nature chemistry

Check for updates

OPEN
Molecular interactions of FG nucleoporin repeats at high resolution

Alain Ibáñez de Opakua [1], James A. Geraets [2], Benedikt Frieg[2], Christian Dienemann[3], Adriana Savastano[1], Marija Rankovic[4], Maria-Sol Cima-Omori [1], Gunnar F. Schröder [2,5] ✉ and Markus Zweckstetter [1,4] ✉

Proteins that contain repeat phenylalanine-glycine (FG) residues phase separate into oncogenic transcription factor condensates in malignant leukaemias, form the permeability barrier of the nuclear pore complex and mislocalize in neurodegenerative diseases. Insights into the molecular interactions of FG-repeat nucleoporins have, however, remained largely elusive. Using a combination of NMR spectroscopy and cryoelectron microscopy, we have identified uniformly spaced segments of transient β-structure and a stable preformed α-helix recognized by messenger RNA export factors in the FG-repeat domain of human nucleoporin 98 (Nup98). In addition, we have determined at high resolution the molecular organization of reversible FG–FG interactions in amyloid fibrils formed by a highly aggregation-prone segment in Nup98. We have further demonstrated that amyloid-like aggregates of the FG-repeat domain of Nup98 have low stability and are reversible. Our results provide critical insights into the molecular interactions underlying the self-association and phase separation of FG-repeat nucleoporins in physiological and pathological cell activities.

Phenylalanine-glycine (FG) repeats are present in many intrinsically disordered proteins and have been linked to multiple cellular processes[1–4]. Sequence analysis has identified more than 600 proteins containing FG repeats[5]. Nucleoporins containing FG repeats form the permeability barrier of the nuclear pore complex[4,6–8]. In addition, they are involved in cancer-associated biomolecular condensates, the so-called oncogenic transcription factor condensates[9–12]. FG-repeat proteins are also present in several other membraneless organelles[3]. Increasing evidence further links cellular mislocalization of FG-repeat-containing nucleoporins to pathological protein misfolding and aggregation in neurodegenerative diseases, including Alzheimer's disease, amyotrophic lateral sclerosis and frontotemporal dementia[13,14]. Mutagenesis in combination with functional assays has provided ample support for the critical role of FG repeats in these cellular processes. However, the nature of the underlying molecular interactions between FG repeats is largely unknown.

An important FG-repeat protein associated with multiple physiological and pathological processes is the human nucleoporin 98 (Nup98)[2]. In certain types of leukaemia, the FG-repeat domain of Nup98 is fused to a chromatin-binding domain as a result of recurrent chromosomal translocations[15]. The oncogenic properties of the Nup98 fusion proteins are related to their ability to concentrate into condensates[9–12]. Site-directed mutagenesis demonstrates that the ability to self-associate and form oncogenic transcription factor condensates critically depends on the FG repeats of Nup98 (refs. [9,11,12]). Consistent with the formation of Nup98 condensates in cells, the FG repeats of several nucleoporins phase separate in vitro into liquid-like droplets and solid-like condensates above critical concentrations as low as 20 nM (refs. [5,16]). The biophysical properties of particles and gels of FG-repeat Nups have been char-

acterized previously in great detail[5,17–20]. The FG-repeat domain of Nup98 also facilitates the aggregation of the protein tau associated with Alzheimer's disease in vitro and accumulates in the cell bodies of neurons that contain tau aggregates[14].
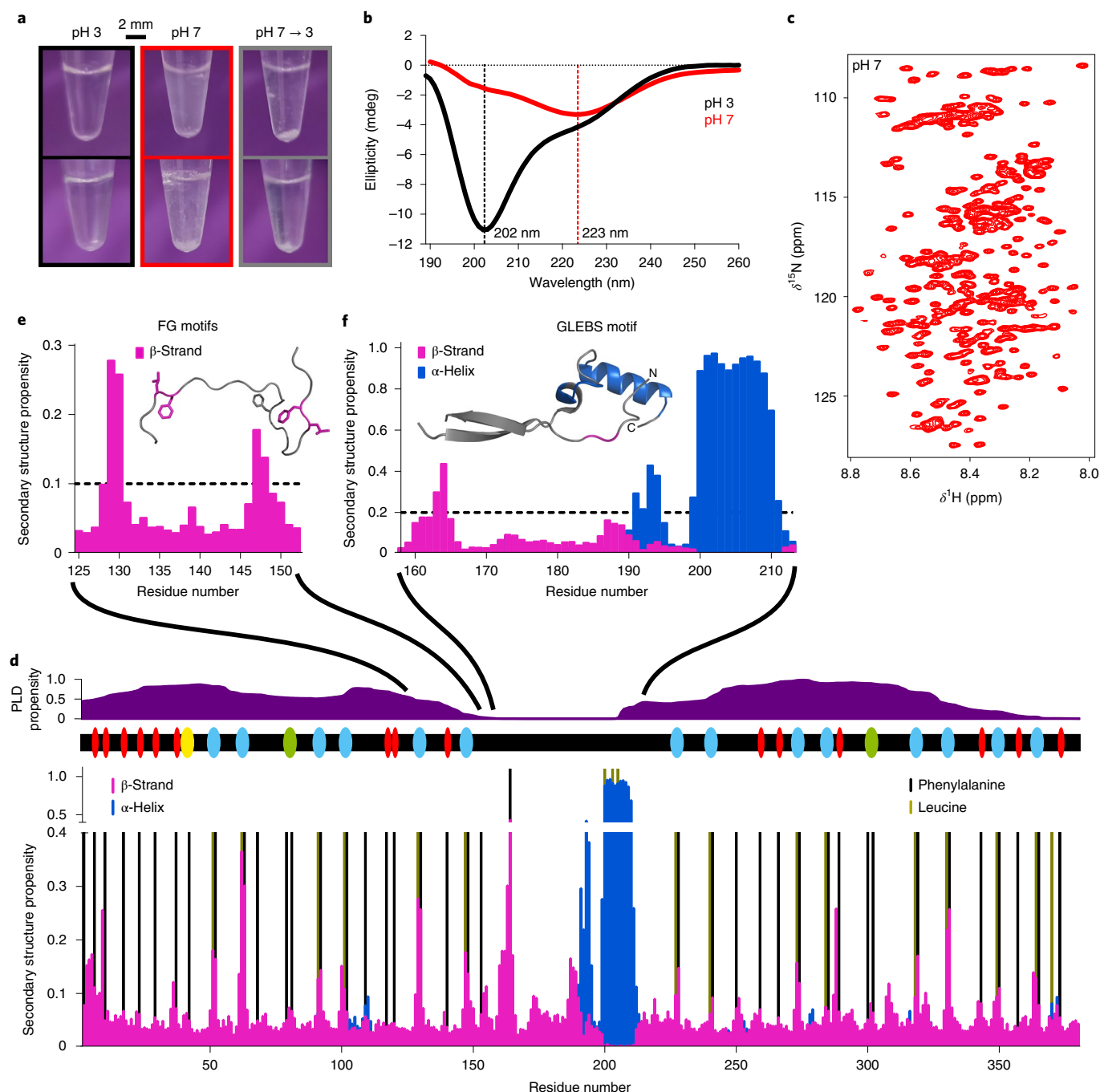
Through a combination of NMR spectroscopy and cryoelectron microscopy (cryo-EM), here we provide insights into the dynamic structure of the FG-repeat domain of Nup98 at the single residue level, reveal a stable preformed structure and define the molecular organization of cohesive FG–FG interactions in reversible FG clusters at high resolution.

Results
Molecular organization of the FG-repeat domain of human Nup98. The amino-terminal 384 residues of human Nup98 (named Nup98[FG]) have a high density of FG repeats comprising in total 41 phenylalanine residues. We recombinantly expressed and purified Nup98[FG] (see Methods and Supplementary Materials). Nup98[FG] is soluble and predominantly disordered at pH 3 (Fig. 1a,b and Supplementary Fig. 1). This is in agreement with previous reports that at acidic pH the net charge of proteins changes and the hydrophobic interactions of aromatic rings are attenuated, promoting the solubility of proteins that are highly insoluble at native pH (refs. [21,22]).

