Barrier properties of Nup98 FG phases ruled by FG motif identity and inter-FG spacer length

Sheung Chun Ng[1], Abin Biswas[2,3], Trevor Huyton[1], Jürgen Schünemann[1], Simone Reber[2], and Dirk Görlich[1]*

[1]Department of Cellular Logistics, Max Planck Institute for Multidisciplinary Sciences, Göttingen, Germany; [2]Quantitative Biology, IRI Life Sciences, Humboldt-Universität zu Berlin, Germany; [3]Department of Biological Optomechanics, Max Planck Institute for the Science of Light, Erlangen, Germany

* To whom correspondence should be addressed. E-mail: goerlich@mpinat.mpg.de

Abstract (150/150)

Nup98 FG repeat domains comprise hydrophobic FG motifs linked through uncharged spacers. FG motifs capture nuclear transport receptors (NTRs) during nuclear pore complex (NPC) passage, confer inter-repeat cohesion, and condense the domains into a selective phase with NPC-typical barrier properties. We found that shortening inter-FG spacers enhances cohesion, increases phase density, and tightens such barrier, all consistent with a sieve-like phase. Phase separation tolerated mutating the Nup98-typical GLFG motifs, provided domain-hydrophobicity remained preserved. NTR-entry, however, was sensitive to (certain) deviations from canonical FG motifs, suggesting co-evolutionary adaptation. Unexpectedly, we found arginines to promote FG-phase-entry apparently also by hydrophobic interactions/ H-bonding and not just through cation-π interactions. Although incompatible with NTR·cargo complexes, a YG phase displayed remarkable transport selectivity, particularly for engineered GFP^{NTR}-variants. GLFG to FSFG mutations made the FG phase hypercohesive, precluding NTR-entry. Extending spacers relaxed this hypercohesion. Thus, antagonism between cohesion and NTR·FG interactions is key to transport selectivity.

Introduction

The nuclear envelope (NE) separates the nucleus from the cytoplasm, confining the exchange of macromolecules to NE-embedded nuclear pore complexes (NPCs). NPCs are elaborate ~120 MDa structures assembled in 8-fold rotational symmetry from 30 different nucleoporins or Nups for short.

By combining cryo-electron microscopy, X-ray crystallography, and biochemical and genetic approaches, the architecture of the NPC scaffold has been elucidated in remarkable detail[1–7]. The NPC scaffold encloses a ~60 nm wide central channel[6,8,9], which is controlled by a sieve-like permeability barrier to suppress an uncontrolled intermixing of nuclear and cytoplasmic contents[10,11]. Traditionally, two modes of NPC passage have been distinguished, 'passive' and 'facilitated'. The passive passage is efficient for small molecules but already severely restricted for 5 nm (30 kDa)-sized objects. Nuclear transport receptors (NTRs) are not bound to this size limit. Instead, they can engage in facilitated translocation and carry even very large cargoes through NPCs.

NTRs of the importin β superfamily[12] circulate between the two compartments, draw energy from the RanGTPase system, and pump cargoes against concentration gradients. Importins, such as importin β itself or transportin, capture cargo in the cytoplasm, traverse NPCs, and release cargo upon RanGTP-binding into the nucleus. Following their return to the cytoplasm, GTP-hydrolysis, and RanGDP-release, they can import the next cargo. Exportins (e.g., Exportin 1/Xpo1 a.k.a. CRM1) function similarly but recruit cargoes from the nucleus (along with RanGTP) and release them into the cytoplasm.

Each of such transport cycles transfers one Ran molecule from the nucleus to the cytoplasm. NTF2 (Nuclear transport factor 2) returns RanGDP to the nucleus[13] and thus has to pass NPCs more frequently than any other NTR. A single NPC can accommodate up to 1000 facilitated events or a mass flow of 100 MDa per second[14], with transit times in the order of 10 milliseconds for simple importin·cargo complexes[15,16] and ~25 milliseconds for 60S pre-ribosomes[17,18]. NPCs have the capacity to transport many NTR·cargo complexes in parallel[19]. NPC passage *per se* is reversible and energy-independent. Active, directed transport requires additional energy input, e.g., by the aforementioned RanGTPase system.

FG nucleoporins (FG Nups; ~10 different ones in any given species) contain so-called FG repeat domains[20–22]. These domains comprise numerous FG (phenylalanine-glycine) motifs, are of low sequence complexity, and are intrinsically disordered[23,24]. FG domains are

anchored to the inner scaffold of the NPCs[2,3] and bind NTRs during facilitated translocation[25,26]. Moreover, they can engage in multivalent, cohesive interactions via FG motifs and other hydrophobic groups to form a 'selective FG phase'[14,27–33].

State of matter similar to the FG phase was later described also as biomolecular condensates or liquid-liquid phase separation (LLPS)[34–40], e.g., in germline granules. Recently, it has become clear that multivalent interactions between interaction sites, called "stickers" drive phase separation and valence, strength, and patterning of stickers are key determinants of phase separation or properties of these condensates[39,41–43].

FG phases can easily be reconstituted and indeed behave like the permeability barrier of NPCs, i.e., they are a good solvent for NTRs and NTR·cargo complexes but exclude 'inert' macromolecules that lack FG interactions[28,29,32,44,45]. Transport through NPCs can thus be described as mobile species partitioning into this FG phase and exiting on the opposite side. Another perspective considers that multivalent FG contacts result in a 3D sieve whose mesh size determines the (passive) sieving properties of the phase. The binding of NTRs to FG motifs competes with FG-FG interactions and allows NTRs to 'melt' (along with their cargoes) through the meshes.

Traditionally, NPCs have been assumed to be sieve-like barriers with a 5 nm size limit for passive passage[11]. This corresponds to the diameter of GFP. Later, however, it became evident that there is actually a continuum between passive and facilitated NPC-passage[44,46]. Indeed, transport rates are not only determined by the size but also by surface properties of the mobile species, with negative charges or lysines impeding NPC passage, while "FG-philic" residues (like R, H, C, or hydrophobic ones) speed up. Based on these principles, GFP has been engineered to pass NPCs either very slowly (super-inert GFP variants) or very rapidly (GFP^{NTR} variants), with the extreme versions (of identical size) differing more than 10 000-fold in their NPC passage rates[44].

High transport selectivity of an FG phase requires the local FG repeat concentration to exceed a threshold of ~200 mg/ml[28]. Such high local concentration can only partially be imposed by anchoring the ~15 MDa of FG domain mass[47] to the NPC scaffold. The other key contribution is localized phase separation, in particular, of Nup98 FG domains[32].

FG domains from Nup98 or its homologs[22,32,48] (e.g., Nup116, Nup100, and Nup145 in yeast) are the most cohesive[32,49] and crucial ones for maintaining the permeability barrier of NPCs[31,46,50,51]. They are typically dominated by GLFG motifs and comprise ~500-700

residues. Of all FG domains, they feature the highest number (~50) and density of FG or FG-like motifs (one per ~12 residues). They are extremely depleted of charged residues, experience water as a poor solvent, and readily phase-separate from low μM or even sub-μM concentrations to form condensed FG phases with 200-300 mg/ml or 100-300 mM FG repeat units[32]. Such self-assembled Nup98 FG phases recapitulate the transport selectivity of NPCs particularly well[32,44,45]. They exclude inert (i.e., "FG-phobic") macromolecules (such as the 25 kDa-sized mCherry) to partition coefficients below 0.05 while allowing rapid entry of NTRs and NTR·cargo complexes. NTF2, for example, reaches a partition coefficient of ~2000. RanGTPase-controlled cargo import and export scenarios can also be recapitulated faithfully with a single Nup98 FG phase[45]. Therefore, Nup98 FG phases capture the essential features and represent simple experimental models of NPC-typical transport selectivity.

