

## The nuclear localization sequence mediates hnRNPA1 amyloid fibril formation revealed by cryoEM structure

Yunpeng Sun<sup>1,2#</sup>, Kun Zhao<sup>1,2#</sup>, Wencheng Xia<sup>1</sup>, Jinge Gu<sup>1,2</sup>, Yeyang Ma<sup>1,2</sup>, Xinrui Gui<sup>1,2</sup>, Xia Zhang<sup>3</sup>, Yanshan Fang<sup>1,2</sup>, Bo Sun<sup>3</sup>, Cong Liu<sup>1,2\*</sup>, Dan Li<sup>4,5\*</sup>

<sup>1</sup>Interdisciplinary Research Center on Biology and Chemistry, Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences, Shanghai 201210, China,

<sup>2</sup>University of Chinese Academy of Sciences, Beijing 100049, China;

<sup>3</sup> School of Life Science and Technology, ShanghaiTech University, Shanghai, 201210, China;

<sup>4</sup> Bio-X-Renji Hospital Research Center, Renji Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai, 200240, China

<sup>5</sup>Bio-X Institutes, Key Laboratory for the Genetics of Developmental and Neuropsychiatric Disorders, Ministry of Education, Shanghai Jiao Tong University, Shanghai, 200030, China;

# These authors contributed equally to this work.

\* To whom correspondence should be addressed. E-mails: [liulab@sioc.ac.cn](mailto:liulab@sioc.ac.cn) (C.L.), [liudan2017@sjtu.edu.cn](mailto:liudan2017@sjtu.edu.cn) (D.L.).

## Abstract

Human heterogeneous nuclear ribonucleoprotein A1 (hnRNPA1) serves as a key regulating protein in RNA metabolism. Malfunction of hnRNPA1 in nucleocytoplasmic transport or dynamic phase separation leads to abnormal amyloid aggregation and neurodegeneration. The low complexity (LC) domain of hnRNPA1 drives both dynamic phase separation and amyloid aggregation. Here, we use cryo-electron microscopy to determine the amyloid fibril structure formed by hnRNPA1 LC domain. Remarkably, the structure reveals that the nuclear localization sequence of hnRNPA1 (termed PY-NLS), which is initially known to mediate the nucleocytoplasmic transport of hnRNPA1 through binding with karyopherin- $\beta$ 2 (Kap $\beta$ 2), represents the major component of the fibril core. The residues that contribute to the binding of PY-NLS with Kap $\beta$ 2 also exert key molecular interactions to stabilize the fibril structure. Notably, hnRNPA1 mutations found in familial amyotrophic lateral sclerosis (ALS) and multisystem proteinopathy (MSP) are all involved in the fibril core and contribute to fibril stability. Our work illuminates structural understandings on the pathological amyloid aggregation of hnRNPA1 and the amyloid disaggregase activity of Kap $\beta$ 2, and highlights the multiple roles of PY-NLS in hnRNPA1 homeostasis.

## Introduction

Heterogeneous nuclear ribonucleoprotein A1 (hnRNPA1) is a multifunctional protein playing key roles in the regulation of gene expression and RNA metabolism<sup>1-3</sup>. It is involved in a wide-range of RNA processing events under normal and pathological cellular conditions including transcription, constitutive and alternative splicing, stability, nuclear export and translation<sup>4-7</sup>. HnRNPA1 contains two globular RNA recognition motifs (RRM1 and RRM2) for the specific binding with mRNA precursor, followed by an unstructured low complexity (LC) C-terminal domain. HnRNPA1 LC consists of a cluster of RGG repeats with interspersed aromatic residues and a nuclear localization sequence (termed PY-NLS)<sup>8-10</sup>. HnRNPA1 LC is intrinsically disordered and exhibits a characteristic feature of homogenous and heterogeneous supermolecular polymerization<sup>11-14</sup>. It can mediate protein-protein and protein-RNA interactions to undergo liquid-liquid phase separation to participate in stress granule assembly<sup>11,12</sup>. The LC domain also exerts a prion-like function to mediate the amyloid aggregation of hnRNPA1, which is causative to neurodegenerative diseases including amyotrophic lateral sclerosis (ALS) and multisystem proteinopathy (MSP)<sup>7,12,15-17</sup>, and enhanced by inheritable mutations<sup>12,15</sup>.

HnRNPA1 predominantly localizes in nucleus, while shuttles between the nucleus and cytoplasm in response to cellular signals<sup>11,18,19</sup>. The transport of hnRNPA1 between nucleus and cytoplasm is mediated by karyopherin- $\beta$ 2 (Kap $\beta$ 2, also known as transportin-1) via direct binding of the PY-NLS sequence with Kap $\beta$ 2<sup>2,9,10</sup>. The complex structure of Kap $\beta$ 2 and PY-NLS reveals a high-affinity binding of PY-NLS in the C-terminal arch of Kap $\beta$ 2<sup>8</sup>. Moreover, recent studies demonstrate that Kap $\beta$ 2 engages PY-NLS or analogous NLSs to inhibit and even reverse the amyloid fibril formation of hnRNPA1 and other RNA-binding proteins<sup>20</sup>.

In this work, we determine the atomic structure of amyloid fibril formed by hnRNPA1 LC domain by cryo-electron microscopy (cryoEM). The structure reveals that the fibril core is mainly formed and stabilized by PY-NLS. Mutations of hnRNPA1 identified in familial ALS and MSP patients are all involved in the fibril core and contribute to fibril stability. This structure not only demonstrates the role of PY-NLS in driving hnRNPA1 amyloid aggregation at the atomic level, and also renders mechanistic understanding for the anti-amyloid activity of Kap $\beta$ 2.

## Results

### CryoEM structure determination of the amyloid fibril formed by hnRNPA1 LC domain

Recombinant hnRNPA1 LC (residues 186-320) was incubated at 16 °C for 2 days to form amyloid fibrils (**Fig. 1a**). The fibrils were fixed on a carbon grid and frozen in liquid ethane. CryoEM data were collected by using a 300 keV Titan Krios microscope. 62,656 fibrils from 3,227 micrographs were selected for the reconstruction of hnRNPA1 LC fibril (**Table 1**). 2D classification showed one dominant species of left-handed twist fibril (~97 % of the selected fibrils) (**Supplementary Fig. 1**). We then performed helical reconstruction by Relion<sup>21</sup> and obtained the 3D density map of the hnRNPA1 LC fibril to an overall resolution of 2.8 Å (**Fig. 1b and Supplementary Fig. 2**). The density map showed two protofilaments intertwining along an approximate 2-fold screw axis to form a left-handed helix with a width of ~7 nm and a half pitch of ~45 nm (**Fig. 1c**). The helical twist and rise between neighboring hnRNPA1 LC subunits are 179.05° and 2.37 Å, respectively (**Fig. 1c and Table 1**). Based on the high-resolution cryoEM density map, we were able to unambiguously build an atomic structure model for hnRNPA1 LC fibril (**Fig. 1d**). Despite that the fibrils were formed by hnRNPA1 residues 186-320, only residues 251-295 were involved in the fibril core (FC) (**Fig. 1d**).

