK were maintained at 4°C or 25°C during measurement. Samples were injected to SEC-SY-SAXS system and eluted directly into a continuous flow cell capillary where scattering data were collected using a Pilatus 1M detector [65]. Data was reduced using ScatterBrain [66]. Primary data processing was performed using the ATSAS software package [67]. SAXS data analysis statistics are presented in Table 1. Additional ensemble optimization were carried out using EOM [68].

 doi:10.20944/preprints202412.2428.v1

4.10. Small-Angle Neutron Scattering (SANS)

SANS measurements were carried out at the Australian Nuclear Science and Technology Organization (ANSTO) OPAL reactor, using the QUOKKA neutron scattering instrument. hnRNPK samples were loaded into quartz Hellma QS-120 cells with a 2 mm pathlength for measurement. Data was collected using a neutron wavelength of 6 Å and a wavelength resolution of $\Delta\lambda/\lambda$= 0.1. Data was reduced using Igor Pro software (Wavemetrics). Scattering data was processed using SASView software [69]. *Further analysis was perform using PRIMUS* [70], *GNOM* [71], *and DAMMIF* [36] *from the ATSAS package* [67].

4.11. Transmission Electron Microscopy (TEM)

For grid preparation, 5µl of the diluted sample were adsorbed to glow-discharged 400 mesh carbon coated Parlodion copper grids washed with three drops of water, incubated with 3 µl of tobacco mosaic virus containing solution (kindly provided by R. Diaz Avalos, Institute of Molecular Biophysics, Florida State University) further washed with two drops of water and finally stained with two drops of 2% (w/v) uranyl acetate. Grids were imaged using a Tecnai12 transmission electron microscope (FEI, Eindhoven, the Netherlands) operating at 120 kV. Electron micrographs were recorded on a 4096 by 4096 pixel charged-coupled device camera (TVIPS F416) at a nominal magnification of ×67 000 yielding a final pixel size of 0.198 nm on the specimen level. Micrographs were processed by reference-free alignment on manually selected particles from micrographs using the EMAN2 image processing package [72]. The extracted particles were aligned and classified by multivariate statistical analysis yielding two-dimensional (2D) class averages.

Author Contributions: My contribution to this work was the mutagenesis, expression of hnRNPK in hydrogenated media, purification, DLS, SEC-LS, SAXS data analysis of hnRNPK, SANS data analysis using *ATSAS*, and TEM data collection of hnRNPK monomer. Moreover, I carried out saturation concentration determination, and fluorescence microscopy of hnRNPK phase separation. Additionally, I performed hnRNPK hydrogel formation experiments, X-ray diffraction of eGFP-hnRNPK hydrogel. Amanda Lewis and Philippe Ringler provided technical advice and performed TEM data analysis of hnRNPK monomer. In addition, Amanda Lewis and Diwei Ho carried out SEM data collection and analysis of eGFP-hnRNPK hydrogel, as well as TEM data collection and analysis of hnRNPK aggregate. Alice D Matthews expressed, purified hnRNPK in hydrogenated media for SANS experiment, and performed SANS data collection. Anthony Duff expressed hnRNPK protein in deuterated media for SANS experiments. Manuel Brunner analysed SANS data using SASView and wrote one part of the first draft of the manuscript on SANS results using SASView. I wrote the manuscript apart for the SANS analysis using SASView, prepared figures and tables, and edited for the finalized version of the manuscript. Charles S. Bond contributed to editing this manuscript. The other authors contributed reagents and/or technical advice.

Funding: This work was supported by the National Health and Medical Research Council (NHMRC APP1147496) and Australian Research Council (ARC DP160102435, LE120100092 and LE140100096). Vietnam International Education Development (VIED) and the University of Western Australia. Bacterial expression vectors pETM11 was a gift from Dr. Gunter Stier (EMBL, Heidelberg, Germany; https://www.embl.de).