To gain insight into the soluble structure of Nup98[FG], we performed extensive NMR measurements on both Nup98[FG] and a large number of shorter Nup98[FG] segments (Supplementary Figs. 2 and 3). In addition, NMR spectra were recorded at both native pH and pH 3 (Fig. 1c and Supplementary Fig. 2). Through this integrative approach, we overcame the challenges of a highly repetitive sequence and strong aggregation tendency[5] and determined the sequence-specific resonance assignment of Nup98[FG] at native pH

Fig. 1 | Dynamic structure of Nup98[FG]. a, Macroscopic changes in samples of the FG-repeat domain of Nup98[FG] at pH 3, after adjusting to pH 7 and then back to pH 3 before incubation (top row) and after incubation at 65 °C for 30 min (bottom row). **b**, CD spectra of Nup98[FG] in the soluble phase (pH 3) and in the condensed/aggregated phase (pH 7). **c**, Two-dimensional ¹H–¹⁵N heteronuclear single quantum coherence spectrum of Nup98[FG] at pH 7 (started ~5 min after adjusting the pH from 3 to 7). **d**, The conformational properties of soluble Nup98[FG] at pH 6.8. The likelihood of residue-specific backbone torsion was determined from the experimental NMR chemical shifts using TALOS-N. The propensity for prion-like domain (PLD) structure and the location of FG motifs are shown above (red, FG; yellow, SAFG; cyan, GLFG; green, FXFG). **e**, The β-strand motifs in the N-terminal prion-like domain (taken from **d**). The conformation derived from TALOS-N is shown as the inset; the colouring is based on the threshold propensity shown in the graph. Phenylalanine and leucine side chains are displayed. **f**, Preformed secondary structure of the GLEBS-binding motif of monomeric unbound Nup98[FG] (taken from **d**). The inset shows the crystal structure (Protein Data Bank identification (PDB ID): 3MMY) of the GLEBS-binding motif in the complex with the mRNA export factor Rae1. Regions of the crystal structure that are preformed (>0.2) prior to binding to Rae1 are coloured (the α-helical structure is shown in blue and the β-structure in magenta).

(Supplementary Table 1). The sequence-specific assignment showed that uniformly spaced, short segments of extended β-structure (Fig. 1d,e), which coincide with phenylalanine residues, are abundant in the FG-repeat domain of Nup98.

Preformed structure for messenger RNA export factor binding. The nuclear membrane spatially separates transcription and translation. The mRNA synthesized in the nucleus has to cross the nuclear membrane by passing through the permeability barrier of
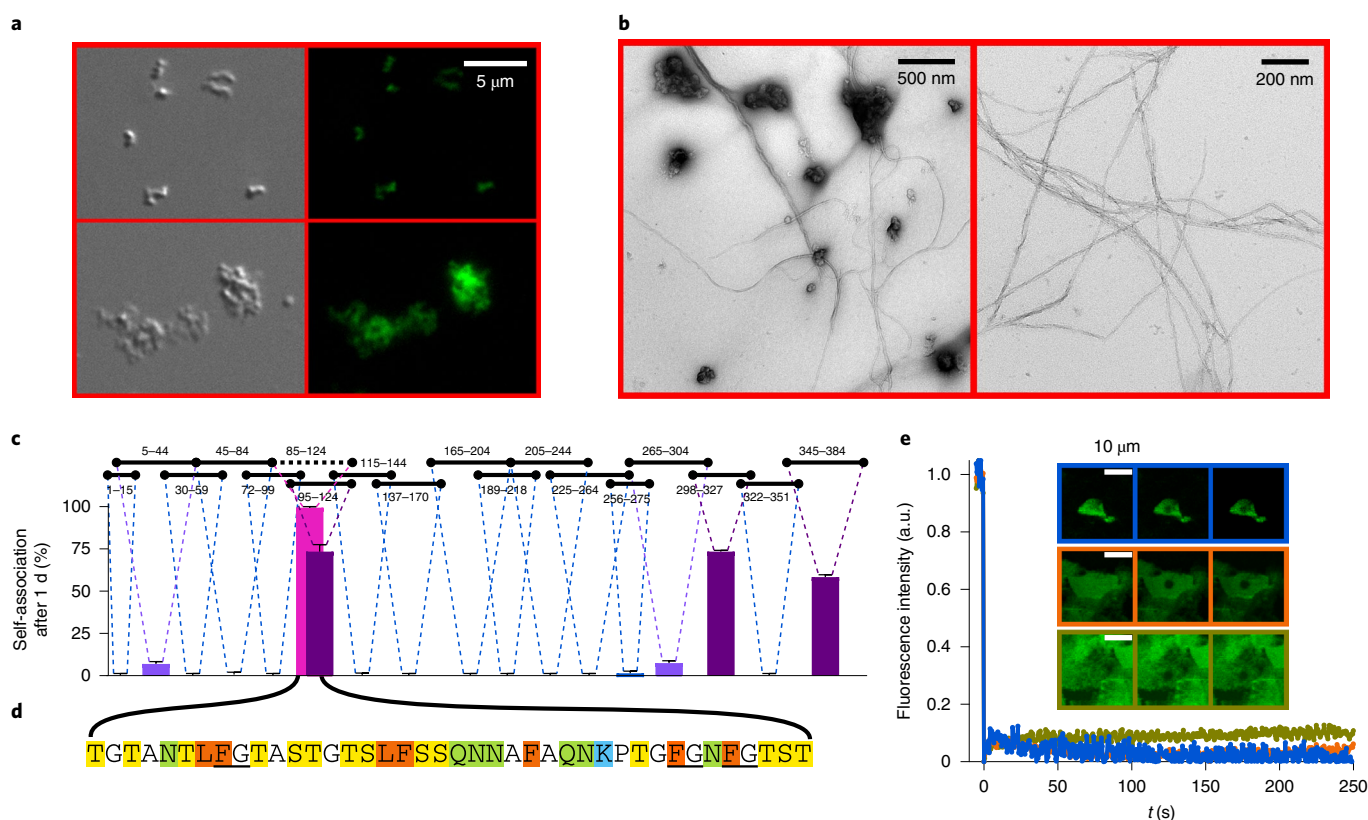
Fig. 2 | Nup98FG forms an amyloid-like structure. a, Differential interference contrast (left) and fluorescence (right) microscopy images of two examples of aggregates/condensates formed by Nup98FG. The fluorescence originates from staining with ThT. **b**, Negative-stain EM images of Nup98FG aggregates. **c**, Self-association propensities of 18 Nup98FG fragments. The residue numbers of each fragment in the Nup98FG sequence are indicated above the plot. The percentage propensity to self-association (condensation/aggregation) corresponds to the difference between the NMR signal expected for that peptide concentration and that observed after 1 day of incubation at 5 °C. High, intermediate and low self-association propensities are coloured magenta, purple and violet, respectively. For the fragments showing no propensity data, the NMR signal remained unchanged and the connecting dashed lines are shown in blue. The error bars represent the s.d. based on the NMR signal-to-noise ratio. **d**, Residues 85–124 of Nup98 (Nup98FG85), which are the most prone to self-association, as indicated in **c** (black horizontal dotted line). Phenylalanine and leucine are highlighted in orange, asparagine and glutamine in green, lysine in blue, and serine and threonine in yellow. The canonical FG motifs are underlined (notably, the glycine in FG motifs is sometimes replaced by other small amino acids, for example, serine[5]). **e**, Nup98FG85 forms solid-like particles that can be stained with ThT. The time plot shows that the bleached ThT fluorescence of three different particles did not recover within 250 s (the plot colour corresponds to the micrograph with the same colour outline).

the nuclear pore complex. This passage is achieved with the help of mRNA export factors. The mRNA export factor Rae1 binds to the conserved Gle2-binding sequence (GLEBS; residues 157–213)[23], which is located between the two prion-like FG-rich regions of Nup98 (Fig. 1d,f). When complexed with Rae1, the GLEBS domain of Nup98 folds into two short β-strands, a short α-helix and a longer α-helix comprising residues 200–210 (the structure is displayed in the inset of Fig. 1f)[23]. Residue-specific analysis of the NMR chemical shifts (Fig. 1d) revealed that residues 200–210 of Nup98FG, which contain the core GLEBS motif[24] and form the long α-helix in the complex with Rae1 (ref. [23]), are folded into a stable α-helix prior to binding to Rae1 (Fig. 1f). The prefolded α-helical structure of the GLEBS motif may decrease the entropic costs of binding to Rae1, thereby promoting the interaction of the GLEBS motif of Nup98 with Rae1.

Amyloid-like interactions of Nup98 FG repeats. To gain insight into the structure of cohesive FG–FG interactions, we studied the molecular properties of aggregated Nup98FG. When we changed the pH from 3 to 7, solutions of Nup98FG rapidly turned turbid (Fig. 1a) and the circular dichroism (CD) spectrum changed markedly (Fig. 1b). For a 'regular' β-structure, a minimum at 218 nm and a maximum at 200 nm are expected. The CD spectrum of Nup98FG

at pH 7, however, does not exhibit a maximum at 200 nm and the minimum is located at 223 nm. A minimum at ~223 nm was previously observed for fibrillar amyloid-β(1–40)[25], while the lack of the maximum at 200 nm is probably due to a substantial amount of remaining disordered structure in Nup98FG at pH 7.