We now systematically tested the relevance of conserved Nup98 sequence features and found charge depletion and balanced hydrophobicity to be key parameters for the phase system. Introducing even a moderate negative charge into the repeats abolished phase separation. Surprisingly, however, there is no fundamental requirement for FG motifs. The GLFG→GLFS mutation, for example, preserved the assembly of a selective phase that excluded mCherry and still allowed a very efficient influx of all NTR·cargo species tested. Even a YG phase (with phenylalanines changed to tyrosines) displayed remarkable transport selectivity, e.g., for the engineered 3B7C GFP^{NTR}-variant. YG motifs appear, however, rather incompatible with NTR-binding, suggesting co-evolutionary adaptations between wild-type FG domains and NTRs. We now matched this preference pattern of native NTRs with an advanced $tetraGFP^{NTR}$ variant that is particularly well adapted to GLFG motifs. Shrinking inter-FG spacers favoured phase separation, increased the intra-phase protein density, caused a stricter exclusion of FG-phobic species, and impeded the entry of NTRs carrying cargo. Extending spacers had the opposite effects, supporting the model of a sieve-like barrier whose mesh size is ruled by the distance of connection points. NTRs bind FSFG motifs well; however, mutating GLFG in the Nup98 FG domain to more hydrophobic FSFG motifs resulted in a hypercohesive phase repelling not only inert species but also precluding any NTR entry. Extending the inter-FSFG spacers relaxed this hypercohesive phenotype. This is consistent with a permeability barrier maintained by cohesive repeat contacts and with NTRs traversing the barrier by resolving such cohesive interactions locally.

Results

Experimental approach

Nup98 is central to the NPC permeability barrier. Its FG domain is highly cohesive and comes with several remarkable evolutionarily conserved features, such as depletion of charges, a dominance of GLFG motifs, and a rather high FG density with about eight motifs per 100 residues (Supp. Table 1). Conservation implies evolutionary pressure and functional relevance, and so we wondered how changing Nup98-typical features would impact the phase behaviour.

In order to characterize FG domain variants, we used codon-optimized expression vectors to produce them in *E. coli*, performed purifications under denaturing conditions, and prepared 1 mM FG domain stocks in 4 M guanidinium hydrochloride (to keep them initially non-interacting). Phase separation was then initiated by a 100-fold dilution in a guanidinium-free buffer[32,45]. This setup was used to determine saturation concentrations (C_{sat}) by measuring the fractions that remained soluble after pelleting the condensed FG phases. In separate assays, we used confocal laser scanning microscopy (CLSM) to determine how various fluorescent probes partition into the FG phases.

A simplified FG domain as a starting point

Our starting point was the FG domain from *Tetrahymena thermophila* MacNup98A (Fig. 1a) because it is already well characterized and avoids complications such as an elevated amyloid propensity or a need for O-GlcNAc modifications[49]. For conceptual clarity, we used a derivative that had been simplified[45], namely by (1) converting all FG and FG-like motifs to GLFG motifs and (2) shifting the GLFG motifs to equidistant positions in the sequence. This simplified sequence contains 52 GLFG motifs in total (each separated by a spacer of eight *variable* amino acids); it has a repeat length of 12 residues, containing one GLFG motif per 12 residues, and is referred to as "$GLFG_{52x12}$" (Fig. 1b). We tried to preserve the original overall hydrophobicity and amino acid composition and kept the intervening Gle2-binding sequence (GLEBS) unchanged (see Supp. Note 1 for complete sequences of the FG domains/ variants). The simplification had only marginal effects on the phase separation properties and the transport selectivity of the resulting FG phase (Figs. 1-3; and ref.[45]), suggesting that $GLFG_{52x12}$ captures the barrier properties of the original domain indeed very well.

A negative net charge antagonizes FG phase separation

We first tested whether the observed uncharged nature of Nup98 FG repeats is relevant and observed that introducing one negative charge (as an aspartate) per repeat completely abolished phase separation (see GLFG//D_{52x12} variant in Fig. 1c). This can be rationalized by electrostatic repulsion and thus an increase in water solubility of the spacers and also the entire domain[52,53]. It is consistent with the extreme selection against Asp and Glu in native Nup98 FG domains and other cohesive FG domains[54,55], and against a negative net charge in FG domains in general.

Shorter inter-FG spacers enhance cohesive interactions

In the next step, we varied the length of the inter GLFG-spacers while keeping their amino acid composition and the number of GLFG motifs constant (Fig. 1d). For shorter spacer variants ($GLFG_{52x10}$ and $GLFG_{52x11}$), residues were deleted such that the remaining spacers matched the original composition as closely as possible (Supp. Table 2). Likewise, amino acid insertions to extend the spacers ($GLFG_{52x13}$, $GLFG_{52x14}$, and $GLFG_{52x15}$) were chosen according to the same constraint.

These spacer length variations had a striking effect on phase separation, with shortening by one residue (per spacer) decreasing C_{sat} by a factor of ~3-5 and thus favouring phase separation. C_{sat} of $GLFG_{52x10}$ (with six-amino-acid spacers between the GLFG motifs) was so low that it could hardly be measured by the pelleting assay. Conversely, extending the spacers increased the saturation concentration and thus disfavoured phase separation.

Above their saturation concentrations, all FG domains of the spacer length series assembled into µm-sized, near-spherical FG phases ("FG particles") that stained bright with Hoechst 33342 (Fig. 2a). This dye is environmentally-sensitive and shows a greatly enhanced fluorescence upon shielding from water. It is best known as a DNA-specific probe[56,57] but also stains FG phases in a way that depends on the presence of cohesive interactions[45]. Indeed, the Hoechst signals in Fig. 2a correlate clearly with the cohesiveness of the FG domains: the variant with the shortest spacers and the lowest C_{sat} ($GLFG_{52x10}$) stained 2.5 times brighter than the standard $GLFG_{52x12}$ phase and ~13 times brighter than the least cohesive one ($GLFG_{52x15}$). Thus, the local FG motif concentration in the phase and/or occupancy of cohesive interactions appear to increase with shorter spacers.

To determine the protein content of the phases (C_{dense}), we applied optical diffraction tomography (ODT), which exploits the refractive index (RI) as a robust concentration measure (Fig. 2b, d; refs.[58–60]). This way, we obtained a C_{dense} of ~470 mg/ml for the original Mac98A and the standard $GLFG_{52x12}$ variant. This number is higher than previous estimates based on signal-matching between Alexa488-labelled FG domains and a dilute phase marker[32]; but some fluorescence quenching within such a dense phase might explain this difference. The protein concentration of 470 mg/ml (380 mM FG motifs in the $GLFG_{52x12}$ phase) leaves a solvent content of ~67% (assuming a partial specific volume of 0.7 ml per gram of FG domain). Shrinking the spacers between the GLFG motifs from 8 to 6 residues (in $GLFG_{52x10}$) increased C_{dense} considerably to ~600 mg/ml (550 mM FG motifs), leaving a solvent content of 58%. Conversely, extending the spacers to 11 residues decreased the FG domain concentration to ~310 mg/ml (210 mM FG motifs), leaving 78% solvent.