### Overall structural features of the hnRNPA1 LC fibril

HnRNPA1 LC is rich in glycine residues dispersed throughout the sequence. This property is reflected in a high content of random coils in the fibril structure (**Fig. 2a and 2b**). Although the coil regions lack rigid secondary structures, they orderly stack along the fibril axis to enhance the stability of each protofilament (**Fig. 2b and 2c**). In particular, coil 2 at the C terminus swaps to the next subunit to enhance the intra-protofilament stability (**Fig. 2c**); coils 1 and 2 form inter-protofilament contacts further enhancing the stability of the whole fibril (**Fig. 2b and 2c**).

The serpentine fold of hnRNPA1 FC remarkably contains a single but long  $\beta$  strand composed of residues 260-273 (**Fig. 2a and 2b**).  $\beta$  strands of the same protofilament form a parallel  $\beta$  sheet with a kink at <sup>263</sup>FG<sub>264</sub> (**Fig. 2b**). A pair of  $\beta$  sheets mate to provide major inter-protofilament interactions for the fibril formation (**Fig. 2b**). Of note, ordered coil and kinked  $\beta$  sheet structures have also been found in the reversible

fibrils formed by FUS and hnRNPA1 segments<sup>12,22-24</sup>. Thus, these structural features might be general in Gly-rich LC domains.

### **The hnRNPA1 LC fibril features a remarkably large protofilamental interface**

Based on the fibril structure, residues 254-275, which represent half of the FC, are involved in the protofilamental interface (**Fig. 3a and 3b**). The interface renders a large buried surface area of 1,814 Å<sup>2</sup> per layer. According to the local resolution estimation, the cryoEM density in this area is of the highest resolution of the entire fibril structure indicating that the protofilamental interface is the most stable region of the fibril (**Supplementary Fig. 2b**). The interface is tightly packed with a shape complementarity (Sc)<sup>25</sup> of 0.844 excluding any cofactor or water molecules. It is formed by a typical parallel in-register steric-zipper in the middle sandwiched by Phe clusters (**Fig. 3b**). In the steric zipper, the opposing polar side chains of N265, N267, Q269, and S271 are tightly interdigitated to each other via van der Waals forces to form a shape complimentary dry interface (**Fig. 3c**), which is reminiscent of the first atomic structure of steric zipper formed by segment GNNQQNY of yeast prion Sup35<sup>26</sup> (**Supplementary Fig. 3**). Flanking this steric zipper, F263 and the opposing F273 form T-shaped  $\pi$  stacking and further stack along the fibril axis (**Fig. 3a, 3b and 3d**). Outside the  $\pi$  stacking, F254 and the opposing P275 form hydrophobic interactions, which further enlarge the protofilamental interface, and together with the  $\pi$  stacking of F263-F273, stabilize the fibril by blocking the contact of solvent with the hydrophilic steric-zipper interface (**Fig. 3a, 3b and 3e**).

### **PY-NLS is the main component of the fibril core**

PY-NLS is structurally disordered in free hnRNPA1. In the complex with Kap $\beta$ 2, residues 263–289 of PY-NLS forms an extended conformation to line in the C-terminal arch of Kap $\beta$ 2<sup>8</sup> (**Fig. 4a and 4b**). Strikingly, upon amyloid fibril formation, this same region forms the main part of FC (**Fig. 4a and 4c**). In contrast to the overall extended conformation in the complex with Kap $\beta$ 2, PY-NLS forms a more compact structure in the fibril (**Fig. 4d**). F263-F273 further extends to a  $\beta$  strand, referring to which the rest of PY-NLS curls into a serpentine fold (**Fig. 4d**). The distance between the backbones of the terminal residues reduces from 42.0 Å to 22.7 Å.

Moreover, PY-NLS residues that form extensive interactions with Kap $\beta$ 2 also form important intra- and inter-molecular interactions to stabilize the amyloid fibril structure. In detail, segments  $_{263}\text{FGNY}_{266}$  and  $_{273}\text{FGPM}_{276}$  that contact hydrophobic patches on Kap $\beta$ 2 are essential for the formation of the protofilamental interface (**Fig. 4b, 4c, 4e and 4f**). In addition to the inter-protofilamental  $\pi$ -stacking between F263 and F273 (**Fig. 3d and 4e**), and hydrophobic interaction between F254 and P275 (**Fig. 3e and 4f**) mentioned above, Y266 is involved in the inter-molecular electrostatic interaction between D262 and R284, and forms  $\pi$ -cation interaction with R284 (**Fig. 4e**). This triad of D262, Y266 and R284 provides major forces to lock coil 2 with the  $\beta$  strand (**Supplementary Fig. 4**). R284 also contributes to the binding of PY-NLS with Kap $\beta$ 2 by forming salt links with the acidic interface of Kap $\beta$ 2<sup>8</sup> (**Fig. 4b**). Upon Kap $\beta$ 2 binding, F281 and  $_{288}\text{PY}_{289}$  also bind hydrophobic surfaces of Kap $\beta$ 2. PY has been suggested as a consensus motif in various NLSs including hnRNP D, F, M, FUS, HuR, TAP, and PQBP-1, mutation of which validates the importance of this motif in the Kap $\beta$ 2 binding<sup>8,9,27-29</sup>. Upon fibril formation, F281 forms intra- and inter-molecular hydrophobic interactions with the main chains of G279 and G293 (**Fig. 4g**); F289 of the PY motif forms  $\pi$ -stacking with Y295 (**Fig. 4h**). These interactions provide important forces for the folding of coil 2 (**Supplementary Fig. 4**).

Therefore, these structures demonstrate that hnRNPA1 uses the same sequence of PY-NLS for Kap $\beta$ 2 binding and amyloid aggregation. Thus, the binding of PY-NLS with Kap $\beta$ 2 can directly prevent it from driving the amyloid fibril formation of hnRNPA1, which provides structural understanding for the chaperone activity of Kap $\beta$ 2 in the inhibition and disaggregation of hnRNPA1 fibril formation<sup>20</sup>.