Further support for the pH-induced changes in the molecular properties of Nup98FG was obtained by microscopy. At pH 7, clusters of particles were observed by differential interference contrast microscopy (Fig. 2a, left). Notably, the particles were fluorescent when we exposed the sample to the amyloid-specific dye thioflavin T (ThT) (Fig. 2a and Supplementary Fig. 4). In addition to these ThT-positive particles/particle clusters, amyloid fibrils were detected by electron microscopy (EM; Fig. 2b). Taken together, these experiments showed that the FG-repeat domain of Nup98, similar to the FG-repeat domains of other nucleoporins[26,27], readily forms ThT-positive particles and amyloid-like structures at native pH.

To investigate whether specific regions in Nup98FG favour the formation of amyloid-like aggregates, we studied the aggregation propensity of 18 shorter segments, which together cover the FG-repeat domain of Nup98 (Fig. 2c). We quantified the aggregation propensity of each segment by the time-dependent changes in NMR signal intensity. The analysis showed that the aggregation
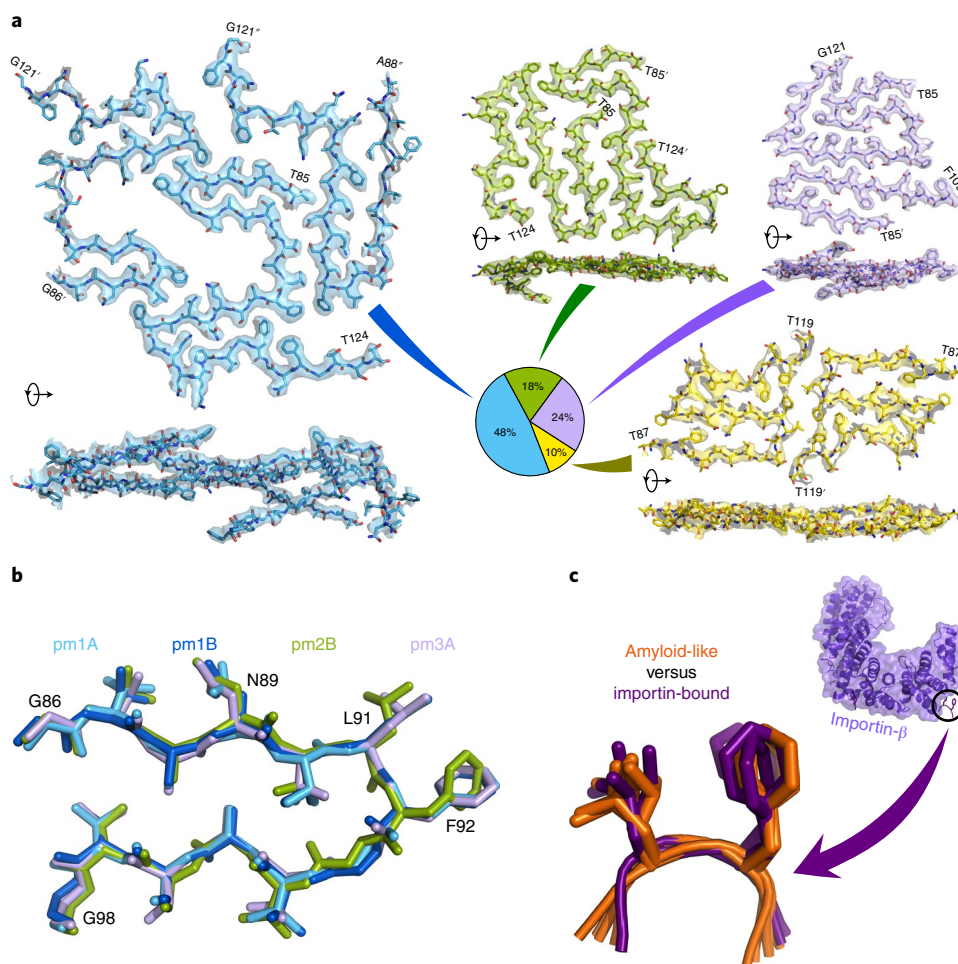
Fig. 3 | Cryo-EM structures of nucleoporin FG fibrils. a, Single cross-sections of the fibril structures of the Nup98^{FG85} polymorphs 1, 2, 3 and 4, shown in blue, green, violet and yellow, respectively, together with the cryo-EM density maps (from the top and side of the fibril axis). Zoned density maps are shown for clarity; unzoned maps are presented in Supplementary Figs. 6–9. Initial and final residues of each peptide are indicated, using ' and '' for the second and third chains, respectively. Polymorph populations, as a percentage of fibrillar segments, are displayed in the pie chart. The percentage of peptides forming each polymorph is 58, 15, 19 and 8% for pm1–pm4, respectively. **b**, The most common structural motif in Nup98^{FG85} fibrils: residues 86–98 of Nup98^{FG85} pm1 chain A (pm1A) and chain B (pm1B), pm2 chain B (pm2B) and pm3 chain A (pm3A) are aligned with a maximum root-mean-squared deviation of 0.492 Å. **c**, Superposition of the amyloid-like structures of six XLFX motifs, present in the Nup98^{FG85} pm1 fibrils (shown in **b**), on five GLFG motifs bound to importin-β (PDB ID: 1O6P).

propensity varies markedly along the Nup98^{FG} sequence (Fig. 2c). The most aggregation-prone segment comprises residues 85–124 (named Nup98^{FG85}; Fig. 2c). Nup98^{FG85} contains five phenylalanine residues, three in FG motifs (Fig. 2d). Two of the phenylalanines are preceded by leucine. Similar to the 384-residue Nup98^{FG} domain (Fig. 2a), Nup98^{FG85} forms ThT-positive particles/clusters (Fig. 2e and Supplementary Fig. 4). Fluorescence recovery after photobleaching (FRAP) experiments showed that bleached fluorescence did not recover with time (Fig. 2e), consistent with the solid-like nature of the Nup98^{FG85} particles/clusters. We have thus identified a highly aggregation-prone FG-rich segment in Nup98.

We also quantified the time-dependent decrease in the NMR signal of a slightly shorter peptide (residues 95–124 of Nup98; Fig. 2c) and compared this with the decrease in signal of two mutant peptides in which either two phenylalanine residues (F117 and F120) or two glutamine residues (Q105 and Q111) were replaced by serine (Supplementary Fig. 5). The glutamine-to-serine mutations reduced slightly the aggregation kinetics and resulted in a higher residual NMR-observable peptide concentration at the end of the incubation period. The impact of the phenylalanine-to-serine mutations was more pronounced, leading to a nearly twofold increase in the

time for 50% of the peptide to become unobservable. This analysis suggests that both phenylalanine and glutamine are important contributors to the aggregation of Nup98^{FG}.

Cryo-EM structures of Nup98(85–124) fibrils. To gain insight into the molecular organization of cohesive FG-repeat interactions, Nup98^{FG85} was allowed to aggregate in pure water. EM revealed the formation of well-defined and mostly separated amyloid fibrils (Supplementary Fig. 6). Using cryo-EM, we determined the structures of four different fibril polymorphs (pm1–pm4; Fig. 3a and Supplementary Tables 2 and 3).

The most abundant polymorph fibril, pm1, was resolved at a resolution of 2.8 Å (Supplementary Fig. 7a) and is formed by three asymmetrically arranged protofilaments (Supplementary Fig. 8). Polymorphs pm2 and pm3, which are almost equally populated, were solved at a resolution of 3.3 and 2.8 Å, respectively, and are formed by two protofilaments (Supplementary Figs. 7b,c, 9 and 10). The least populated polymorph, pm4, was reconstructed at a resolution of 3.4 Å (Supplementary Fig. 7d) and contains two protofilaments related by an approximate 2_1 screw symmetry (Supplementary Fig. 11).
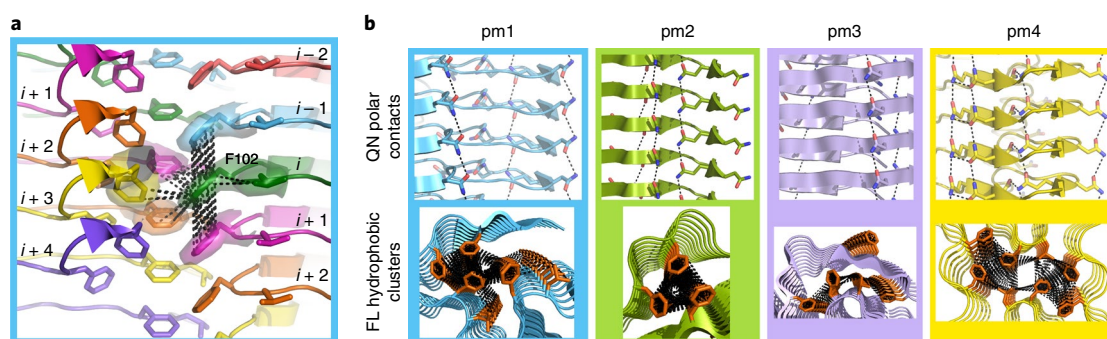
Fig. 4 | Cohesive FG–FG interactions at high resolution. a, Interactions of F102 with other phenylalanine residues from different chains and layers (shown in different colours) in pm1. The parameter *i* refers to the layer of the reference residue F102 (shown in green), *i* − 1 and *i* − 2 refer to the layers above the reference, and *i* + 1, *i* + 2, *i* + 3 and *i* + 4 to the layers below it. Associated cryo-EM densities are displayed. The dashed lines show F102 contacts below 5 Å. **b**, Selected views of glutamine-asparagine polar contacts (dashed lines, 4 Å cut-off; top) and major phenylalanine-rich clusters (5 Å cut-off; bottom) in pm1–4.
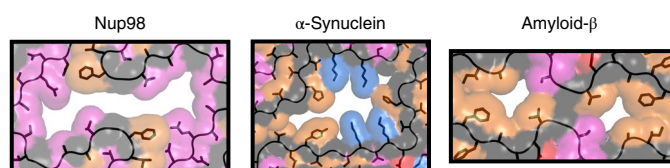
Fig. 5 | Polar residues line the cavity in nucleoporin FG fibrils. The cavity of Nup98[FG85] fibril pm1 is compared with the cavities in the disease-associated fibrils of α-synuclein (PDB ID: 6XYO) and amyloid-β (PDB ID: 6SHS). Hydrophobic/aromatic residues (LIVWFYHMA) are shown in orange, polar residues (STQN) in magenta and charged residues in blue (KR) and red (ED), respectively.