There was a remarkable log-linear relationship between C_{sat} (Fig. 2c) or the C_{sat} : C_{dense} ratio and the spacer length. On the other hand, the C_{sat} : C_{dense} ratio is also linearly related to the free energy change of phase separation[61]:

$$\Delta G = RT \cdot \ln\left(\frac{C_{sat}}{C_{dense}}\right) \qquad (1)$$

where R is the gas constant (8.31 J/mol·K) and T the absolute temperature in Kelvin. ΔG thus depends inversely on the spacer length (Fig. 2e). Perhaps, shorter spacers entropically favour hydrophobic interactions of the LF clusters[62,63] and/or enhance the local cooperativity of such cohesive contacts[52,53]. The effect can also be interpreted as short spacers increasing the overall hydrophobicity of the FG domain and thus decreasing its water solubility.

Shorter spacers make the FG phase a stricter barrier

Nup98 FG phases provide transport selectivity because they can exclude non-interacting ("FG-phobic") macromolecules, thereby impeding their passage and thus controlling fluxes through NPCs. One mechanistic interpretation is that the phase-entry of mobile species requires a local disruption of cohesive interactions, which imposes a thermodynamic penalty. By providing a spectrum of cohesion strengths, the spacer length series now allowed to explicitly test this concept.

For a first test (Fig. 3), we used four monomeric fluorescent proteins (of ~30 kDa) as probes, whose surfaces ranged from very inert/FG-phobic to weakly FG-philic (mCherry, EGFP, efGFP_8Q, efGFP_8R; ref.[44]). The $GLFG_{52x10}$ phase with the shortest spacers excluded all four proteins very well (with partition coefficients of ≤0.05 to 0.07). The spacer extension increased the partitioning within the $GLFG_{52x15}$ phase, namely to 0.12 (for mCherry) or even to 1.5 (for efGFP_8R). Thus, less cohesion appears to lower the energetic barrier for phase entry.

We next asked whether the partitioning of NTRs is also sensitive to spacer length and observed that NTF2 (coupled with Alexa488 for tracking) partitioned in all FG phases equally well with very high partition coefficients between 2700 and 3300. This contrasts the behaviours of the just mentioned fluorescent proteins but can be explained by NTF2 engaging in more FG contacts when immersed in a phase with a higher FG density. We would assume that the entire NTF2 surface is available for such interactions and provides a sufficient number of binding sites to respond to the higher FG density.

Indeed, we observed a rather different trend when NTF2 was complexed with its cognate cargo, RanGDP (in this case, Ran was labelled with Atto488). The NTF2·Ran complex partitioned well (with coefficients > 200) in the $GLFG_{52x12}$ phase and any phase with longer spacers; however, it arrested at the surface of the $GLFG_{52x10}$ phase and essentially failed to enter. This is the expected phenotype when the FG-philic NTF2 immerses into the phase while Ran still faces the bulk solvent.

Next, we tested three additional NTR·cargo complexes, namely (1) transportin·M9-EGFP (with M9 being the nuclear import signal of hnRNP A1[64]), (2) importin β·IBB-EGFP (with IBB being the Importin Beta-Binding domain from importin α - a very strong import signal[65]), and (3) the exportin Xpo1 bound to RanGTP and NES-EGFP (with the NES being a PKI nuclear export signal[66]). All three complexes were arrested at the surface of the short-spacer $GLFG_{52x10}$ phase but efficiently entered the long-spacer phase variants.

The inefficient entry of these species into the short-spacer FG phases could be explained by their larger sizes (53-170 kDa) as compared to NTF2 alone (30 kDa). An alternative explanation is that entry is impeded by larger FG-phobic surface regions of Ran and/or EGFP. Indeed, the partitioning of the importin β·IBB-fusion into the $GLFG_{52x10}$ or $GLFG_{52x11}$ phases increased 30 or 1000-fold, respectively, when EGFP was exchanged for $sffrGFP7^{NTR}$, a GFP derivative engineered for an FG-philic surface[44]. Likewise, the tetrameric (110 kDa)

GFP^{NTR}_3B7C, which was also engineered for FG-philicity and rapid NPC passage[44], partitioned very well in all FG phases. Therefore, we assume that shorter spacers not only increase cohesion but, in consequence, also cause a stricter exclusion of macromolecules with inert surfaces. In line with this, such a strict selectivity can also be relaxed by just lowering the salt concentration in the assay buffer, which weakens hydrophobic inter-repeat contacts (Supp. Fig. 1; refs.[33,61]). Increasing the salt concentration had the opposite effect, in particular for the long-spacer FG phases. These data show that selectivity, in general, depends on the strength of cohesive interactions in the FG phase.

Phase separation tolerates FG motif mutations if overall hydrophobicity is preserved

We next analysed a series of GLFG motif mutations, again in the context of the standard-spacer GLFG$_{52x12}$ variant. In line with previous observations[27,31,33,54,67], a reduction in the motif hydrophobicity essentially abolished phase assembly (Fig. 4a). This applied to the GLFG→GAFG mutation (Leu replaced by Ala), the similar GLFG→GAYG mutation, as well as to the GLFG→GLLG mutation (Phe changed to a less hydrophobic Leu). Phase separation of the latter was, however, restored when a spacer residue was mutated to a more hydrophobic leucine (GLLG//L$_{52x12}$). Note that the double mutation is roughly neutral with respect to the count of hydrophobic (aromatic+ aliphatic) carbons per repeat unit.

Likewise, shifting the leucine from the GLFG motif to the spacer without altering the amino acid composition (GXFG//L$_{52x12}$) preserved phase separation. In addition, we tested several additional motifs (GFLG, GLYG, GIFG, GLFA, SLFG, and GLFS) with similar hydrophobicity as the GLFG motif and observed phase separation as well. The protein concentration in the resulting dense phases was also remarkably similar (Fig. 4b, c). Thus, phase separation tolerates deviations from the GLFG motif sequence as long as the overall hydrophobicity is not reduced. This is consistent with the cohesive behaviours of FG domains that lack GLFG motifs but are sufficiently hydrophobic (e.g., Nup42, the N-terminal Nsp1, or the Nup62/54/58 FG domains; refs.[49,54,55,67,68]).

Changing to GLYG motifs lowered C_{sat} below the detection limit of our assay. This is in line with the higher interaction strength of Tyr than Phe, as reported in other systems[69–72].

FG-like phases can show high transport selectivity even without FG motifs or phenylalanines

We then tested all the above-mentioned FG phases (with GLFG, SLFG, GIFG, GXFG//L, or SLFG motifs) and FG-like phases (with GFLG, GLFA, GLFS, GLYG, or GLLG//L motifs) in permeation assays. All of them excluded mCherry (29 kDa) very well (with partition coefficients of less than 0.05 to 0.1) but allowed the accumulation of the tetrameric (110 kDa) GFP^{NTR}_3B7C variant to partition coefficients of 200 to 2900 (Fig. 5 and Fig. 6). This even applied to repeats where the FG motifs were replaced by YG motifs ($GLYG_{52x12}$) or repeats lacking Phe and Tyr altogether ($GLLG//L_{52x12}$). FG- and FG-like phases can thus show very high transport selectivity – independently of which hydrophobic motif confers the cohesive interaction. The engineered GFP^{NTR}_3B7C variant appears here to be a rather promiscuous client of these phases.

Effects of FG motif mutations on phase entry of NTRs

We then tested how various NTRs and NTR·cargo complexes partition into our set of FG and FG-like phases. Mutating the GLFG to GIFG motifs (i.e., Leu replaced by Ile) had no noticeable effect on transport selectivity (Fig. 5). Given that leucine and isoleucine are very similar amino acids, this exchange appears indeed rather subtle. The change to SLFG motifs also had no detrimental effect on NTR entry but increased the partition coefficients for most NTRs and NTR·cargo complexes (up to four-fold). This is consistent with SLFG motifs being rather common, e.g., in the wildtype Nup98 FG domain from *C. elegans* (Supp. Table 1), and that the slightly weaker cohesion of this motif (as measured by phase separation; Fig. 4) might promote NTR influx (see also below).