### **Disease-causative hereditary mutations in the hnRNPA1 LC fibril structure**

Several ALS- or MSP-causative familial mutations of hnRNPA1 have been identified including D262N/V, N267S, and P288S/A<sup>15-17</sup>. In terms of Kap $\beta$ 2 binding for nucleocytoplasmic transport, the complex structure of PY-NLS and Kap $\beta$ 2 shows the binding of residues 263-289 with Kap $\beta$ 2<sup>8</sup>. Thus, being outside of the binding region, the D262N/V mutation is less likely to influence the cellular localization of hnRNPA1. In contrast, D262 forms salt bridge with R284 in the fibril structure (**Fig. 4e and Supplementary Fig. 4**). However, the role of this salt bridge in stabilizing the fibril structure is arguable. Previous structural study indicates that D262 is one of the key

Asp residues that enable the reversibility of the hnRNPA1 fibril formed during liquid-liquid phase separation via charge repulsion<sup>12</sup>. D262 mutations may remove the charge repulsion and enhance the formation of pathological irreversible fibril<sup>12,30</sup>. Without D262 mutation, wild-type hnRNPA1 is still prone to form irreversible fibrils, in which the charge repulsion of aligned D262 along the fibril is neutralized by the aligned R284 to stabilize the fibril (**Fig. 4e and Supplementary Fig. 5a**). In the context of D262N/V mutation, R284 may still form  $\pi$ -cation interaction with Y266 to stabilize the fibril (**Fig. 4e**). Therefore, the pathology of D262N/V mutation may mainly come from the impairment of reversible fibril formation during phase separation<sup>12</sup>, which consequently enhances pathological irreversible fibril formation.

N267 and P288 are involved in both Kap $\beta$ 2 binding and fibril formation. Upon Kap $\beta$ 2 binding, N267 forms intra- and inter-molecular hydrophilic interactions with neighboring Q269 of hnRNPA1 and N803 of Kap $\beta$ 2 (**Supplementary Fig. 5b**). Upon fibril formation, N267 locates at the center of the protofilamental interface, and together with N265, Q269, and S271 forms steric zipper to stabilize the fibril interface (**Fig. 3c and Supplementary Fig. 3**). Thus, N267S mutation may influence both Kap $\beta$ 2 binding and fibril formation. As for the P288S/A mutation, although P288 is involved in the fibril core, its side chain does not interact with other residues except for stacking along the fibril (**Supplementary Fig. 5a**). While, P288 is essential for Kap $\beta$ 2 binding<sup>8</sup> (**Fig. 4b**). Thus, the P288S/A mutation may not directly influence the fibril formation of hnRNPA1. Rather, it disrupts the proper cellular localization of hnRNPA1<sup>17</sup>, which may lead to cytoplasmic accumulation and amyloid aggregation of hnRNPA1.

## Discussion

HnRNPA1 is one of the most abundant nuclear proteins and shuttles between nucleus and cytoplasm to regulate gene expression and RNA metabolism<sup>2,4,5,31</sup>. It is composed of 320 amino acids, and nearly half of the sequence at the C-terminus is of low complexity and intrinsically disordered. The LC domain underpins the nucleo-cytoplasmic transport, liquid-liquid phase separation, and amyloid aggregation of hnRNPA1<sup>8,11,12,14,15</sup>. Our work reveals that the PY-NLS in the LC domain of hnRNPA1 is key not only for the nucleo-cytoplasmic transport, but also for the self-association of amyloid fibril formation. Intriguingly, hnRNPA1 can form both reversible and irreversible amyloid fibrils. The former is suggested to play a functional role in the

phase separation of hnRNPA1<sup>12</sup>, while the latter is pathological associated with neurodegeneration<sup>11,15</sup>. Our previous work has identified three reversible amyloid-forming core segments (so-called RACs) of hnRNPA1, which share consensus motifs of (N)D(N) and (G)F/Y(G), and form highly reversible amyloid fibrils<sup>12</sup> (**Supplementary Fig. 6**). Truncation or mutation of the key residues in RACs alters the phase separation behavior and stress granule involvement of hnRNPA1<sup>12</sup>. Among the three RACs, RAC3 (residues 260-266) is within the the irreversible fibril core and overlaps with PY-NLS (**Fig. 5a and Supplementary Fig. 6**), which is in line with the previous result showing that truncation of RAC3, but not RAC1 and RAC2, completely eliminates the irreversible amyloid fibril formation of hnRNPA1<sup>12</sup>.

Combining previous and the current work, we hypothesize that hnRNPA1 forms different structures via its LC domain to adapt to different functions. Under physiological conditions, hnRNPA1 shuttles between nucleus and cytoplasm via its PY-NLS forming an extended structure lining in Kap $\beta$ 2 (**Fig. 5b**). In response to cellular stimuli, the LC domain self-associates via multivalent interactions<sup>11,14</sup> and reversible amyloid-like interactions<sup>12,14,30</sup> to drive the liquid-liquid phase separation and incorporation of hnRNPA1 into stress granules (**Fig. 5b**). Note that the intrinsically disordered sequences of other RNA-binding proteins, e.g. TDP43, may also associate with hnRNPA1 for phase separation<sup>13</sup>. The involvement of the LC domain in either the Kap $\beta$ 2 binding or self-association in the dynamic liquid-liquid phase separation may protect hnRNPA1 from irreversible aggregation. However, under pathological conditions such as hereditary mutations, decreased binding affinity with Kap $\beta$ 2 or disruption of phase separation dynamics may expose hnRNPA1 LC, particularly the PY-NLS, and thus lead to the pathological amyloid fibril formation and neurodegeneration (**Fig. 5b**).



## Methods

### Expression and purification of hnRNPA1 LC

The gene encoding hnRNPA1 LC (residues 186-320) was inserted into pET32M vector with a Trx1 tag. HnRNPA1 LC was overexpressed in *E. coli* BL21 (DE3) at 25 °C overnight after induction by 0.4 mM isopropyl-1-thio-D-galactopyranoside (IPTG). Cells were harvested by centrifugation and lysed in the lysis buffer (50 mM Tris-HCl and 6 M guanidine hydrochloride, pH 8.0) at room temperature. Cell lysate was sonicated for 30 min, followed by centrifugation at 30,966 g for 45 min at 4 °C. The supernatant was loaded into Ni column (HisTrap FF, GE Healthcare), and HnRNPA1 LC was eluted with the elution buffer containing 50 mM Tris-HCl, 6 M guanidine hydrochloride and 50 mM imidazole, pH 8.0. The protein was further purified by HPLC (Agilent 1260 Infinity, Agilent Technologies), with the elution buffer containing 20%-80% (v/v) acetonitrile. After freeze-dried, the purified hnRNPA1 LC was freeze-dried and resuspended with 50 mM Tris-HCl, 100 mM NaCl, pH 7.5. The Trx1 tag was cleaved by 3C enzyme with a 50:1 (protein: enzyme) molar ratio at room temperature overnight. After centrifugation at 3,724 g for 30 min at room temperature, the precipitate was dissolved by the lysis buffer. The hnRNPA1 LC protein without tag was obtained by additional purification by HPLC, and freeze-dried.