All four fibril types display a parallel in-register β-structure comprising short and kinked β-strands (Fig. 3a). Kinked aromatic-rich structures were previously named low-complexity aromatic-rich kinked segments[28]. The most common motif shared by the four polymorphs is a 13-residue β-turn/β-arch structure (Fig. 3b). A leucine-phenylalanine pair (L91-F92) is located at the tip of the β-turn, with both side chains pointing away from the turn (Fig. 3b). The other LF motifs in the four fibril structures show a very similar structure (Supplementary Fig. 12). Notably, this specific amyloid structure of the LF motif closely overlaps the structure of the GLFG motifs bound to importin-β (Fig. 3c).

Cohesive FG–FG interactions at high resolution. The structure of the most abundant polymorph, pm1, deviates considerably from the flat two-dimensional (2D) arrangement of cross-β-structure sheets (Fig. 4a). This leads to one chain forming contacts with several layers of the opposing protofilaments and the formation of a cluster of phenylalanine side chains (Fig. 4a,b, bottom). For example, the aromatic ring of F102 contacts three other phenylalanine rings located up to three layers away (Fig. 4a). Similar interactions can be found in the other polymorphs, where clusters of up to six phenylalanine rings, in some cases including leucine side chains, establish a tight network of molecular contacts (Fig. 4b, bottom). In all four Nup98[FG85] fibril structures, side chains stacked along the fibril axes build glutamine and asparagine ladders (Fig. 4b, top). Previous studies showed that gels formed by Nup98 FG domains, which are rich in glutamine and asparagine, exhibit a strong amyloid-like character[5,27].

The pm1 structure contains a large cavity (Fig. 3a) that is predominantly lined by polar residues (Fig. 5). Because the Nup98[FG85] fibrils were prepared in pure water, it is likely that the cavity is filled

with water molecules. In contrast, the cavity observed in the structure of α-synuclein fibrils purified from the brain of a patient with multiple system atrophy is lined by hydrophobic and charged residues (Fig. 5). Finally, mostly hydrophobic residues are observed at the rim of the cavity in amyloid-β fibrils purified from the brain of a patient with Alzheimer's disease (Fig. 5). Hydrophobic and/or charged cofactors are therefore likely to fill the cavities of these disease-associated amyloid fibrils[29].

FG–FG interactions have low stability. Next, we estimated the atomic solvation energies[30] as a measure of the stability of the four Nup98[FG85] fibril structures and compared them with those of previously resolved fibril structures (Fig. 6a). We obtained values of approximately −20 kcal mol[−1] per chain for the four Nup98[FG85] fibril structures. This value is comparable to the stability of amyloid fibrils formed by the RNA-binding proteins hnRNPA1 and hnRNPA2 (Fig. 6a). By contrast, most disease-associated fibrils are predicted to be more stable (Fig. 6a). Considering that Nup98[FG85] has the strongest condensation propensity in the FG-repeat domain of Nup98 (Fig. 2c), −20 kcal mol[−1] per chain is likely to be the maximum value that is reached for cohesive conformations within the FG-repeat domain of Nup98.

Mapping of the calculated solvation energies onto the structures shows that the stabilizing interactions are distributed non-uniformly (Fig 6b and Supplementary Fig. 13). Most of the stabilizing interactions arise where there are clusters of two to four phenylalanine rings and one or more leucine side chains. Other regions contribute little to the stability of the FG nucleoporin fibrils (Fig. 6b and Supplementary Fig. 13).

The Nup98[FG85] fibril structures are predicted to have higher stability than the amyloid fibrils formed by the low-complexity region of the stress granule-associated protein fused in sarcoma (FUS; Fig. 6a). The low-complexity region of FUS (FUS-LC) contains many polar residues (Supplementary Fig. 14). In addition, the predominant aromatic residue in FUS-LC is tyrosine, which has the ability to form polar contacts through its hydroxy group and does not form as many hydrophobic clusters as the phenylalanine residues in Nup98[FG85] (compare Supplementary Fig. 14 with Fig. 6b). The high content of polar residues might be responsible for the predicted lower stability of FUS-LC fibrils (Fig. 6a).

FG-based interactions are reversible. Pathogenic amyloid fibrils do not resolubilize when exposed to very high temperature (>100 °C) over several hours[31–33]. In contrast, functional fibrils/fibres associated with gels and membraneless organelles are less stable (Fig. 6a) and can dissociate upon changes in pH/ionic strength and elevated
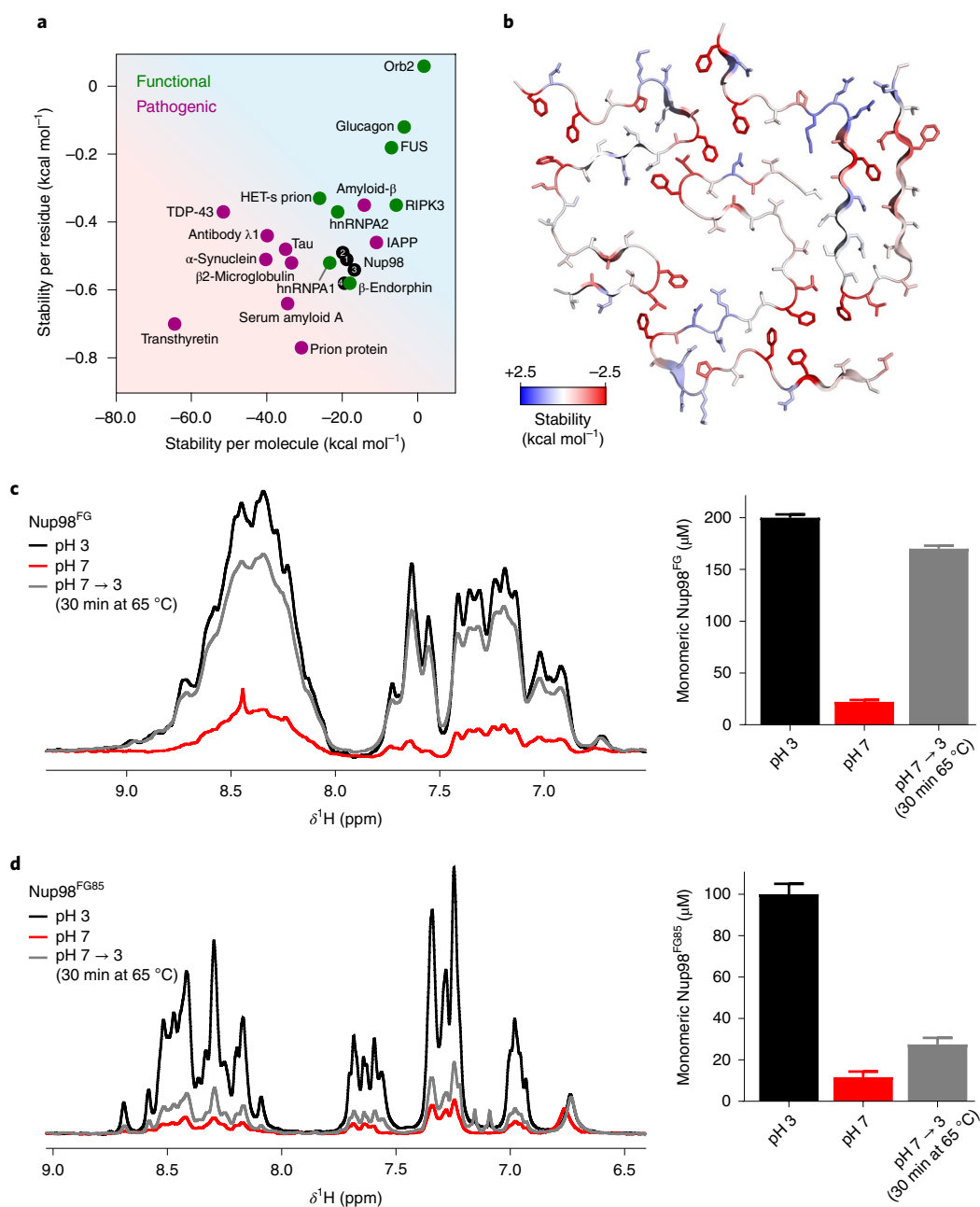
Fig. 6 | Cohesive FG clusters have low stability. a, Stability of Nup98[FG85] fibrils (black) per residue and per molecule compared with the stability of previously determined functional and pathogenic fibril structures. The energies were calculated on the basis of cryo-EM structures. **b**, Residue-specific stabilities based on solvation energies mapped onto the structure of Nup98[FG85] fibril pm1. Red and blue represent high and low stability, respectively. **c**, One-dimensional ^{1}H NMR spectra of Nup98[FG] at pH 3, after adjusting to pH 7 and after adjusting back to pH 3 and incubating the sample for 30 min at 65 °C (left). The corresponding concentrations of monomeric Nup98[FG] calculated from the integral of the NMR signals (right). **d**, One-dimensional ^{1}H NMR spectra of Nup98[FG85] at pH 3, after adjusting to pH 7 and after adjusting back to pH 3 and incubating the sample for 30 min at 65 °C (left). The corresponding concentrations of monomeric Nup98[FG85] calculated from the integral of the peaks (right). The error bars in **c** and **d** represent the s.d. based on NMR signal-to-noise ratios.