The GLFG→GLFS mutation marks a departure from FG motifs. Indeed, GLFS motifs are very rare in Nup98 FG domains (Supp. Table 1). Nevertheless, it had a surprisingly small effect: just a ~three-fold lower partition of NTF2 and ~two-fold higher of importin β, transportin, and Xpo1 complexes.

Other mutations, however, were very detrimental to NTR interactions (Fig. 6): The GFLG mutation (with swapped Phe and Leu positions) reduced phase entry of importin β and the exportin Xpo1 drastically (20- to 200-fold). The GXFG//L mutation (Fig. 4) was created by swapping the leucine of the GLFG motif with a residue of the spacer (X= G, N, Q, S, T, A, P…). This reduced the influx of the importin β and transportin·cargo complexes nearly to the

extent of phase-exclusion – though it still allowed a quite efficient accumulation of NTF2 and the Xpo1 complex. The tyrosine GLYG mutant, as well as the also phenylalanine-free GLLG//L mutant, essentially excluded importin β, transportin, and Xpo1 complexes. Likewise, phase entry of NTF2 and the NTF2·RanGDP complex was 50-fold reduced. These observations suggest co-evolutionary adaptations between FG domains and FG-binding sites in native NTRs, with FG-binding pockets[73–76] being probably more constrained than surface sites.

Arginines promote FG phase entry not just through cation-π interactions

Lysine and arginine are very similar amino acids with positively charged sidechains. On the surface of a mobile species, however, the two have profoundly different effects. Exposed lysines render a mobile species FG-phobic and slow in NPC passage. By contrast, arginines favour partitioning in the FG phase and can speed up NPC passage by a large factor[44]. The $GLFG_{52x12}$ phase recapitulated this behaviour (Fig. 7). It accumulated the sffrGFP4^{NTR} variant (with 25 solvent-accessible arginines and no lysines) to a partition coefficient of ~20, but it excluded the sffrGFP4 25R→K variant (with 25 solvent-exposed lysines instead) down to a partition coefficient of 0.1. The enhancing effect by arginines can be explained by the positively charged guanidinium group engaging in cation-π interactions with the π electron cloud of phenylalanines[77]. The multiplicity of arginines amplifies the effect to a ~200-fold difference in partition coefficient between the two GFP variants.

The GLYG phase also accumulated sffrGFP4 while excluding the 25R→K variant. This was expected - given that tyrosine is also known for forming similar cation-π interactions with arginine. The GLLG//L phase, however, contains neither phenylalanines nor tyrosines, and yet it accumulated sffrGFP4 80-fold higher (to a partition coefficient of 8) than the 25R→K variant (of 0.1). This unexpected observation strongly argues for functionally relevant arginine interactions that are distinct from cation-π (see below).

TetraGFP^{NTR}: an engineered FG (GLFG) phase marker

We previously described the GFP^{NTR}_3B7C variant, which was engineered to partition particularly well into FG phases, to cross NPCs rapidly, and to provide a crisp and specific NPC stain when incubated with digitonin permeabilised HeLa cells[44]. We used the X-ray structure of this homotetramer for Rosetta optimization and introduced a (buried) M225F

mutation to stabilize the interface between the subunits. We refer to the new variant as TetraGFP^{NTR}. The mutation also made the probe more similar to native NTRs, namely by selectively suppressing the partitioning of TetraGFP^{NTR} into the GFLG, GXFG, and GLYG phases while maintaining the very high accumulation in the GLFG phase (Fig. 8a). This enhanced specificity is also evident from NPCs-stains in digitonin-permeabilised cells, where TetraGFP^{NTR} gave a crisp signal with less background than its progenitor 3B7C (Fig. 8b, c). TetraGFP^{NTR} also showed a preference for GLFG over the FSFG phase (see the section below). How the mutation at the tetramer interface translates to enhanced GLFG selectivity is still unclear. However, TetraGFP^{NTR} can serve as a very useful phase marker and translocation probe.

Changing GLFG to FSFG motifs results in hyper-cohesiveness

GLFG motifs are dominant in Nup98 domains and their homologs, like the yeast Nup116 FG domain. This is, however, different in other Nup FG domains, an example being yeast Nsp1[20]. Its regular sub-domain (residues 274-601) contains exclusively (more hydrophobic) FSFG motifs separated by rather long, highly charged, and thus anti-cohesive spacers[68].

We aimed to compare FSFG and GLFG motifs directly and mutated all GLFG motifs in the $GLFG_{52x12}$ domain to FSFG motifs (Fig. 9a). The $FSFG_{52x12}$ domain phase separated very well and perfectly excluded mCherry, indicating that it forms an effective barrier. NTRs bound to the surface but failed to reach the interior of the phase. This applied to all NTR species and, strikingly, even to cargo-free NTF2 and to GFP^{NTR}_3B7C, which had entered the above-described, very strict $GLFG_{52x10}$ phase very well.

Two possible explanations for this phase behaviour came to our mind. First, the cohesion between FSFG motifs might be so strong that NTRs cannot pass these contacts. Second, the FSFG motifs are poor NTR binders when brought into the context of the uncharged Nup98 inter-FG spacers. To distinguish between these scenarios, we mutated (in two variants) one-third of the FSFG motifs to SSSG motifs (Fig. 9b, c). This lowered the FG density and reduced the overall hydrophobicity approximately to the level of the original GLFG domain. These partial F→S mutants indeed allowed a very efficient accumulation of all tested NTR species with the importin β·IBB-EGFP complex reaching a partition coefficient of 200-300, NTF2 of ~2000, and 3B7C of ~1000. As the remaining FSFG motifs remained in their initial context, we can conclude that they are fully proficient in NTR-binding, also in the context of

the non-charged Nup98 spacers. In turn, this suggests that the cohesion between FG motifs of the $FSFG_{52x12}$ domain is so strong that even the otherwise productive FG·NTR interactions cannot disengage them.

Antagonism between FG cohesion and NTR·FG interactions

The very strong cohesion of the $FSFG_{52x12}$ domain is also evident from its extreme phase separation propensity. C_{sat} here is so low that the soluble fraction remained undetectable in our phase separation experiments. C_{sat} decreases exponentially with the number of repeats[61], and even when reducing the number of $FSFG_{x12}$ repeats from 52 to 29 ($FSFG_{29x12}$, Fig. 9d), C_{sat} (0.2 μM) was just above the detection limit. The very high protein concentration (650 mg/ml) in the $FSFG_{52x12}$ phase and the very strong Hoechst signal (Fig. 10b) are also consistent with extreme cohesion and high occupancy of inter-repeat contacts.

In the next step, we asked if an NTR-permissive FSFG phase can also be generated with regular repeats. For this, we extended the inter-FSFG spacer length of the $FSFG_{52x12}$ domain from initially 8 to 11 residues ($FSFG_{52x15}$) or 14 residues ($FSFG_{52x18}$; Fig. 10a). Indeed, the resulting dense phases dropped in protein concentration (to 530 and 410 mg/ml, respectively) and in Hoechst signal (7 and 60-fold, respectively). This came with relaxed barrier properties: The $FSFG_{52x15}$ phase accumulated NTF2 and 3B7C to partition coefficients of 2600 and 1300, respectively. Extending the spacer length to 14 residues ($FSFG_{52x18}$) then allowed also the importin β·IBB-EGFP complex to reach an intra-phase partition coefficient of 90. The hypercohesive nature of the FSFG motifs (Fig. 10b and Supp. Fig. 2) can thus be balanced by longer inter-FG spacers that lower the cohesion propensity, probably through entropic effects. In any case, this set of experiments illustrates nicely the antagonism between inter FG-cohesion and FG·NTR interactions and suggests that selective NPC passage indeed happens by NTRs transiently disengaging cohesive contacts within the FG phase.