### Fibril formation

25  $\mu$ M hnRNPA1 LC in the buffer containing 50 mM Tris-HCl, 100 mM NaCl, pH 7.5 was incubated at 16 °C for 2 days without agitation. The fibrils were harvested and dialyzed into 10 mM acetate buffer (pH 5.0) overnight for cryoEM sample preparation.

### CryoEM data collection

HnRNPA1 LC fibrils were loaded onto glow-discharged copper Quantifoil grids (R1.2/1.3, 300 mesh) and then plunge-frozen in liquid ethane precooled by liquid nitrogen using Vitrobot Mark IV(FEI). Micrographs were captured using a Gatan K2 Summit camera on a Titan Krios transmission electron microscope (FEI) operated at 300 kV with a GIF Quantum energy filter (slit width 20 eV), and collected in dose fractionation mode (32 frames per micrograph) by Serial EM software. A magnification of 130,000 $\times$  with physical pixel size 0.52 Å in super-resolution mode was applied. More details are shown in **Table1**.

## **Imaging processing**

Motion correction implement of MotionCorr2 was used to correct beam induced motion of movie frames with dose weighting implement<sup>32</sup>. All 32 frames per micrograph were aligned with 1.356 Å physical pixel size. CTFFIND4.1.8 was then used to estimate the contrast transfer function (CTF)<sup>33</sup>. All filaments were picked manually using manual picking method of RELION3.0<sup>21</sup>.

## **Helical reconstruction**

Helical reconstruction was performed by RELION3.0<sup>21</sup>. A box size of 288 pixels with an inter-box distance of 30 Å was used in extraction method and 801,707 segments were obtained for subsequent reference-free 2D classification. A staggered arrangement was observed in 2D classification results suggesting a pseudo-2<sub>1</sub> symmetry. An initial helical rise of 2.4 Å and helical twist of 179.01° calculated by the apparent crossover distance of filaments were set in 3D classification. 259,615 segments selected after 2D classification were used to construct an initial 3D model de novo. The initial 3D model low-pass filtered to 60 Å was then applied as reference map in 3D classification. Local search of symmetry to optimize of helical twist and rise was not carried out until separation of β-strands along the helical axis was shown in 3D classification. Several rounds of 3D classification with K=3 was performed to remove heterogeneous segments. Optimized parameters and selected segments were applied for 3D refinement and final reconstruction with appropriate reference. Post-processing with a soft-edge solvent mask in 20 % central Z length was performed to sharpen final reconstruction. Fourier shell correlations at FSC = 0.143 criteria were used to estimate resolution of map. A soft spherical mask without masking a 3D helix was used to gain local resolution estimates.

## **Atomic model building and refinement**

The central region of the density map, which exhibits the highest resolution was used to build atomic model de novo in COOT<sup>34</sup>. The model containing 3 layers was generated and then refined by PHENIX using the real-space refinement program<sup>35</sup>.

## **Buried surface area calculation**

The buried surface area was calculated using areaimol from the CCP4 package<sup>36</sup>. with a probe radius of 1.4 Å. The difference between the solvent accessible surface areas of one hnRNPA1 LC chain alone and within the 5-layer fibril structure represents half of the reported buried area.

### **Data availability**

Density maps of the hnRNPA1 LC fibril are available through EMDB with entry code: EMD-30235. The structural model was deposited in the Protein Data Bank with entry code: 7BX7. Structural model of Kap $\beta$ 2 used in this study is available in the Protein Data Bank with entry codes: 2H4M.

### **Acknowledgements**

This work was supported by the Major State Basic Research Development Program (2016YFA0501902 to C. Liu), the National Natural Science Foundation (NSF) of China (91853113 and 31872716 to D. Li and C. Liu), the Science and Technology Commission of Shanghai Municipality (18JC1420500 to C. Liu), Shanghai Pujiang Program (18PJ1404300 to D. Li), the “Eastern Scholar” project supported by Shanghai Municipal Education Commission (to D. Li). The Shanghai Municipal Science and Technology Major Project (2019SHZDZX02 to C. Liu). We acknowledge the support provided by the cryoEM facility of ShanghaiTech University.

### **Author contributions**

D.L., and C.L. designed the project. K.Z., Y.S., and X.G. prepared the hnRNPA1 LC fibril samples. K.Z and X.G. performed the biochemical assays. Y.S., K.Z., and B.S. performed the cryo-EM experiments, built and refined the structure model. All the authors are involved in analyzing the data and contributed to manuscript discussion and editing. D.L. and C.L. wrote the manuscript.