temperature (~65 °C)[31–33]. To experimentally study the stability of Nup98[FG] aggregates, we freshly prepared the 384-residue FG-repeat domain of Nup98 at pH 3. In agreement with the CD and turbidity experiments (Fig. 1a,b), Nup98[FG] at pH 3 displays an NMR spectrum that is typical of a flexible protein (Fig. 6c). When we then raised the pH to 7, we observed an immediate and strong signal loss. The signal loss was partially reversed on lowering the pH back to 3 and raising the temperature to 65 °C (Fig. 6c, left). In addition, the sample became less turbid (Fig. 1a). Quantification of the NMR

signal intensities showed that ~85% of the original protein signal was regained (Fig. 6c, right).

We then performed the same experiments with the highly aggregation-prone Nup98[FG85] segment (Fig. 6d). Again, strong signal loss occurred at native pH, but the NMR signal could be regained on returning to pH 3 and raising the temperature to 65 °C (Fig. 6d, left). Notably, however, only 22% of the original NMR signal was regained (Fig. 6d, right), indicating that the more aggregation-prone Nup98[FG85] segment forms more stable aggregates. This combination

of experiments (Fig. 6c,d) and stability estimations (Fig. 6a,b) demonstrates that amyloid-like aggregates of the FG-repeat domain of Nup98 have low stability.

Discussion

The human nucleoporin Nup98 forms fusion proteins in certain types of malignant leukaemia, is a critical component of the permeability barrier of the nuclear pore complex and is mislocalized in cytosolic deposits in Alzheimer's disease[2,9–12,14,15]. Key to these Nup98 activities is its large FG-repeat domain with a high number of FG motifs. In leukaemia-associated transcription factor condensates, the Nup98 FG-repeat domain is fused to DNA-binding domains, whereas it is attached to scaffold proteins inside the nuclear pore complex. In both cases, the FG-repeat domain of Nup98 critically influences the molecular properties of these cellular compartments. High-resolution information about the interactions between FG motifs either within a single Nup98[FG] chain or between multiple Nup98[FG] chains has, however, been largely elusive. In this study we combined NMR spectroscopy with four cryo-EM structures to gain detailed insights into the structural organization of both the more transient liquid-like and the more stable cohesive interactions of the FG-repeat domain of Nup98. The data reveal a preformed structure recognized by mRNA export factors and the molecular basis of FG cluster stabilization, as well as establishing a structural mimicry of FG motifs inside cohesive FG clusters and when bound to nuclear transport receptors.

The condensates formed by the FG-repeat domains of Nup98 from different species display different amounts of amyloid-like structure, that is, different levels of liquid- and solid-like properties[5]. Our data show that the ability of Nup98 to self-associate varies strongly along its long FG-repeat domain. The strongest self-association propensity is observed in 3 segments comprising ~30–40 residues (Fig. 2c). This suggests that Nup98 phase separation and condensate formation may be explained by a well-balanced network of interactions involving less cohesive and transient interactions as well as cohesive and stable interactions. Depending on the precise sequence composition, the relative contribution of liquid- and solid-like interactions is thus likely to vary in Nup98 FG-repeat domains from different species. The combination of NMR spectroscopy and cryo-EM provides insights into the least and most stable interactions from the spectrum of possible FG-repeat interactions. When more liquid-like interactions dominate, the FG-repeat domain is very dynamic, the FG motifs transiently populate β-structure and the GLEBS motif folds into a stable α-helix. Cryo-EM determines with high resolution the most stable interactions from the spectrum of possible FG-repeat interactions. The combined data thus advance our structural knowledge of both the local and long-range structure of cohesive FG interactions to a new level compared with previous studies (for example, ref. [34]). Because the amyloid-like interactions between FG motifs are reversible (Fig. 6), it is likely that liquid-like transient and amyloid-like cohesive interactions are not completely distinct/disconnected molecular properties but can interchange even within a single region of an FG-repeat protein.

The central channel of the nuclear pore complex is filled with FG-repeat-containing nucleoporins that form the permeability barrier[6,7]. Several different (sometimes competing) models have been suggested for the molecular organization of the FG permeability barrier of the nuclear pore complex (reviewed, for example, in refs. [7,16,35,36]). In some of these models, cohesive interactions between FG motifs do not play a role in the transport selectivity of the FG-filled channel, while in other models cohesive FG–FG interactions are critical. In particular, in a model that considers the FG-filled channel as a phase-separated biomolecular condensate, cohesive FG–FG interactions are essential[17,19]. In addition, it has been shown that self-assembled Nup98[FG] particles, which display key features of transport selectivity, can contain cross-β-structure, as evidenced

by ThT staining[5]. Indeed, we confirmed here the presence of cross-β-structure in the Nup98[FG] particles by ThT staining (Fig. 1 and Supplementary Fig. 4). Solid-state NMR analysis of nucleoporin FG gels further indicated that regions containing asparagine and glutamine residues form the cross-β-structure[27]. Consistent with these results, we found that the most aggregation-prone segment of the FG-repeat domain of Nup98, identified in this study and structurally characterized at high resolution using cryo-EM, contains five asparagine and two glutamine residues.

Self-association and cohesive FG–FG interactions underlie the phase separation and formation of leukaemia-associated Nup98 fusion protein condensates[9,11,12] and might influence the mislocalization and accumulation of Nup98 in Alzheimer's disease[14]. However, high-resolution structural information has not been available for any of these FG-repeat assemblies. Furthermore, it has often been unclear whether they are more liquid- or solid-like, or whether they can be described at all by the physicochemical process of phase separation[7,9,14,16,35,36]. A recent study, for example, suggested that an artificial nucleoporin sequence based on Nup98 and comprising 52 repeats of 12 residues remains disordered in the phase-separated state[34], while another study based on a solid-state NMR analysis identified an amyloid-like β-structure inside gels formed by the yeast nucleoporin Nsp1p (ref. [27]). Importantly, changes in the material properties of condensates, from liquid- to solid-like, have been linked to human diseases[37], suggesting that cohesive FG-repeat interactions could be most relevant for the pathological states of FG-repeat-based compartments and the mislocalization of FG-repeat-containing nucleoporins in neurodegenerative diseases.

In summary, our study represents a critical advance in the understanding of the molecular interactions determining the self-association of Nup98 and thus provides an important reference for future studies of FG-repeat proteins.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41557-022-01035-7.