Discussion

Conserved Nup98 FG domain features

Nup98 FG repeat domains readily phase-separate to form an FG phase with NPC-typical transport selectivity. They are intrinsically disordered, but in contrast to typical intrinsically disordered domains, they show a high degree of sequence conservation[32]. In fact, the Nup98 FG domain of vertebrates is as conserved as its globular anchor domain. This points to considerable evolutionary pressure to maintain its interactions. Conserved features of Nup98 FG domains include (1) an overall length of about 500-700 residues, (2) a low complexity sequence with an extreme selection against charged residues, against negatively charged ones in particular, but also against tyrosines, histidines, cysteines, tryptophans, valines or isoleucines, (3) a rather constant hydrophobicity with a density of one FG motif per 12 residues, and (4) a prevalence of GLFG motifs.

The preferred overall length of Nup98 FG domains probably relates to providing enough FG mass per NPC and a sufficient contour length for spanning the central NPC channel. Likewise, the phase separation propensity depends on the number of repeats[61] .

We now found that the absence of negative charges is a fundamental requirement for inter-repeat cohesion. Similarly, the hydrophobicity and FG motif density are crucial. Replacing the GLFG with more hydrophobic FSFG motifs or increasing the FG motif density by shrinking the inter-FG spacers resulted in an intriguing phenotype, namely hypercohesive FG phases. This increase in cohesion was evident by orthogonal readouts: (1) a higher phase separation propensity/ lowered saturation concentration, (2) an increased protein concentration in the dense phase, and (3) an increased fluorescent signal by Hoechst 33342, which scales not just with the local FG motif concentration but apparently also reveals the occupancy of cohesive contacts.

The most striking functional consequence of hypercohesion is a stricter barrier – to the extent that NTR·cargo complexes or even cargo-free NTRs get arrested at the FG phase's surface and fail to enter. An interpretation is given by the selective phase model[14], which assumes that FG-mediated inter-repeat contacts are obstacles blocking the undesired passage of inert materials. NTRs bind FG motifs, thereby resolving inter-repeat contacts in their immediate vicinity and thus catalysing their own passage. In a hypercohesive phase, these contacts are stabilised by a higher energetic penalty than NTRs can release by FG-binding. As a result, even NTRs remain excluded. If this happened in living cells, then transport of cargo would

get blocked – with lethal consequences or at least with impaired fitness. This consideration explains the evolutionary pressure for keeping the parameters 'FG motif density' and 'hydrophobicity' of the Nup98 FG domains within their remarkably narrow tolerance.

Arginines appear to promote NPC passage also by general hydrophobic and H-bonding interactions

Lysine and arginine have long hydrophobic sidechains with terminal positive charges; both engage in salt bridges and are highly soluble in water. Yet, the two have opposite effects on NPC passage: Exposed lysines render a mobile species slow in NPC passage, while arginines can drastically increase the rate of NPC passage[44]. This behaviour is well recapitulated with *in vitro* assembled FG phases (see Fig. 7), where e.g., the sffrGFP4 variant (with 25 exposed Arg) shows a 100-fold higher partition coefficient than its 25-lysine counterpart.

This is so far best explained by the positively charged guanidinium group of arginine engaging in cation-π interactions[77] with the phenylalanines of the FG domain. Lysine can engage in similar interactions; however, the planar guanidinium of arginine is favoured here by better shape complementarity to the planar phenylalanine ring. Furthermore, a planar guanidinium is more easily released from the tetragonal hydrogen-bonding network of water than the (tetragonal) ammonium group of lysine.

We have now characterized an analogous dense phase, the GLLG//L phase, whose phase separation is driven by the hydrophobicity of leucines. It accumulated sffrGFP4^{NTR} 80-fold higher than the 25R→K variant (Fig. 7). This is a similar selectivity factor as found for the analogous FG or YG phase. However, since this domain contains neither phenylalanines nor tyrosines, we can exclude cation-π interactions as an explanation.

We suspect two kinds of alternative interactions. First, arginines might engage more easily also in general hydrophobic interactions (i.e., between the aliphatic portions of its side chain and the three leucines and other aliphatic groups in the repeats) – again because its planar guanidium group interacts less favourably with water than the lysyl ammonium. Second, the guanidinium group is an excellent hydrogen bond donor and probably bonds to carbonyl oxygens from the polypeptide backbone as well as from asparagine or glutamine sidechains of the phase. The extremely high concentration of repeat units of ~400 mg/ml would favour such interactions. Taken together, this suggests a mixed-mode interaction of arginines with

the *in vitro* assembled FG phases and - by extrapolation - also during passage through the permeability barrier of NPCs.

Although extremely depleted in FG domains, arginines are common in other phase-separating intrinsically disordered domains, where (cation-π) interactions with tyrosines/ phenylalanines have been reported[63,72,78]. Based on the above considerations, it might be safe to speculate that in these cases, arginines engage also in general hydrophobic and H-bonding interactions.

Why is the NPC permeability barrier formed by FG and not by YG repeats?

We previously reported that F→Y mutated Nsp1 FG repeats can complement an otherwise lethal FG domain deletion[27]. This is best explained by this YG mutant being able to contribute to the NPC permeability barrier, at least in the presence of the remaining eight FG domains. We now found that an FG→YG mutated domain not only phase-separates very well, but also the resulting dense phase perfectly excludes mCherry while strongly accumulating the four times larger GFP^{NTR}_3B7C variant (Figs. 6-8). This remarkably high transport selectivity poses the question of why the NPC barrier relies on FG but not on YG repeats.

Resilins[79], a class of naturally occurring YG repeat proteins, might provide one answer. These extracellular elastomers store mechanical energy to allow dragonflies a very energy-efficient flight and fleas their remarkable jumping abilities. Resilins mature by a radical-induced formation of covalent Y-Y bonds between the YG repeat units. The same reaction within the NPC permeability barrier (e.g., under conditions of oxidative stress) would lead to an irreversible clogging of the barrier. Since NPC are very long-lived structures (at least in terminally differentiated cells), even a slow rate of this reaction would deteriorate NPC function over time. Thus, the inherent propensity of the phenolic tyrosine side chains to covalent crosslinking could explain the strong selection against tyrosines in Nup FG repeat domains.

Other reasons might be that tyrosines are more promiscuous than phenylalanines in their interactions[44], and that YG-rich proteins, such as the yeast prion Sup35[80,81], are prone to form potentially detrimental amyloid-like structures. A final reason might be that an early co-evolutionary adaptation of NTRs to FG repeats was a hurdle for a later switch to YG repeats.

FG phase arrangement in NPCs

Phase separation of a Nup98 FG domain in bulk buffer leads to a dense phase with spherical FG particles (Figs. 2-10), suggesting that surface tension minimizes the contact area between the hydrophobic FG phase and the surrounding water. These FG particles also feature a rather homogenous protein concentration indicating that the phase has reached thermodynamic equilibrium. We now measured this protein concentration by ODT and found it to be remarkably high – in the 450 mg/ ml range for a repeat domain with standard FG density ($GLFG_{52x12}$ or wild-type MacNup98A FG). This number has interesting implications for the organization of the permeability barrier in intact NPCs, which – according to the latest structures[8,9] – has to span a central channel with a diameter of 60 nm.