### **Competing interests**

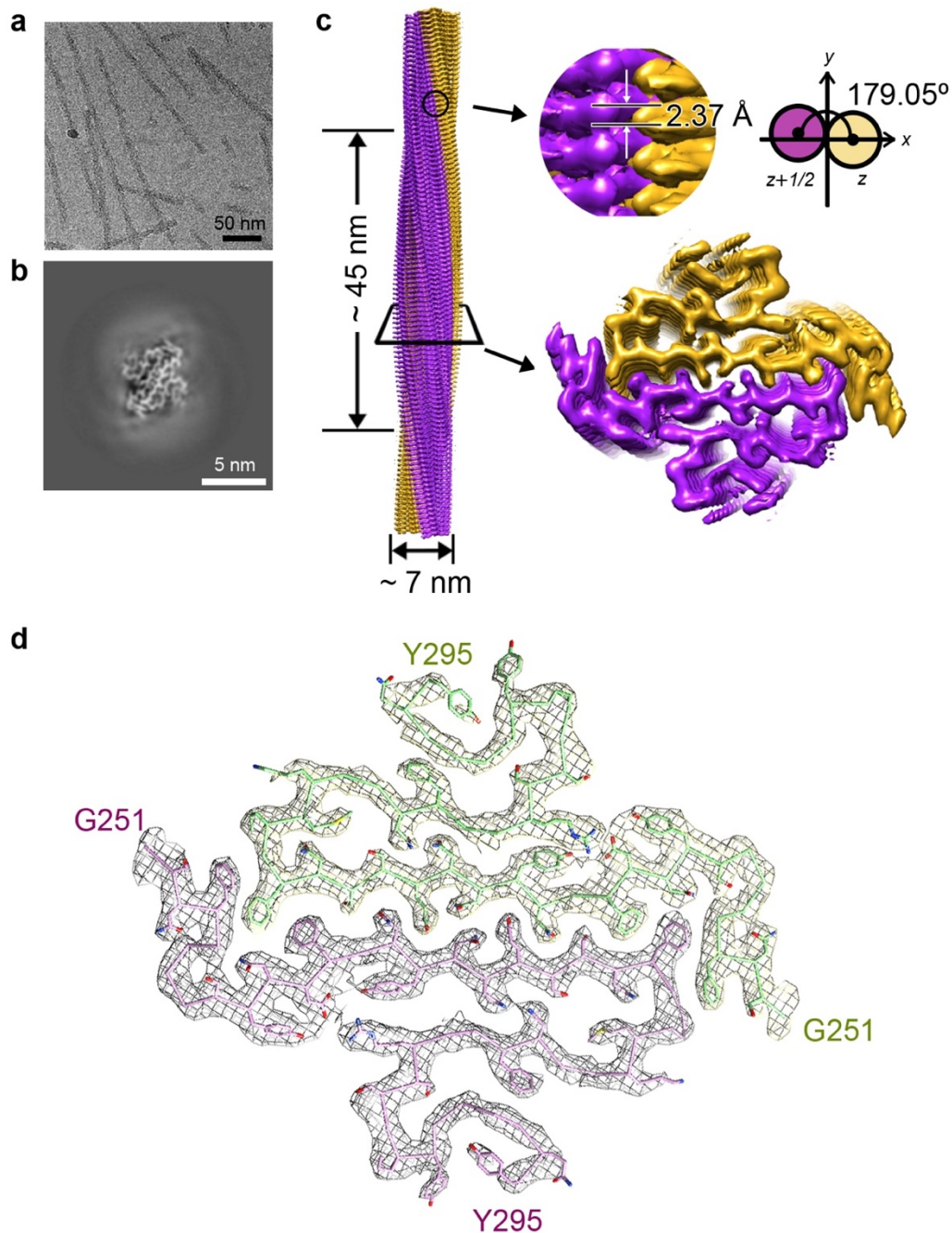
The authors declare no competing interests.

## References

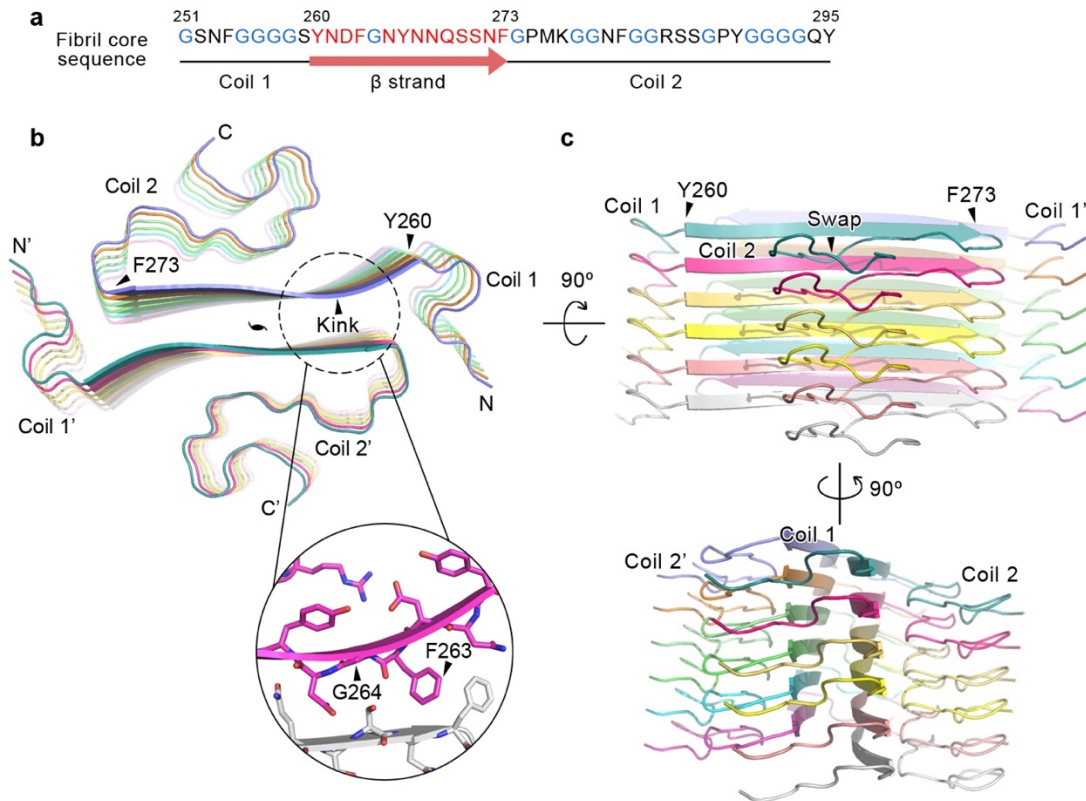
1. Beusch, I., Barraud, P., Moursy, A., Clery, A. & Allain, F.H. Tandem hnRNP A1 RNA recognition motifs act in concert to repress the splicing of survival motor neuron exon 7. *Elife* **6**(2017).
2. Jean-Philippe, J., Paz, S. & Caputi, M. hnRNP A1: the Swiss army knife of gene expression. *Int J Mol Sci* **14**, 18999-9024 (2013).
3. Allemand, E. et al. Regulation of heterogenous nuclear ribonucleoprotein A1 transport by phosphorylation in cells stressed by osmotic shock. *Proc Natl Acad Sci U S A* **102**, 3605-10 (2005).
4. Jain, N., Lin, H.C., Morgan, C.E., Harris, M.E. & Tolbert, B.S. Rules of RNA specificity of hnRNP A1 revealed by global and quantitative analysis of its affinity distribution. *Proc Natl Acad Sci U S A* **114**, 2206-2211 (2017).
5. Guil, S. & Caceres, J.F. The multifunctional RNA-binding protein hnRNP A1 is required for processing of miR-18a. *Nat Struct Mol Biol* **14**, 591-6 (2007).
6. Burd, C.G. & Dreyfuss, G. RNA binding specificity of hnRNP A1: significance of hnRNP A1 high-affinity binding sites in pre-mRNA splicing. *EMBO J* **13**, 1197-204 (1994).
7. Purice, M.D. & Taylor, J.P. Linking hnRNP Function to ALS and FTD Pathology. *Front Neurosci* **12**, 326 (2018).
8. Lee, B.J. et al. Rules for nuclear localization sequence recognition by karyopherin beta 2. *Cell* **126**, 543-58 (2006).
9. Xu, D., Farmer, A. & Chook, Y.M. Recognition of nuclear targeting signals by Karyopherin-beta proteins. *Curr Opin Struct Biol* **20**, 782-90 (2010).
10. Siomi, H. & Dreyfuss, G. A nuclear localization domain in the hnRNP A1 protein. *J Cell Biol* **129**, 551-60 (1995).
11. Molliex, A. et al. Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. *Cell* **163**, 123-33 (2015).
12. Gui, X. et al. Structural basis for reversible amyloids of hnRNPA1 elucidates their role in stress granule assembly. *Nat Commun* **10**, 2006 (2019).
13. Duan, Y. et al. PARylation regulates stress granule dynamics, phase separation, and neurotoxicity of disease-related RNA-binding proteins. *Cell Res* **29**, 233-247 (2019).
14. Xiang, S. et al. The LC Domain of hnRNPA2 Adopts Similar Conformations in Hydrogel Polymers, Liquid-like Droplets, and Nuclei. *Cell* **163**, 829-39 (2015).
15. Kim, H.J. et al. Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS. *Nature* **495**, 467-73 (2013).
16. Naruse, H. et al. Molecular epidemiological study of familial amyotrophic lateral sclerosis in Japanese population by whole-exome sequencing and identification of novel HNRNPA1 mutation. *Neurobiol Aging* **61**, 255 e9-255 e16 (2018).
17. Liu, Q. et al. Whole-exome sequencing identifies a missense mutation in hnRNPA1 in a family with flail arm ALS. *Neurology* **87**, 1763-1769 (2016).
18. Kim, H.J. & Taylor, J.P. Lost in Transportation: Nucleocytoplasmic Transport Defects in ALS and Other Neurodegenerative Diseases. *Neuron* **96**, 285-297 (2017).