References

1. Gorlich, D. & Kutay, U. Transport between the cell nucleus and the cytoplasm. *Annu. Rev. Cell Dev. Biol.* **15**, 607–660 (1999).
2. Iwamoto, M., Asakawa, H., Hiraoka, Y. & Haraguchi, T. Nucleoporin Nup98: a gatekeeper in the eukaryotic kingdoms. *Genes Cells* **15**, 661–669 (2010).
3. Shinkai, Y., Kuramochi, M. & Miyafusa, T. New family members of FG repeat proteins and their unexplored roles during phase separation. *Front. Cell Dev. Biol.* **9**, 708702 (2021).
4. Terry, L. J. & Wente, S. R. Flexible gates: dynamic topologies and functions for FG nucleoporins in nucleocytoplasmic transport. *Eukaryot. Cell* **8**, 1814–1827 (2009).
5. Schmidt, H. B. & Gorlich, D. Nup98 FG domains from diverse species spontaneously phase-separate into particles with nuclear pore-like permselectivity. *Elife* **4**, e04251 (2015).
6. Heinß, N., Sushkin, M., Yu, M. & Lemke, E. A. Multifunctionality of F-rich nucleoporins. *Biochem. Soc. Trans.* **48**, 2603–2614 (2020).
7. Schmidt, H. B. & Gorlich, D. Transport selectivity of nuclear pores, phase separation, and membraneless organelles. *Trends Biochem. Sci.* **41**, 46–61 (2016).
8. Beck, M. & Hurt, E. The nuclear pore complex: understanding its function through structural insight. *Nat. Rev. Mol. Cell Biol.* **18**, 73–89 (2017).
9. Ahn, J. H. et al. Phase separation drives aberrant chromatin looping and cancer development. *Nature* **595**, 591–595 (2021).
10. Mittag, T. & Ansari, A. Z. Fusion proteins form onco-condensates. *Nat. Struct. Mol. Biol.* **28**, 543–545 (2021).

11. Terlecki-Zaniewicz, S. et al. Biomolecular condensation of NUP98 fusion proteins drives leukemogenic gene expression. *Nat. Struct. Mol. Biol.* **28**, 190–201 (2021).

12. Chandra, B. et al. Phase separation mediates NUP98 fusion oncoprotein leukemic transformation. *Cancer Discov.* **12**, 1152–1169 (2022).

13. Chou, C. C. et al. TDP-43 pathology disrupts nuclear pore complexes and nucleocytoplasmic transport in ALS/FTD. *Nat. Neurosci.* **21**, 228–239 (2018).

14. Eftekharzadeh, B. et al. Tau protein disrupts nucleocytoplasmic transport in Alzheimer's disease. *Neuron* **99**, 925–940 (2018).

15. Gough, S. M., Slape, C. I. & Aplan, P. D. NUP98 gene fusions and hematopoietic malignancies: common themes and new biologic insights. *Blood* **118**, 6247–6257 (2011).

16. Celetti, G. et al. The liquid state of FG-nucleoporins mimics permeability barrier properties of nuclear pore complexes. *J. Cell Biol.* **219**, e201907157 (2020).

17. Frey, S. & Gorlich, D. A saturated FG-repeat hydrogel can reproduce the permeability properties of nuclear pore complexes. *Cell* **130**, 512–523 (2007).

18. Frey, S. et al. Surface properties determining passage rates of proteins through nuclear pores. *Cell* **174**, 202–217 (2018).

19. Frey, S., Richter, R. P. & Gorlich, D. FG-rich repeats of nuclear pore proteins form a three-dimensional meshwork with hydrogel-like properties. *Science* **314**, 815–817 (2006).

20. Labokha, A. A. et al. Systematic analysis of barrier-forming FG hydrogels from *Xenopus* nuclear pore complexes. *EMBO J.* **32**, 204–218 (2013).

21. Van Lindt, J. et al. A generic approach to study the kinetics of liquid-liquid phase separation under near-native conditions. *Commun. Biol.* **4**, 77 (2021).

22. Xia, F., Nagrath, D. & Cramer, S. M. Effect of pH changes on water release values in hydrophobic interaction chromatographic systems. *J. Chromatogr. A* **1079**, 229–235 (2005).

23. Ren, Y., Seo, H. S., Blobel, G. & Hoelz, A. Structural and functional analysis of the interaction between the nucleoporin Nup98 and the mRNA export factor Rae1. *Proc. Natl Acad. Sci. USA* **107**, 10406–10411 (2010).

24. Wang, X. et al. The mitotic checkpoint protein hBUB3 and the mRNA export factor hRAE1 interact with GLE2p-binding sequence (GLEBS)-containing proteins. *J. Biol. Chem.* **276**, 26559–26567 (2001).

25. Terol, P. A., Kumita, J. R., Hook, S. C., Dobson, C. M. & Esbjorner, E. K. Solvent exposure of Tyr10 as a probe of structural differences between monomeric and aggregated forms of the amyloid-β peptide. *Biochem. Biophys. Res. Commun.* **468**, 696–701 (2015).

26. Milles, S. et al. Facilitated aggregation of FG nucleoporins under molecular crowding conditions. *EMBO Rep.* **14**, 178–183 (2013).

27. Ader, C. et al. Amyloid-like interactions within nucleoporin FG hydrogels. *Proc. Natl Acad. Sci. USA* **107**, 6281–6285 (2010).

28. Hughes, M. P. et al. Atomic structures of low-complexity protein segments reveal kinked β sheets that assemble networks. *Science* **359**, 698–701 (2018).

29. Li, D. & Liu, C. Hierarchical chemical determination of amyloid polymorphs in neurodegenerative disease. *Nat. Chem. Biol.* **17**, 237–245 (2021).

30. Eisenberg, D. & McLachlan, A. D. Solvation energy in protein folding and binding. *Nature* **319**, 199–203 (1986).

31. Kato, M. et al. Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. *Cell* **149**, 753–767 (2012).

32. Sawaya, M. R., Hughes, M. P., Rodriguez, J. A., Riek, R. & Eisenberg, D. S. The expanding amyloid family: structure, stability, function, and pathogenesis. *Cell* **184**, 4857–4873 (2021).

33. Ragonis-Bachar, P. & Landau, M. Functional and pathological amyloid structures in the eyes of 2020 cryo-EM. *Curr. Opin. Struct. Biol.* **68**, 184–193 (2021).

34. Najbauer, E. E., Ng, S. C., Griesinger, C., Gorlich, D. & Andreas, L. B. Atomic resolution dynamics of cohesive interactions in phase-separated Nup98 FG domains. *Nat. Commun.* **13**, 1494 (2022).

35. Stanley, G. J., Fassati, A. & Hoogenboom, B. W. Biomechanics of the transport barrier in the nuclear pore complex. *Semin. Cell Dev. Biol.* **68**, 42–51 (2017).

36. Sakiyama, Y., Panatala, R. & Lim, R. Y. H. Structural dynamics of the nuclear pore complex. *Semin. Cell Dev. Biol.* **68**, 27–33 (2017).

37. Nedelsky, N. B. & Taylor, J. P. Bridging biophysics and neurology: aberrant phase transitions in neurodegenerative disease. *Nat. Rev. Neurol.* **15**, 272–286 (2019).

Methods

Protein preparation. The FG-repeat domain of human Nup98 (Nup98[FG], residues 1–384) was cloned into bacterial expression vector pET28a (Novagen) between the restriction sites NheI and XhoI, keeping the N-terminal His-Tag for its purification. The recombinant Nup98[FG] was expressed in the *Escherichia coli* expression strain BL21(DE3). For the unlabelled protein, the bacteria were grown in Luria-Bertani (LB) medium supplemented with kanamycin to an optical density at 600 nm (OD_{600}) of ~0.8 and induced with 0.5 mM isopropyl-β-D-thiogalactoside (IPTG) overnight at 37 °C. For the uniform $^{13}C,^{15}N$ labelling of Nup98[FG], 5 ml of an overnight LB culture was inoculated into 1 l of M9 medium containing 4 g l^{-1} D-[$^{13}C_6$]glucose and 2 g l^{-1} $^{15}NH_4Cl$, supplemented with kanamycin and $^{13}C,^{15}N$-labelled ISOGRO powder growth medium. At an OD_{600} of ~0.8, the culture was induced with 0.5 mM IPTG overnight at 37 °C. For purification, cell pellets were sonicated in lysis buffer containing 100 mM Na_2HPO_4, 10 mM Tris and 10 mM 2-mercaptoethanol at pH 8.5, and the lysate was clarified by centrifugation. After centrifugation, the supernatant was discarded and the pellet containing Nup98[FG] was resuspended in denaturing buffer containing 8 M urea, 100 mM Na_2HPO_4, 10 mM Tris and 10 mM 2-mercaptoethanol at pH 8.5. After a second centrifugation step, the supernatant was loaded onto a self-packed nickel-nitrilotriacetic acid (Ni-NTA) column (Qiagen) equilibrated with denaturing buffer and bound protein was eluted with a buffer solution of 6 M urea, 100 mM Na_2HPO_4, 10 mM Tris and 10 mM 2-mercaptoethanol at pH 4.0. The same elution buffer was used to perform size exclusion chromatography on a Superdex75 26/600 column (GE Healthcare). A second size exclusion chromatography (using the same column equilibrated with 50 mM sodium phosphate buffer containing 1 mM tris(2-carboxyethyl)phosphin (TCEP) at pH 3) was conducted to remove any remaining impurities. The pure protein was concentrated by ultracentrifugation using a 5 kDa molecular weight cut-off membrane.