Nup98 occurs in 48 copies per NPC, has a 50 kDa FG domain, and contributes thus 2.5 Megadalton FG mass. If the Nup98 FG phase would contract to its preferred concentration of 400 mg/ml, then this mass would allow only a 3.5 nm thick disc to span the NPC's cross-section. Such a thin layer would allow for a very rapid NTR passage. However, it would be rather unstable because surface tension would force such a thin disc into a shape with a smaller surface area, ultimately leading to barrier collapse.

The Nup98 FG phase volume probably increases by Gle2/Rae1 binding the GLEBS domain, the presence of the Nup98 anchor domain, and partitioning NTRs that all occupy volume. Nevertheless, we assume additional stabilization mechanisms to be in place. First, single-molecule tracking has identified a sizeable obstacle in the very centre of the NPC channel, which NTR-cargo complexes circumnavigate[19]. This central object might function in 'spanning' a thin Nup98 FG phase layer across the channel (like the hub in a spoked wheel). This is pure speculation at the moment, but it can explain the presence of such an object and how a thin barrier layer could be stabilized. A further stabilization might arise from GLFG repeats binding to the NPC scaffold[82].

Finally, ten other FG Nups (Nups 358, 214, 153, 62, 58, 54, 50, 42, Pom121, and Ganp in humans) provide another 12 Megadalton in FG mass. This increases the FG phase volume and should reduce the shape-changing effect of surface tension. We do not expect a homogenous mixing between all these FG domains because their anchor points already dictate a spatial separation and because these domains come with a range of cohesivity. The most cohesive ones will co-phase-separate, forming the barrier's core. The less cohesive and

more water-soluble ones, however, should stay more peripheral, reduce (in a surfactant-analogous manner) the surface tension at the water-FG interface, and thereby stabilize the arrangement of a thin cohesive FG layer further. A nice illustration of this principle is soap allowing water to form thin films on flat plastic rings or as soap bubbles.

Methods

Nomenclature:

All sequence-regularized variants of the Nup98 FG domain in this study are named $XXXX_{N \times M}$, where XXXX is the sequence of a perfectly repeated tetrapeptide motif (e.g., GLFG or FSFG), *N* is the number of the repeat unit, and *M* is the number of amino acid residues per repeat unit (i.e., *M* = inter-tetrapeptide spacer length (*L*)+ 4). Full sequences of variants are shown in Supplementary Note 1.

DNA sequences of FG domain variants:

DNA sequences encoding the variants were generated with the assistance of the *Gene Designer* program, which minimized the repetition of local DNA sequences and optimized codon usage for *E. coli* expression. DNA fragments were synthesized by *GenScript* and cloned into a bacterial expression vector for overexpression and purification (see below). The *Tetrahymena thermophila* GLEBS domain (44 amino acid residues) of the Mac98A FG domain was included in the amino acid sequences (inserted between the 29th and 30th repeat unit) of all the variants because the presence of the corresponding non-repetitive DNA sequence improved the ease of DNA syntheses.

Recombinant protein expression and purification:

Mac98A FG protein domain and variants: The protein domains were recombinantly expressed in histidine-tagged form with an additional C-terminal cysteine each and purified as described[45]. The proteins were expressed in *E. coli* NEB Express at 30°C for 4 hr with induction by 0.4 mM IPTG. The Mac98A FG domain and most of the FG domain variants (except those specified below) phase-separated *in vivo* to form inclusion bodies. Cells were resuspended in cold 50 mM Tris/HCl pH 7.5, 300 mM NaCl, 1 mg/ml lysozyme, and lysed by a freeze-thaw cycle followed by mild sonication. Inclusion bodies were recovered by centrifugation (k-factor: 3158; 10 min) and washed once in 50 mM Tris/HCl pH 7.5, 300 mM NaCl, 5 mM DTT. The FG domain/variant was extracted with 40% formamide, 50 mM Tris/HCl pH 7.5, 10 mM DTT. The extract was cleared by ultracentrifugation (k-factor: 135; 90 min) and applied for 3 hr at room temperature to a Ni(II) chelate column. The column was washed in extraction buffer, followed by 200 mM ammonium acetate pH

7.5 (as a volatile buffer). The target protein was eluted with 30% acetonitrile, 265 mM formic acid, 10 mM ammonium formate, and lyophilized directly. The lyophilized powder was weighted for quantification. Unless specified otherwise, the lyophilized material was dissolved to 1 mM protein concentration in 4 M guanidinium hydrochloride (GuHCl).

GLFG//D_{52x12} and $GLLG_{52x12}$: Each was recombinantly expressed with an N-terminal His_{14}-ZZ-*sc*SUMO tag (which improved the expression) in *E. coli* NEB Express at 30 °C for 4 hr with induction by 0.4 mM IPTG. Cells were resuspended and lysed as described above. However, these target proteins remained in the soluble fraction of cell lysate and did not form inclusion bodies *in vivo*, indicating a lack of cohesiveness. The soluble fraction of cell lysate was cleared by ultracentrifugation and applied directly to a Ni(II) chelate column. The column was washed extensively in 50 mM Tris/HCl pH 7.5, 300 mM NaCl, 20 mM imidazole, 20 mM DTT, and then in protease buffer: 50 mM Tris/HCl pH 7.5, 300 mM NaCl, 5 mM DTT. 50 nM SUMO protease in protease buffer was applied for overnight on-column cleavage[83,84]. The cleaved targets eluted with the protease, remained soluble, were re-buffered to 30% acetonitrile on a PD10 Sephadex column (GE Healthcare), and finally lyophilized.

NTRs, EGFP, GFP variants, and mCherry: Most were expressed as His-tagged-fusions (Supp. Table 3) and purified by native Ni(II) chelate chromatography, as described previously[32,44]. Elution was performed by on-column protease cleavage[83,84].

Analysis of phase separation by centrifugation:

For each, 1 µl of a fresh stock of FG domain or variant (typically 1 mM protein in 4 M GuHCl; for $GLFG_{29x12}$ and $FSFG_{29x12}$: 2 mM protein in 4 M GuHCl) was rapidly diluted with assay buffer (50 mM Tris/HCl, 5 mM DTT, 150 mM NaCl, unless specified otherwise), to the concentration(s) stated in the figures, at 25°C. After incubation for 1 min, the FG phase (insoluble content) was pelleted by centrifugation (21130g, 30 mins, using a temperature-controlled Eppendorf 5424 R centrifuge equipped with a FA-45-24-11 rotor) at 25°C. Equivalent ratios of the pellet (condensed FG phase) and the supernatant were analysed by SDS-PAGE/ Coomassie blue-staining (the exact amount loaded for SDS-PAGE was adjusted individually such that the loaded amount of pellet + supernatant equalled 4.1 µg). Saturation concentration (a.k.a. threshold concentration or critical concentration) for phase separation of a given sample was taken as the concentration that remained in the supernatant, which was

estimated with a concentration series loaded onto the same gel. All tests were performed at least two times independently with similar results. Representative gel images are shown.