Data Availability Statement: SAXS and SANS data is in progress of submission to SASBDB.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

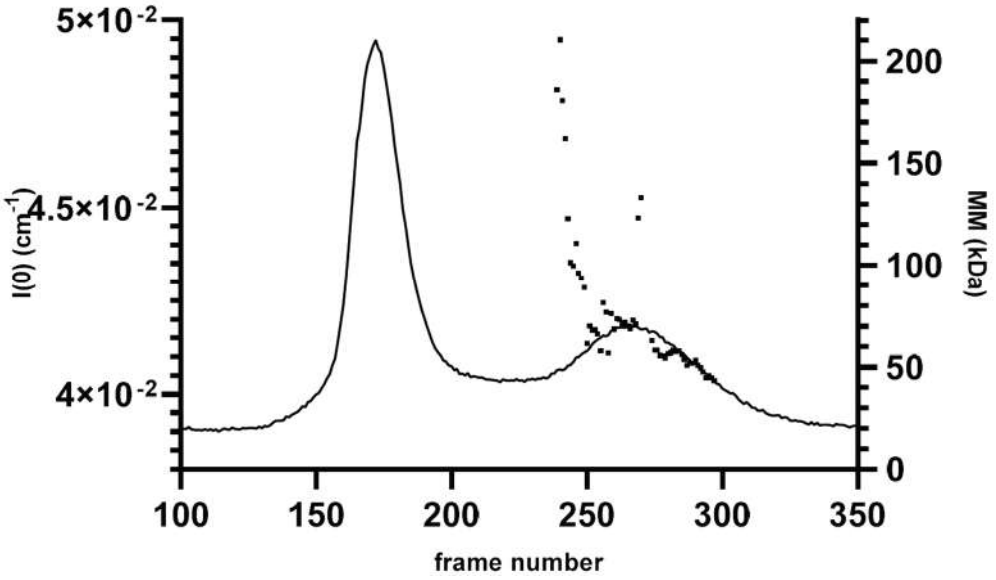
The following abbreviations are used in this manuscript:

hnRNPK	Heterogeneous nuclear ribonucleoprotein K
DLS	Dynamic light scattering
SEC-LS	Size-exclusion chromatography coupled to light scattering
SAXS	Small-angle X-ray scattering

 doi:10.20944/preprints202412.2428.v1

SANS	small-angle neutron scattering
TEM	Transmission electron microscopy
SEM	Scanning electron microscopy
LCD	Low complexity domain
KH	K-homologous domains
KI	K-interactive region
ALS	Amyotrophic lateral sclerosis
DTT	Dithiothreitol
MM	Molecular mass
C_{sat}	Saturation concentration

Appendix



Scheme A1. Scattering profile of hnRNPK at 4°C. Plot of *I*(0) trace and molecular mass distribution from SEC-SAXS of hnRNPK at 4°C. The scattering data were selected from frames 360–371 for buffer, and 239–297 for hnRNPK sample.

References

1. Zimmerman, S.B. and S.O. Trach, *Estimation of macromolecule concentrations and excluded volume effects for the cytoplasm of Escherichia coli.* J Mol Biol, 1991. **222**(3): p. 599-620.
2. Kuznetsova, I.M., K.K. Turoverov, and V.N. Uversky, *What macromolecular crowding can do to a protein.* Int J Mol Sci, 2014. **15**(12): p. 23090-140.
3. Leibly, D.J., et al., *Stabilizing additives added during cell lysis aid in the solubilization of recombinant proteins.* PLoS One, 2012. **7**(12): p. e52482.
4. Foit, L., et al., *Optimizing protein stability* in vivo. Mol Cell, 2009. **36**(5): p. 861-71.
5. Arakawa, T. and S.N. Timasheff, *Theory of protein solubility.* Methods Enzymol, 1985. **114**: p. 49-77.
6. Schmidt, C.L.A., *Proteins, Amino Acids and Peptides as Ions and Dipolar Ions (Cohn, Edwin J.; Edsall, John T.).* Journal of Chemical Education, 1943. **20**(8): p. 415.
7. Rousseau, F., J. Schymkowitz, and L. Serrano, *Protein aggregation and amyloidosis: confusion of the kinds?* Curr Opin Struct Biol, 2006. **16**(1): p. 118-26.
8. Jonker, A.M., D.W.P.M. Löwik, and J.C.M. van Hest, *Peptide- and Protein-Based Hydrogels.* Chemistry of Materials, 2012. **24**(5): p. 759-773.
9. Boeynaems, S., et al., *Protein Phase Separation: A New Phase in Cell Biology.* Trends Cell Biol, 2018. **28**(6): p. 420-435.
10. Díaz-Villanueva, J.F., R. Díaz-Molina, and V. García-González, *Protein Folding and Mechanisms of Proteostasis.* Int J Mol Sci, 2015. **16**(8): p. 17193-230.