Nup98 peptides were prepared by solid-phase synthesis (Genscript). For EM and optical microscopy measurements, powdered lyophilized Nup98[FG85] was dissolved in pure water to reach a concentration of 1.5 mM, followed by incubation at 25 °C for 1 h.

Circular dichroism. Nup98[FG] samples were prepared at a concentration of 5 µM in 50 mM sodium phosphate buffer containing 1 mM TCEP at pH 3. The experiments at pH 7 were performed with the same samples, adjusting the pH to 7 just before measurement. CD data were collected in the range 185–280 nm using a Chirascan-plus qCD spectrometer (Applied Photophysics) at 25 °C with 1.5 s per point in 1 nm steps. The datasets are averages of ten repeat experiments. All spectra were baseline-corrected against buffer and smoothed (window size: 4) using GraphPad Prism.

Dynamic light scattering. Nup98[FG] samples were prepared at a concentration of 2.5 µM in 50 mM sodium phosphate buffer containing 1 mM TCEP at pH 3. Samples at pH 7 were prepared in the same buffer by adjusting the pH to 7 just before measurement and reducing the protein concentration to 100 nM. Both concentrations were optimized to avoid saturation of the detector. Measurements were conducted at 25 °C using a DynaPro NanoStar instrument (Wyatt Technologies) and NanoStar disposable microcuvettes. The samples were illuminated with a 120 mW air-launched laser at a wavelength of 662 nm and the intensity of light scattered at an angle of 90° was detected with an actively quenched, solid-state single-photon counting module. Data were acquired with an acquisition time of 5 s with a total of five acquisitions per measurement. The hydrodynamic radii were determined using the Dynamics (version 7.10.0.23) software package. The final values are given as the average and standard error of 12 measurements.

Light microscopy. For optical microscopy, Nup98[FG] samples were prepared at a concentration of 50 µM in 50 mM sodium phosphate buffer containing 1 mM TCEP at pH 3. The samples measured at pH 7 were prepared by adjusting the pH to 7 just before measurement. Nup98[FG85] samples were prepared by dissolving the powder in water to a concentration of 1.5 mM, incubating them for 1 h at 25 °C and then diluting to a concentration of 200 µM. Where indicated, ThT was added to reach a concentration of 50 µM. A total 5 µl of sample was loaded onto a slide and covered with a 18 mm coverslip. Differential interference contrast micrographs as well as fluorescent micrographs were acquired on a Leica DM6B microscope with a ×63 objective (water immersion) and processed using ImageJ. The exposure time of the fluorescent micrographs in the absence of ThT was 100 times longer.

FRAP measurements of Nup98[FG85] hydrogels were conducted with a Leica TCS SP8 confocal microscope using a ×63 objective (oil immersion) and a 488 nm argon laser line. Samples were loaded onto a slide and covered with a 18 mm coverslip. FRAP curves were acquired in 64×64 format at a scan speed of 1,000 Hz. Ten frames were collected for pre-bleaching and bleaching, and 500 frames for post-bleaching. Each frame corresponded to 523 ms. Regions of interest (ROIs) were bleached with 80% laser power, while a low laser intensity (5%) was used during recovery. Data were processed using FIJI software (NIH).

FRAP recovery curves were obtained by the standard protocols. Briefly, for each FRAP measurement, the intensities of the pre-bleaching, bleaching and post-bleaching ROIs were measured. A pre-bleaching ROI corresponds to a selected region in the droplet before bleaching, a bleaching ROI corresponds to the bleached area, and a reference ROI corresponds to an area that did not experience bleaching. The fluorescence intensity measured for each of the described ROIs was corrected for background by subtraction; a region where no fluorescence was detected was used to calculate the background.

NMR spectroscopy. NMR spectra were recorded at 5 °C on Bruker 700, 800, 900 and 950 MHz spectrometers equipped with triple-resonance cryogenic probes. For the full-length Nup98[FG] protein (200 µM in 50 mM sodium phosphate buffer, 1 mM TCEP, 0.01% NaN_3, 5% D_2O, pH 3), one-dimensional ^1H NMR and 2D ^1H–^1H TOCSY spectra were acquired at 800 MHz. In addition, ^1H–^{15}N and ^1H–^{13}C heteronuclear single quantum coherence (HSQC) and three-dimensional (3D) HNCO, HNcaCO, HNCA, HNcoCA, HNCACB and HNcoCACB spectra were recorded at 900 MHz. Samples were incubated each day for 30 min at 65 °C to resolubilize the protein.

The ^1H–^{15}N HSQC spectrum at pH 7 was acquired after adjusting the pH of the same sample from 3 to 7 (the dead time from changing the pH, mixing, transferring to the NMR spectrometer to starting the ^1H–^{15}N HSQC acquisition was ~5 min). The acquisition time of the ^1H–^{15}N HSQC spectrum at pH 7 was 15 min.

To assign the backbone resonances of the FG-repeat domain of Nup98 at pH 6.8, 17 overlapping Nup98 peptides (N-terminal acetylated, except residues 1–15, and carboxy-terminal amidated, except residues 5–44) were used covering the sequence of residues 1–384: 1–15, 5–44, 30–59, 45–84, 72–99, 95–124, 115–144, 137–170, 165–204, 189–218, 205–244, 225–264, 256–275, 265–304, 298–327, 322–351 and 345–384. Peptide concentrations of 2 mM were used for resonance assignment. The peptides were dissolved in 50 mM sodium phosphate buffer containing 0.01% NaN_3 and 5% D_2O at pH 6.8. Samples of Nup98(165–204) also included 1 mM TCEP. Several samples of Nup98(95–124), Nup98(298–327) and Nup98(345–384) were used due to their rapid aggregation. One-dimensional ^1H NMR, 2D ^1H–^1H TOCSY, NOESY, and ^1H–^{15}N and ^1H–^{13}C HSQC spectra of the 17 overlapped Nup98 peptides were acquired. All the NMR data were processed using TopSpin 3.6.1 (Bruker) and analysed with Sparky[38]. The ^1H NMR chemical shifts were referenced to 2,2-dimethyl-2-silapentane-5-sulfonate (DSS, 0 ppm), and the ^{13}C and ^{15}N NMR chemical shifts were indirectly referenced to DSS. All HA, HN, N, CA and CB resonances were assigned except for the N of prolines and the CA of the three asparagine residues before prolines. The overlapping peptide assignments were used to negate the influence of peptide ends on the final chemical shift assignments (Supplementary Fig. 15). The secondary structure and φ/ψ angles were calculated from the experimental HA, HN, N, CA and CB chemical shifts using TALOS-N (ref. [39]).

To study the reversibility of the amyloid-like structure in the FG-repeat domain of Nup98, samples of Nup98[FG] (200 µM) and Nup98[FG85] (100 µM) were prepared in 50 mM sodium phosphate buffer containing 1 mM TCEP, 0.01% NaN_3 and 5% D_2O at pH 3. Subsequently, the pH was adjusted to 7 and then back to 3, followed by incubation of the samples for 30 min at 65 °C in a water bath.

Transmission electron microscopy. Nup98[FG] samples at a concentration of 20 µM were dialysed in 50 mM HEPES buffer containing 1 mM TCEP at pH 3, and just before preparation of the grid, the pH was adjusted to pH 7 by adding NaOH. Samples were adsorbed onto 400-mesh carbon-coated copper grids and the buffer removed using filter paper. Subsequently, samples were stained by the addition of 1% uranyl acetate solution and dried with filter paper. The grids were imaged using a Talos L120C G2 electron microscope.

Cryoelectron microscopy. Nup98[FG85] fibrils were prepared by dissolving the peptide in water to reach a concentration of 1.5 mM, followed by incubation at 25 °C for 1 h. Subsequently, 3 µl of sample was applied to freshly glow-discharged R2/1 holey carbon grids (Quantifoil) and vitrified in liquid ethane using a Mark IV Vitrobot (Thermo Fisher) operated at 100% relative humidity and 4 °C. Cryoelectron microscopy was conducted with a Titan Krios transmission electron microscope (Thermo Fisher) operated at an accelerating voltage of 300 keV. Images were recorded at a nominal magnification of ×81,000 using a Quantum LS energy filter (Gatan) with the slit width set to 20 eV and a K3 direct electron detector (Gatan) in non-super-resolution counting mode, corresponding to a calibrated pixel size of 1.05 Å at the specimen level. In total, 4,180 images with defocus values in the range of –0.7 to –2.0 µm were acquired in movie mode with an acquisition time of 2.5 s. Each movie contained 40 frames with an accumulated dose of approximately 41 electrons per Å^2. The resulting dose-fractionated image stacks, containing all 40 frames, were subjected to beam-induced motion correction on the fly using Warp[40].