Optical diffraction tomography (ODT):

For each, 1 µl of a fresh stock of FG domain or variant (typically 1 mM protein in 4 M GuHCl; for $GLFG_{52x15}$ 1 mM protein in 2 M GuHCl), containing 0.5% molecules coupled with Atto488 via C-terminal cysteine, was rapidly diluted with assay buffer (500-fold for $GLFG_{52x13}$; 300-fold for $GLFG_{52x14}$ and GLLG//L_{52x12}; 100-fold for $GLFG_{52x15}$ and 600-fold for all others). See ref.[33] for the preparation of Atto488-coupled FG domains. The resulting mixture was placed on µ-slide 18-well (IBIDI, Germany), which had been passivated with 0.1 mg/ml Maltose-binding protein (MBP). FG particles were allowed to sediment under gravity for 15 mins and analysed with a custom-built correlative epifluorescence-ODT microscopy setup[60]. FG particles emitting Atto488-fluorescence in the well were visualised and located by epifluorescence microscopy coupled with the setup. The three-dimensional (3D) refractive index (RI) was measured using optical diffraction tomography (ODT) employing Mach-Zehnder interferometry similar to the one described[59,60,85,86] by a coherent laser beam with a wavelength of 532 nm. This interferometric technique measured spatially modulated holograms from 150 different angles, from which the complex optical fields were retrieved. By mapping the Fourier spectra of retrieved complex optical fields onto the surface of the Ewald sphere in the 3D Fourier space according to the Fourier diffraction theorem, 3D RI tomograms were reconstructed. Detailed principles for tomogram reconstruction can be found in refs.[58,87,88]. The image acquisition, field retrieval, and RI tomogram reconstruction were performed using custom-written MATLAB scripts (R2020a). Several reconstruction algorithms have also been developed for reconstructing RI tomograms of samples encountering multiple light scattering[89,90],which can extend the characterization of samples with heterogeneous density distribution or higher mass density if required. The mean RI value of each sample was measured by manually segmenting regions of interest (ROIs) from the central slice of the reconstructed tomogram using FIJI[91]. To calculate the mass density of each ROI, we used the following relationship[85]:

$$n = n_{medium} + \alpha C \qquad (2)$$

where n = measured RI in the ROI, n_{medium} = RI of the assay medium (1.336, measured with visible light by an Abbe refractometer: ABBE-2WAJ from Arcarda), α = RI increment (0.190 ml/g for proteins[92]) and C = mass density in the ROI.

Calculation of solvent contents:

$$Solvent\ content\ (\%) = (1 - C_{dense} \cdot 0.70\, ml/g) \times 100\% \quad (3)$$

Data fitting: Fittings in Fig. 2 were by the least squares method by Microsoft Excel 16.42. Mean values were fitted.

FG phase preparation for permeation assays:

For each, 1 µl of a fresh stock of FG domain or variant (1 mM unlabelled protein in 4 M GuHCl, unless specified elsewhere) was rapidly diluted with 50 µl assay buffer, and 7.5 µl of the suspension was mixed with 22.5 µl substrate. As described[45], the substrate contains either 12 µM Hoechst 33342/ 6 µM mCherry/ 3 µM EGFP/ 3 µM efGFP_8Q/ 3 µM efGFP_8R/ 1 µM of an NTR or [1.5 µM NTR pre-incubated with 1 µM cargo], in assay buffer. Note: before mixing NTF2 with RanGDP-Atto488, the Ran protein was incubated with 2 µM RanGAP for 30 mins, which converts all Ran to the GDP-bound form in the assay. The resulting mixture was placed on collagen-coated µ-slides 18-well (IBIDI, Germany). FG particles were allowed to sediment under gravity for 1 hr before imaging by CLSM. The samples for ODT and CLSM were prepared independently.

Confocal laser scanning microscopy (CLSM) and quantification of partition coefficients:

Hoechst 33342, mCherry, and GFP/Alexa488 signals were acquired (as described[44,45]) with 405, 561, and 488 nm excitation, respectively, with a Leica SP5 confocal scanning microscope equipped with a 63x oil immersion objective and hybrid detectors (standard mode, in which non-linear response of the detector was auto-corrected) at 21°C. Scanning settings (e.g., laser power) were adjusted individually such that the quantification was within the dynamic range. Typically, the powers of the 405, 561, and 488 nm lasers were set to 1-5%, 20%, and 2-8%, respectively. For optimal signal-to-noise ratios, line-averaging was applied in all cases. Frame accumulation was additionally applied for recording weak signals, e.g., when the probe was excluded by the FG phase.

To determine the Hoechst fluorescence signals, raw signals in the centres of 3-5 (in focus) FG particles were quantified using Leica Application Suite X 3.3.0. Standard deviations were typically about 10-15% between individual particles. For validation, the mean signal of a given sample was compared to that of $GLFG_{52x12}$, which was recorded under the same settings. To determine the partition coefficients, raw signals in the centres of 3-5 (in focus) FG particles (IN) and three reference areas within the aqueous region (OUT) were quantified using the same software. If necessary, XZ scans were acquired to determine the signals along the equators of FG particles. Partition coefficients (Part. Coeff.) for each particle were calculated according to: Part. Coeff.= IN / Mean (OUT). As reported previously[44], standard deviations were typically <10% between individual particles and between experiments. Mean values are shown in the figures. All measurements were performed at least two times independently with similar results. Representative images are shown in the figures.

NPC-staining tests of GFP^{NTR} variants:

These were performed two times independently as described[44] - with similar results. Fig. 8b shows representative images.

Code Availability

Custom MATLAB scripts for ODT reconstruction are publicly available at:
https://github.com/OpticalDiffractionTomography

References

1. Hoelz, A., Glavy, J. S. & Beck, M. Toward the atomic structure of the nuclear pore complex: when top down meets bottom up. *Nat Struct Mol Biol* **23**, 624-630 (2016).
2. Hampoelz, B., Andres-Pons, A., Kastritis, P. & Beck, M. Structure and assembly of the nuclear pore complex. *Annual review of biophysics* **48**, 515-536 (2019).
3. Lin, D. H. & Hoelz, A. The Structure of the Nuclear Pore Complex (An Update). *Annu Rev Biochem* **88**, 725-783 (2019).
4. Bley, C. J. et al. Architecture of the cytoplasmic face of the nuclear pore. *Science* **376**, eabm9129 (2022).
5. Petrovic, S. et al. Architecture of the linker-scaffold in the nuclear pore. *Science* **376**, eabm9798 (2022).
6. Schwartz, T. U. Solving the nuclear pore puzzle. *Science* **376**, 1158-1159 (2022).