 doi:10.20944/preprints202412.2428.v1

11. Banani, S.F., et al., *Biomolecular condensates: organizers of cellular biochemistry.* Nat Rev Mol Cell Biol, 2017. **18**(5): p. 285-298.
12. Shin, Y. and C.P. Brangwynne, *Liquid phase condensation in cell physiology and disease.* Science, 2017. **357**(6357).
13. Alberti, S., *The wisdom of crowds: regulating cell function through condensed states of living matter.* J Cell Sci, 2017. **130**(17): p. 2789-2796.
14. Balchin, D., M. Hayer-Hartl, and F.U. Hartl, In vivo *aspects of protein folding and quality control.* Science, 2016. **353**(6294): p. aac4354.
15. Chiti, F. and C.M. Dobson, *Protein Misfolding, Amyloid Formation, and Human Disease: A Summary of Progress Over the Last Decade.* Annu Rev Biochem, 2017. **86**: p. 27-68.
16. Martin, E.W. and T. Mittag, *Relationship of Sequence and Phase Separation in Protein Low-Complexity Regions.* Biochemistry, 2018. **57**(17): p. 2478-2487.
17. Kato, M., Y. Lin, and S.L. McKnight, *Cross-β polymerization and hydrogel formation by low-complexity sequence proteins.* Methods, 2017. **126**: p. 3-11.
18. Xu, Y., et al., *Post-translational modification control of RNA-binding protein hnRNPK function.* Open Biol, 2019. **9**(3): p. 180239.
19. Nakamoto, M.Y., et al., *hnRNPK recognition of the B motif of Xist and other biological RNAs.* Nucleic Acids Res, 2020. **48**(16): p. 9320-9335.
20. Bomsztyk, K., O. Denisenko, and J. Ostrowski, *hnRNP K: one protein multiple processes.* Bioessays, 2004. **26**(6): p. 629-38.
21. Wootton, J.C. and S. Federhen, *Statistics of local complexity in amino acid sequences and sequence databases.* Computers & Chemistry, 1993. **17**(2): p. 149-163.
22. Jarnot, P., et al., *PlaToLoCo: the first web meta-server for visualization and annotation of low complexity regions in proteins.* Nucleic Acids Research, 2020. **48**(W1): p. W77-W84.
23. Poenisch, M., et al., *Identification of HNRNPK as regulator of hepatitis C virus particle production.* PLoS Pathog, 2015. **11**(1): p. e1004573.
24. Moritz, B., et al., *Biophysical and biochemical analysis of hnRNP K: arginine methylation, reversible aggregation and combinatorial binding to nucleic acids.* Biol Chem, 2014. **395**(7-8): p. 837-53.
25. Backe, P.H., et al., *X-ray crystallographic and NMR studies of the third KH domain of hnRNP K in complex with single-stranded nucleic acids.* Structure, 2005. **13**(7): p. 1055-67.
26. Oromendia, A.B., S.E. Dodgson, and A. Amon, *Aneuploidy causes proteotoxic stress in yeast.* Genes & development, 2012. **26**(24): p. 2696-2708.
27. McAlary, L., et al., *Prion-Like Propagation of Protein Misfolding and Aggregation in Amyotrophic Lateral Sclerosis.* Frontiers in Molecular Neuroscience, 2019. **12**(262).
28. Cascarina, S.M., et al., *LCD-Composer: an intuitive, composition-centric method enabling the identification and detailed functional mapping of low-complexity domains.* NAR Genom Bioinform, 2021. **3**(2): p. lqab048.
29. Xue, S., et al., *Low-complexity domain of U1-70K modulates phase separation and aggregation through distinctive basic-acidic motifs.* Science Advances, 2019. **5**: p. eaax5349.
30. Fritzsching, K.J., et al., *Micellar TIA1 with folded RNA binding domains as a model for reversible stress granule formation.* Proceedings of the National Academy of Sciences, 2020. **117**(50): p. 31832-31837.
31. Hennig, S., et al., *Prion-like domains in RNA binding proteins are essential for building subnuclear paraspeckles.* J Cell Biol, 2015. **210**(4): p. 529-39.
32. Johnson, B.S., et al., *TDP-43 is intrinsically aggregation-prone, and amyotrophic lateral sclerosis-linked mutations accelerate aggregation and increase toxicity.* J Biol Chem, 2009. **284**(30): p. 20329-39.
33. Alberti, S., et al., *A User's Guide for Phase Separation Assays with Purified Proteins.* J Mol Biol, 2018. **430**(23): p. 4806-4820.
34. Rahman, S.K., *Minimizing Heterogeneity of Protein Samples for Metal Transporter Proteins Using SAXS and Metal Radioisotopes.* Methods Mol Biol, 2021. **2199**: p. 151-156.
35. Hajizadeh, N.R., et al., *Consensus Bayesian assessment of protein molecular mass from solution X-ray scattering data.* Scientific Reports, 2018. **8**(1): p. 7204.
36. Franke, D. and D.I. Svergun, *DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering.* J Appl Crystallogr, 2009. **42**(Pt 2): p. 342-346.