19. Guil, S., Long, J.C. & Caceres, J.F. hnRNP A1 relocalization to the stress granules reflects a role in the stress response. *Mol Cell Biol* **26**, 5744-58 (2006).
20. Guo, L. et al. Nuclear-Import Receptors Reverse Aberrant Phase Transitions of RNA-Binding Proteins with Prion-like Domains. *Cell* **173**, 677-692 e20 (2018).
21. Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**(2018).
22. Luo, F. et al. Atomic structures of FUS LC domain segments reveal bases for reversible amyloid fibril formation. *Nat Struct Mol Biol* **25**, 341-346 (2018).
23. Hughes, M.P. et al. Atomic structures of low-complexity protein segments reveal kinked beta sheets that assemble networks. *Science* **359**, 698-701 (2018).
24. Murray, D.T. et al. Structure of FUS Protein Fibrils and Its Relevance to Self-Assembly and Phase Separation of Low-Complexity Domains. *Cell* **171**, 615-627 e16 (2017).
25. Norel, R., Lin, S.L., Wolfson, H.J. & Nussinov, R. Shape complementarity at protein-protein interfaces. *Biopolymers* **34**, 933-40 (1994).
26. Nelson, R. et al. Structure of the cross-beta spine of amyloid-like fibrils. *Nature* **435**, 773-8 (2005).
27. Zhang, Z.C. & Chook, Y.M. Structural and energetic basis of ALS-causing mutations in the atypical proline-tyrosine nuclear localization signal of the Fused in Sarcoma protein (FUS). *Proc Natl Acad Sci U S A* **109**, 12017-21 (2012).
28. Imasaki, T. et al. Structural basis for substrate recognition and dissociation by human transportin 1. *Mol Cell* **28**, 57-67 (2007).
29. Yoshizawa, T. et al. Nuclear Import Receptor Inhibits Phase Separation of FUS through Binding to Multiple Sites. *Cell* **173**, 693-705 e22 (2018).
30. Murray, D.T. et al. Structural characterization of the D290V mutation site in hnRNPA2 low-complexity-domain polymers. *Proc Natl Acad Sci U S A* **115**, E9782-E9791 (2018).
31. Geuens, T., Bouhy, D. & Timmerman, V. The hnRNP family: insights into their role in health and disease. *Hum Genet* **135**, 851-67 (2016).
32. Zheng, S.Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nature Methods* **14**, 331-332 (2017).
33. Rohou, A. & Grigorieff, N. CTFIND4: Fast and accurate defocus estimation from electron micrographs. *Journal of Structural Biology* **192**, 216-221 (2015).
34. Emsley, P., Lohkamp, B., Scott, W.G. & Cowtan, K. Features and development of Coot. *Acta Crystallographica Section D* **66**, 486-501 (2010).
35. Adams, P.D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D* **66**, 213-221 (2010).
36. Winn, M.D. et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **67**, 235-42 (2011).