Nup98[FG85] fibrils were reconstructed using RELION-3.1 (ref. [41]) following the helical reconstruction scheme[42]. First, contrast transfer function (CTF) parameters were estimated for each motion-corrected micrograph using CTFFIND4 (ref. [43]), and only micrographs with an estimated resolution of ≤4.0 Å were considered for manual fibril picking. For 2D classification, we extracted particle segments using a box size of 600 pixels downscaled to 200 pixels and an interbox distance of 14 Å; pm1 and pm3 fibrils were successfully separated by this classification procedure, but pm2 and pm4 could not be separated owing to their close similarity. For

3D classification, the segments after 2D classification were re-extracted without downscaling using a box size of 250 pixels. The initial helical rise was estimated from the cross-over distances (180° helical turn), measured from the 2D class averages, and the helical rise was initially set to 4.75 Å. We performed the 3D classification several times, starting from a 60 Å low-pass-filtered featureless cylinder, until the separated β-strands along the helical axis became visible, and then optimized the helical parameters (the final parameters are reported in Supplementary Table 2). We were able to separate the pm2 and pm4 fibrils during the 3D classification. For 3D auto-refinement, all fibril polymorphs were reconstructed individually. Next, standard RELION post-processing with a soft-edged solvent mask that included the central 10% of the box height yielded post-processed maps (sharpening B factors are reported in Supplementary Table 2). The resolution was estimated from the value of the Fourier shell correlation (FSC) curve for two independently refined half maps at 0.143 (refs. [44,45]; Supplementary Fig. 7). Finally, the optimized helical symmetry was applied to the post-processed maps to yield the final maps.

The atomic models of the Nup98[FG85] fibrils were constructed de novo in Coot[46]. The high resolution of the cryo-EM maps allowed reliable modelling of the protein backbone and side chain rotamers. Refinement in real space was conducted using PHENIX[47,48] and Coot[46] in an iterative manner. The resulting models were validated with MolProbity[49] and their construction data are presented in Supplementary Table 3.

Solvation energy calculation. Nup98[FG85] fibril stability was calculated on the basis of the solvation energy using the software accessiblesurfacearea_v07.2d (ref. [30]). The number of layers for each fibril was set to five, and the energy of the middle layer was used. For pm1, nine layers were used because of the interactions between distant layers.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Data supporting the main findings of this study are available in the main text and the Supplementary Information. NMR chemical shifts are included in the Supplementary Information. Cryo-EM maps have been deposited in the Electron Microscopy Data Bank (EMDB) under the accession numbers EMD-13851 (pm1), EMD-13852 (pm2), EMD-13853 (pm3) and EMD-13854 (pm4). The corresponding atomic models have been deposited in the PDB under the accession numbers: 7Q64 (pm1), 7Q65 (pm2), 7Q66 (pm3) and 7Q67 (pm4). PDB accession codes 6XYO, 6IC3, 6SHS, 6NZN, 6VPS, 6SDZ, 6ZRF, 6XFM, 6MST, 6GK3, 7KWZ, 5O3O, 6TUB, 2RNM, 7BX7, 6WQK, 6UUR and 5V7Z cited in the text are publicly available in the PDB. Source data are provided with this paper.

References

38. Goddard, T. D. & Kneller, D. G. SPARKY 3 (Univ. California, 2014).
39. Shen, Y. & Bax, A. Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J. Biomol. NMR* **56**, 227–241 (2013).
40. Tegunov, D. & Cramer, P. Real-time cryo-electron microscopy data preprocessing with Warp. *Nat. Methods* **16**, 1146–1152 (2019).
41. Zivanov, J., Nakane, T. & Scheres, S. H. W. Estimation of high-order aberrations and anisotropic magnification from cryo-EM data sets in RELION-3.1. *IUCrJ* **7**, 253–267 (2020).
42. He, S. & Scheres, S. H. W. Helical reconstruction in RELION. *J. Struct. Biol.* **198**, 163–176 (2017).
43. Rohou, A. & Grigorieff, N. CTFFIND4: fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).
44. Rosenthal, P. B. & Henderson, R. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J. Mol. Biol.* **333**, 721–745 (2003).
45. Scheres, S. H. & Chen, S. Prevention of overfitting in cryo-EM structure determination. *Nat. Methods* **9**, 853–854 (2012).
46. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
47. Afonine, P. V. et al. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr. D* **74**, 531–544 (2018).
48. Liebschner, D. et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. D* **75**, 861–877 (2019).
49. Chen, V. B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D* **66**, 12–21 (2010).

Author contributions

A.I.d.O. performed NMR spectroscopy, biochemical experiments, fluorescence microscopy, prepared samples for cryo-EM, determined the cryo-EM structure of polymorph 1 and analysed the fibril structures. J.A.G., B.F. and G.F.S. performed cryo-EM structure determination of polymorphs 1–4 and analysed the fibril structures. C.D. recorded the cryo-EM datasets, A.S. performed FRAP and fluorescence microscopy, M.R. performed condensation experiments and M.-S.C.-O. prepared recombinant Nup98[FG]. A.I.d.O. and M.Z. designed the project.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41557-022-01035-7.

Correspondence and requests for materials should be addressed to Gunnar F. Schröder or Markus Zweckstetter.

Peer review information *Nature Chemistry* thanks Meytal Landau and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Corresponding author(s): Markus Zweckstetter

Last updated by author(s): Jul 26, 2022

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☒	☐	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	NMR data were acquired using Topspin versions 3.5pl7 or 4.0.8. DLS data were acquired with DYNAMICS v7.10.0.23. CD data were acquired with Pro-Data Chirascan Spectrometer Control Panel Application v4.7.0.194. Optical microscope pictures are acquired with Leica Application Suite X. cryo-EM data were acquired using SerialEM Version 3.9.
Data analysis	NMR data were processed with Topspin 3.6.1 and analyzed using Sparky 3.115. Secondary structure and phi/psi angles were calculated with TALOS-N. PyMOL v1.8.4.0 was used to build the TALOS-N models and represent macromolecular structures. Optical microscope micrographs and FRAP experiments were analyzed with ImageJ v1.52a. DLS data were analyzed with DYNAMICS v7.10.0.23. Nup98FG85 fibrils were reconstructed using RELION-3.1 and the micrographs CTF corrected with CTFFIND4. The atomic models of Nup98FG85 fibrils were built in Coot 0.8.9.1. Refinement in real space was conducted using PHENIX 1.19.2 and Coot 0.8.9.1. The resulting models were validated with MolProbity 4.5. Stability calculations based on solvation energy were performed with the software accessiblesurfacearea_v07.2d. Graphpad Prism 7 was used for plotting data.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Nuclear magnetic resonance chemical shifts are included in the extended data. Cryo-EM maps have been deposited in the Electron Microscopy Data bank (EMDB) under the accession numbers EMD-13851 (pm1), EMD-13852 (pm2), EMD-13853 (pm3), and EMD-13854 (pm4). The corresponding atomic models have been deposited in the Protein Data Bank (PDB) under the accession numbers: 7Q64 (pm1), 7Q65 (pm2), 7Q66 (pm3), and 7Q67 (pm4). All PDB codes cited (6XYO, 6IC3, 6SHS, 6NZN, 6VPS, 6SDZ, 6ZRF, 6XFM, 6MST, 6GK3, 7KWZ, 5O3O, 6TUB, 2RNM, 7BX7, 6WQK, 6UUR and 5V7Z) are publicly available in the PDB.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences   ☐ Behavioural & social sciences   ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample size does not apply for macromolecular structural data.
Data exclusions	No data have been excluded while reporting this study.
Replication	NMR data was acquired with two batches of the Nup98FG protein (natural abundance and 15N, 13C labeled) getting same results. CD and DLS data were acquired twice with same results. EM and cryoEM pictures of the Nup98FG85 fibrils were acquired with 5 samples getting always similar fibrils; the best one was used for the acquisition. Light microscopy images were acquired with two samples getting similar pictures.
Randomization	Randomization was not relevant to this study, because we are not working with any particular population (there is no population to randomly select from) so our experiments can be repeated.
Blinding	Blinding was not relevant to this study, because decision-making has no impact on the experiment and there is no risk of bias.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	Antibodies
☒	Eukaryotic cell lines
☒	Palaeontology and archaeology
☒	Animals and other organisms
☒	Human research participants
☒	Clinical data
☒	Dual use research of concern

Methods

n/a	Involved in the study
☒	ChIP-seq
☒	Flow cytometry
☒	MRI-based neuroimaging