 doi:10.20944/preprints202412.2428.v1

37. Volkov, V.V. and D.I. Svergun, *Uniqueness of* ab initio *shape determination in small-angle scattering.* Journal of Applied Crystallography, 2003. **36**(3-1): p. 860-864.
38. Sarachan, K.L., J.E. Curtis, and S. Krueger, *Small-angle scattering contrast calculator for protein and nucleic acid complexes in solution.* Journal of Applied Crystallography, 2013. **46**(6): p. 1889-1893.
39. Loupiac, C., et al., *High-pressure effects on horse heart metmyoglobin studied by small-angle neutron scattering.* European Journal of Biochemistry, 2002. **269**(19): p. 4731-4737.
40. VanOudenhove, J., et al., *Analysis of PKR structure by small-angle scattering.* J Mol Biol, 2009. **387**(4): p. 910-20.
41. Murray, D.T., et al., *Structure of FUS Protein Fibrils and Its Relevance to Self-Assembly and Phase Separation of Low-Complexity Domains.* Cell, 2017. **171**(3): p. 615-627.e16.
42. Xiang, S., et al., *The LC Domain of hnRNPA2 Adopts Similar Conformations in Hydrogel Polymers, Liquid-like Droplets, and Nuclei.* Cell, 2015. **163**(4): p. 829-39.
43. Makin, O.S. and L.C. Serpell, *Structures for amyloid fibrils.* The FEBS Journal, 2005. **272**(23): p. 5950-5961.
44. Ross, C.A. and M.A. Poirier, *Protein aggregation and neurodegenerative disease.* Nat Med, 2004. **10 Suppl**: p. S10-7.
45. Zbinden, A., et al., *Phase Separation and Neurodegenerative Diseases: A Disturbance in the Force.* Dev Cell, 2020. **55**(1): p. 45-68.
46. Sung, J.J., et al., *Transmission electron microscopy as an orthogonal method to characterize protein aggregates.* J Pharm Sci, 2015. **104**(2): p. 750-9.
47. Fox, A.H., et al., *Paraspeckles: Where Long Noncoding RNA Meets Phase Separation.* Trends in Biochemical Sciences, 2018. **43**(2): p. 124-135.
48. Molliex, A., et al., *Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization.* Cell, 2015. **163**(1): p. 123-33.
49. Alberti, S., A. Gladfelter, and T. Mittag, *Considerations and Challenges in Studying Liquid-Liquid Phase Separation and Biomolecular Condensates.* Cell, 2019. **176**(3): p. 419-434.
50. Julius, K., et al., *Impact of Macromolecular Crowding and Compression on Protein–Protein Interactions and Liquid–Liquid Phase Separation Phenomena.* Macromolecules, 2019. **52**(4): p. 1772-1784.
51. Elbaum-Garfinkle, S., *Matter over mind: Liquid phase separation and neurodegeneration.* J Biol Chem, 2019. **294**(18): p. 7160-7168.
52. Stroo, E., et al., *Cellular Regulation of Amyloid Formation in Aging and Disease.* Front Neurosci, 2017. **11**: p. 64.
53. Kaur, T., et al., *Molecular Crowding Tunes Material States of Ribonucleoprotein Condensates.* Biomolecules, 2019. **9**(2).
54. Qamar, S., et al., *FUS Phase Separation Is Modulated by a Molecular Chaperone and Methylation of Arginine Cation-π Interactions.* Cell, 2018. **173**(3): p. 720-734.e15.
55. St George-Hyslop, P., et al., *The physiological and pathological biophysics of phase separation and gelation of RNA binding proteins in amyotrophic lateral sclerosis and fronto-temporal lobar degeneration.* Brain Research, 2018. **1693**: p. 11-23.
56. Wang, A., et al., *A single N-terminal phosphomimic disrupts TDP-43 polymerization, phase separation, and RNA splicing.* EMBO J, 2018. **37**(5).
57. McGurk, L., et al., *Poly(ADP-Ribose) Prevents Pathological Phase Separation of TDP-43 by Promoting Liquid Demixing and Stress Granule Localization.* Mol Cell, 2018. **71**(5): p. 703-717 e9.
58. Han, X., et al., *Roles of the BRD4 short isoform in phase separation and active gene transcription.* Nature Structural & Molecular Biology, 2020. **27**(4): p. 333-341.
59. Sabari, B.R., et al., *Coactivator condensation at super-enhancers links phase separation and gene control.* Science, 2018. **361**(6400).
60. Kato, M., et al., *Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels.* Cell, 2012. **149**(4): p. 753-67.
61. Jean, L., A.C. Foley, and D.J.T. Vaux, *The Physiological and Pathological Implications of the Formation of Hydrogels, with a Specific Focus on Amyloid Polypeptides.* Biomolecules, 2017. **7**(4).
62. Schindelin, J., et al., *Fiji: an open-source platform for biological-image analysis.* Nature Methods, 2012. **9**(7): p. 676-682.