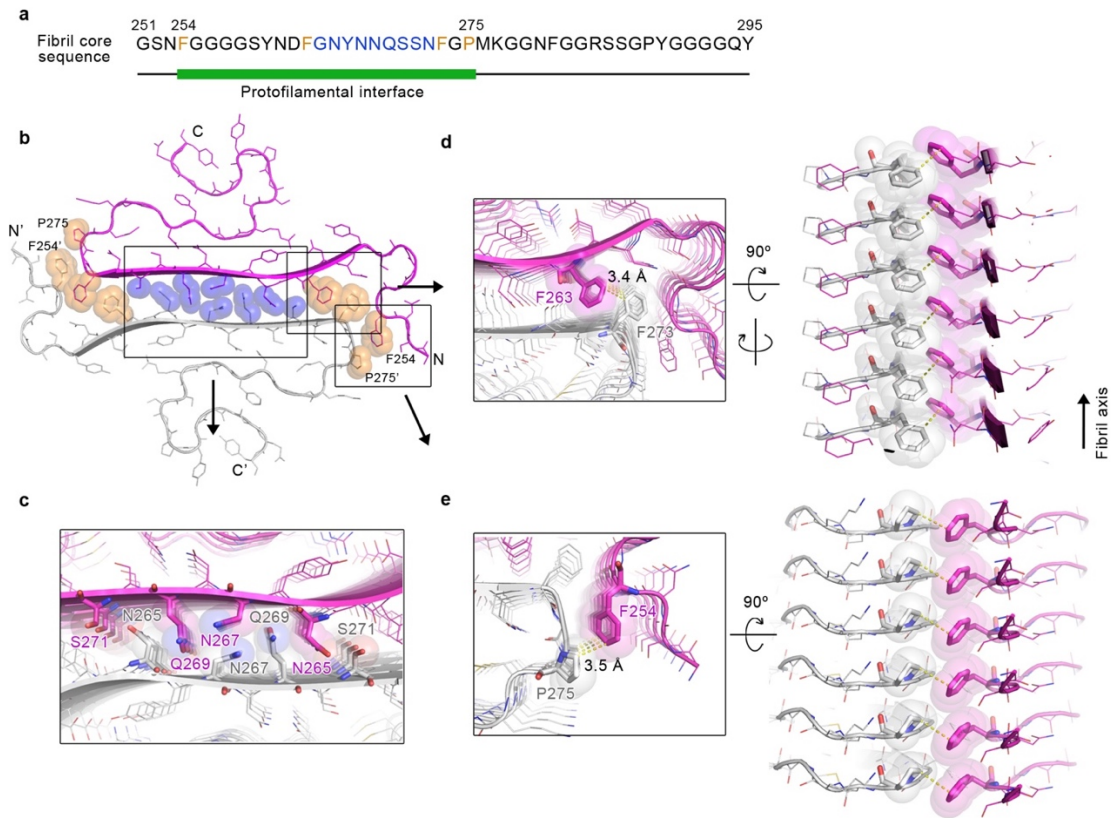
## Figures and tables



**Figure 1 CryoEM structure determination of the amyloid fibril formed by hnRNPA1 LC domain.** **a**, CryoEM image of the amyloid fibrils formed by hnRNPA1 LC. **b**, CryoEM 3D reconstruction of the hnRNPA1 LC fibril. **c**, CryoEM density map of the hnRNPA1 LC fibril. Fibril parameters including fibril width, half pitch, twist angle and rise are marked. Protofilaments are colored in magenta and orange individually. **d**, Cross-section view of the density map with structure model built upon it. Protofilaments are in different colors.

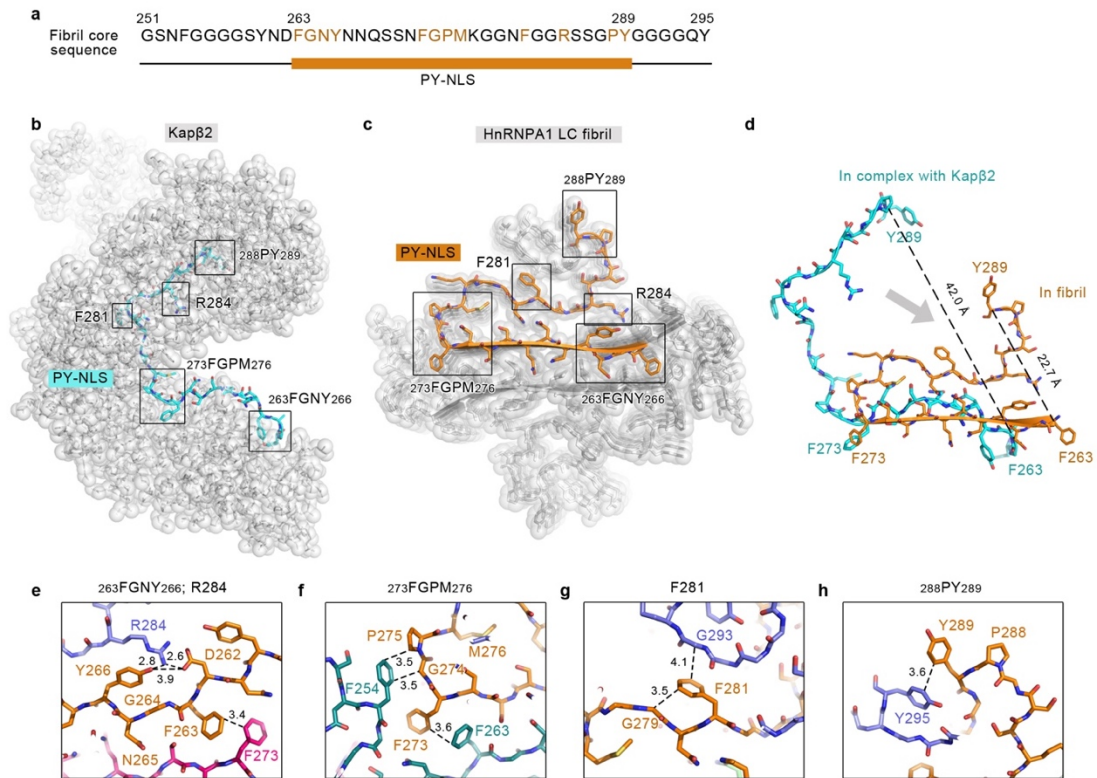


**Figure 2 Overall structure of the hnRNPA1 LC fibril.** **a**, The primary sequence of the hnRNPA1 FC. Residues forming the  $\beta$  strand are highlighted in red. Glycine residues are in blue. Secondary structures are indicated below. **b**, Top view of the cryoEM structure of hnRNPA1 LC fibril. Six layers of backbones are shown. The structure is colored by chain. The pseudo-2-fold screw axis is indicated. The sheet kink region is zoomed in with side chains shown. **c**, Side views of the fibril structure in (**b**).