 doi:10.20944/preprints202412.2428.v1

63. Gasteiger, E., et al., *Protein Identification and Analysis Tools on the ExPASy Server*, in *The Proteomics Protocols Handbook*, J.M. Walker, Editor. 2005, Humana Press: Totowa, NJ. p. 571-607.
64. Gully, B.S., et al., *The solution structure of the pentatricopeptide repeat protein PPR10 upon binding atpH RNA.* Nucleic Acids Research, 2015. **43**(3): p. 1918-1926.
65. Kirby, N.M., et al., *A low-background-intensity focusing small-angle X-ray scattering undulator beamline.* Journal of Applied Crystallography, 2013. **46**(6): p. 1670-1680.
66. http://archive.synchrotron.org.au/aussyncbeamlines/saxswaxs/software-saxswaxs, S.
67. Petoukhov, M.V., et al., *New developments in the ATSAS program package for small-angle scattering data analysis.* J Appl Crystallogr, 2012. **45**(Pt 2): p. 342-350.
68. Tria, G., et al., *Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering.* IUCrJ, 2015. **2**(Pt 2): p. 207-17.
69. https://www.sasview.org/, S.
70. Konarev, P.V., et al., *PRIMUS: a Windows PC-based system for small-angle scattering data analysis.* Journal of Applied Crystallography, 2003. **36**(5): p. 1277-1282.
71. Semenyuk, A.V. and D.I. Svergun, *GNOM– a program package for small-angle scattering data processing.* Journal of Applied Crystallography, 1991. **24**(5): p. 537-540.
72. Tang, G., et al., *EMAN2: an extensible image processing suite for electron microscopy.* J Struct Biol, 2007. **157**(1): p. 38-46.