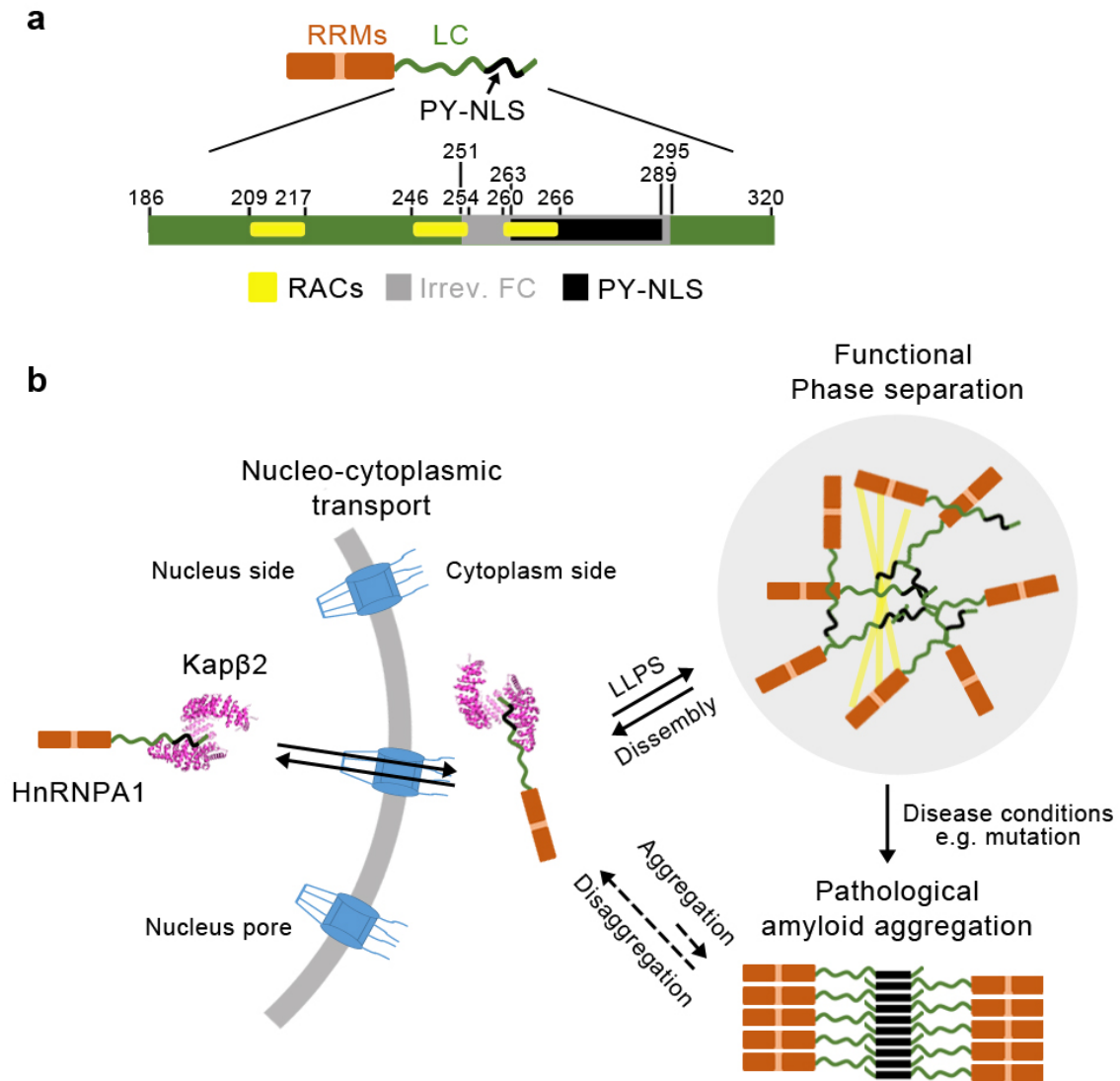


**Figure 3 Protofilamental interface of the hnRNPA1 LC fibril.** **a**, The primary sequence of the hnRNPA1 FC. The green bar marks the region of protofilamental interface. Residues forming the steric zipper in the interface are highlighted in blue. Residues forming  $\pi$ -stacking and hydrophobic interactions flanking the steric zipper are highlighted in orange. **b**, Top view of the hnRNPA1 LC fibril structure. Side chains that are involved in the inter-protofilamental interactions are shown in spheres. Among them, the polar side chains in the steric zipper are colored in blue; the Phe and Pro side chains are colored in orange. The two protofilaments are colored in magenta and gray, respectively. The framed regions are zoomed in and shown in top and side views in (**c-e**).





**Figure 4 Structures of PY-NLS in complex with Kapβ2 and in the hnRNPA1 LC fibril.** **a**, The primary sequence of the hnRNPA1 FC. The orange bar marks the region of PY-NLS. Residues interacting with Kapβ2 are highlighted in orange. **b**, Complex structure of PY-NLS and Kapβ2 (PDB ID: 2H4M). PY-NLS is shown in sticks and colored in cyan. Kapβ2 is shown in gray spheres. Regions of PY-NLS interacting with Kapβ2 are framed and labeled. **c**, Top view of hnRNPA1 LC fibril structure. 6 layers are shown. PY-NLS in one subunit is highlighted in orange. The rest of the structure is shown in gray spheres. The same framed regions of PY-NLS in (**b**) are framed and labeled. **d**, Overlay of PY-NLS structures in complex with Kapβ2 and in fibril. **e-h**, Zoom-in views of the framed regions in (**c**). HnRNPA1 LC subunits are colored differently. Distances between side chains are indicated in Å.



**Figure 5 Hypothesis of hnRNPA1 forming different structures to adapt to different function.** **a**, Domain organization of hnRNPA1. The LC domain is colored in green and zoomed in. Segment 251-295 that is involved in the the irreversible fibril core is colored in gray. The PY-NLS (segment 263-289) is highlighted in black. The reversible amyloid-forming core segments (RACs) are highlighted in yellow. **b**, Schematic illustration of the different structural states of hnRNPA1. HnRNPA1 monomer forms complex with Kap $\beta$ 2 via PY-NLS. The LC domain mediates self-association of hnRNPA1 to undergo liquid-liquid phase separation (LLPS) via multivalent and reversible amyloid-like interactions. Under disease conditions, hnRNPA1 further assembles into highly-ordered amyloid fibrils via PY-NLS.

**Table 1 Statistics of cryoEM data collection and refinement.**

Name	<b>hnRNPA1 LC fibril</b>
PDB ID	7BX7
EMDB ID	EMD-30235
<b>Data Collection</b>	
Magnification	130,000
Pixel size (Å)	0.52
Defocus Range (µm)	-1.3 to -2.1
Voltage (kV)	300
Camera	K2 summit
Microscope	Titan Krios
Exposure time (s/frame)	0.25
Number of frames	32
Total dose (e <sup>-</sup> /Å <sup>2</sup> )	59
<b>Reconstruction</b>	
Micrographs	3,227
Manually picked fibrils	62,656
Box size (pixel)	288
Inter-box distance (Å)	30.0
Segments extracted	801,707
Segments after Class2D	259,615
Segments after Class3D	69,037
Resolution (Å)	2.80
Map sharpening B-factor (Å <sup>2</sup> )	-109.473
Helical rise (Å)	2.37
Helical twist (°)	179.05
<b>Atomic model</b>	
Non-hydrogen atoms	1,896
Protein residues	270
Ligands	0
r.m.s.d. Bond lengths	0.009
r.m.s.d. Bond angles	0.833
All-atom clash score	15.34
Rotamer outliers	0.00%
Ramachandran Outliers	0.00%
Ramachandran Allowed	9.30%
Ramachandran Favored	90.70%