7. Zhu, X. et al. Structure of the cytoplasmic ring of the Xenopus laevis nuclear pore complex. *Science* **376**, eabl8280 (2022).
8. Zimmerli, C. E. et al. Nuclear pores dilate and constrict in cellulo. *Science* **374**, eabd9776 (2021).
9. Schuller, A. P. et al. The cellular environment shapes the nuclear pore complex architecture. *Nature* **598**, 667-671 (2021).
10. Bonner, W. M. Protein migration into nuclei. I. Frog oocyte nuclei in vivo accumulate microinjected histones, allow entry to small proteins, and exclude large proteins. *J Cell Biol* **64**, 421-430 (1975).
11. Mohr, D., Frey, S., Fischer, T., Güttler, T. & Görlich, D. Characterisation of the passive permeability barrier of nuclear pore complexes. *EMBO J* **28**, 2541-2553 (2009).
12. Görlich, D. & Kutay, U. Transport between the cell nucleus and the cytoplasm. *Annu Rev Cell Dev Biol* **15**, 607-660 (1999).
13. Ribbeck, K., Lipowsky, G., Kent, H. M., Stewart, M. & Görlich, D. NTF2 mediates nuclear import of Ran. *EMBO J* **17**, 6587-6598 (1998).
14. Ribbeck, K. & Görlich, D. Kinetic analysis of translocation through nuclear pore complexes. *EMBO J* **20**, 1320-1330 (2001).
15. Yang, W., Gelles, J. & Musser, S. M. Imaging of single-molecule translocation through nuclear pore complexes. *Proc Natl Acad Sci U S A* **101**, 12887-12892 (2004).
16. Kubitscheck, U. et al. Nuclear transport of single molecules: dwell times at the nuclear pore complex. *J Cell Biol* **168**, 233-243 (2005).
17. Delavoie, F., Soldan, V., Rinaldi, D., Dauxois, J.-Y. & Gleizes, P.-E. The path of pre-ribosomes through the nuclear pore complex revealed by electron tomography. *Nature Communications* **10**, (2019).
18. Ruland, J. A. et al. Nuclear export of the pre-60S ribosomal subunit through single nuclear pores observed in real time. *Nat Commun* **12**, 6211 (2021).
19. Chowdhury, R., Sau, A. & Musser, S. M. Super-resolved 3D tracking of cargo transport through nuclear pore complexes. *Nat Cell Biol* **24**, 112-122 (2022).
20. Hurt, E. C. A novel nucleoskeletal-like protein located at the nuclear periphery is required for the life cycle of Saccharomyces cerevisiae. *EMBO J* **7**, 4323-34. (1988).
21. Davis, L. I. & Fink, G. R. The NUP1 gene encodes an essential component of the yeast nuclear pore complex. *Cell* **61**, 965-978 (1990).
22. Wente, S. R., Rout, M. P. & Blobel, G. A new family of yeast nuclear pore complex proteins. *J Cell Biol* **119**, 705-723 (1992).
23. Lemke, E. A. The Multiple Faces of Disordered Nucleoporins. *J Mol Biol* **428**, 2011-2024 (2016).
24. Schmidt, H. B. & Görlich, D. Transport Selectivity of Nuclear Pores, Phase Separation, and Membraneless Organelles. *Trends Biochem Sci* **41**, 46-61 (2016).
25. Iovine, M. K. The GLFG repetitive region of the nucleoporin Nup116p interacts with Kap95p, an essential yeast nuclear import factor. *The Journal of Cell Biology* **131**, 1699-1713 (1995).
26. Bayliss, R. et al. Interaction between NTF2 and xFxFG-containing nucleoporins is required to mediate nuclear import of RanGDP. *J Mol Biol* **293**, 579-593 (1999).
27. Frey, S., Richter, R. P. & Görlich, D. FG-rich repeats of nuclear pore proteins form a three-dimensional meshwork with hydrogel-like properties. *Science* **314**, 815-817 (2006).
28. Frey, S. & Görlich, D. A saturated FG-repeat hydrogel can reproduce the permeability properties of nuclear pore complexes. *Cell* **130**, 512-523 (2007).
29. Frey, S. & Görlich, D. FG/FxFG as well as GLFG repeats form a selective permeability barrier with self-healing properties. *EMBO J* **28**, 2554-2567 (2009).

30. Milles, S. & Lemke, E. A. Single molecule study of the intrinsically disordered FG-repeat nucleoporin 153. *Biophys J* **101**, 1710-1719 (2011).
31. Hülsmann, B. B., Labokha, A. A. & Görlich, D. The permeability of reconstituted nuclear pores provides direct evidence for the selective phase model. *Cell* **150**, 738-751 (2012).
32. Schmidt, H. B. & Görlich, D. Nup98 FG domains from diverse species spontaneously phase-separate into particles with nuclear pore-like permselectivity. *Elife* **4**, e04251 (2015).
33. Najbauer, E. E., Ng, S. C., Griesinger, C., Görlich, D. & Andreas, L. B. Atomic resolution dynamics of cohesive interactions in phase-separated Nup98 FG domains. *Nat Commun* **13**, 1494 (2022).
34. Hyman, A. A., Weber, C. A. & Jülicher, F. Liquid-liquid phase separation in biology. *Annu Rev Cell Dev Biol* **30**, 39-58 (2014).
35. Banani, S. F., Lee, H. O., Hyman, A. A. & Rosen, M. K. Biomolecular condensates: organizers of cellular biochemistry. *Nat Rev Mol Cell Biol* **18**, 285-298 (2017).
36. Shin, Y. & Brangwynne, C. P. Liquid phase condensation in cell physiology and disease. *Science* **357**, eaaf4382 (2017).
37. Hondele, M. et al. DEAD-box ATPases are global regulators of phase-separated organelles. *Nature* **573**, 144-148 (2019).
38. Martin, E. W. & Holehouse, A. S. Intrinsically disordered protein regions and phase separation: sequence determinants of assembly or lack thereof. *Emerging Topics in Life Sciences* **4**, 307-329 (2020).
39. Martin, E. W. et al. Valence and patterning of aromatic residues determine the phase behavior of prion-like domains. *Science* **367**, 694-699 (2020).
40. Musacchio, A. On the role of phase separation in the biogenesis of membraneless compartments. *EMBO J* **41**, e109952 (2022).
41. Schmidt, H. B., Barreau, A. & Rohatgi, R. Phase separation-deficient TDP43 remains functional in splicing. *Nat Commun* **10**, 4890 (2019).
42. Choi, J. M., Holehouse, A. S. & Pappu, R. V. Physical Principles Underlying the Complex Biology of Intracellular Phase Transitions. *Annu Rev Biophys* **49**, 107-133 (2020).
43. Holehouse, A. S., Ginell, G. M., Griffith, D. & Böke, E. Clustering of Aromatic Residues in Prion-like Domains Can Tune the Formation, State, and Organization of Biomolecular Condensates. *Biochemistry* **60**, 3566-3581 (2021).
44. Frey, S. et al. Surface properties determining passage rates of proteins through nuclear pores. *Cell* **174**, 202-217.e9 (2018).
45. Ng, S. C., Güttler, T. & Görlich, D. Recapitulation of selective nuclear import and export with a perfectly repeated 12mer GLFG peptide. *Nat Commun* **12**, 4047 (2021).
46. Popken, P., Ghavami, A., Onck, P. R., Poolman, B. & Veenhoff, L. M. Size-dependent leak of soluble and membrane proteins through the yeast nuclear pore complex. *Mol Biol Cell* **26**, 1386-1394 (2015).
47. Ori, A. et al. Cell type-specific nuclear pores: a case in point for context-dependent stoichiometry of molecular machines. *Mol Syst Biol* **9**, 648 (2013).
48. Powers, M. A., Forbes, D. J., Dahlberg, J. E. & Lund, E. The vertebrate GLFG nucleoporin, Nup98, is an essential component of multiple RNA export pathways. *J Cell Biol* **136**, 241-250 (1997).
49. Labokha, A. A. et al. Systematic analysis of barrier-forming FG hydrogels from Xenopus nuclear pore complexes. *EMBO J* **32**, 204-218 (2013).

50. Strawn, L. A., Shen, T., Shulga, N., Goldfarb, D. S. & Wente, S. R. Minimal nuclear pore complexes define FG repeat domains essential for transport. *Nat Cell Biol* **6**, 197-206 (2004).
51. Lord, C. L., Timney, B. L., Rout, M. P. & Wente, S. R. Altering nuclear pore complex function impacts longevity and mitochondrial function in S. cerevisiae. *J Cell Biol* **208**, 729-744 (2015).
52. Harmon, T. S., Holehouse, A. S., Rosen, M. K. & Pappu, R. V. Intrinsically disordered linkers determine the interplay between phase separation and gelation in multivalent proteins. *Elife* **6**, (2017).
53. Borcherds, W., Bremer, A., Borgia, M. B. & Mittag, T. How do intrinsically disordered protein regions encode a driving force for liquid-liquid phase separation. *Curr Opin Struct Biol* **67**, 41-50 (2021).
54. Yamada, J. et al. A bimodal distribution of two distinct categories of intrinsically disordered structures with separate functions in FG nucleoporins. *Mol Cell Proteomics* **9**, 2205-2224 (2010).
55. Dekker, M., van der Giessen, E. & Onck, P. R. Liquid-liquid phase separation of intrinsically disordered FG-Nups is driven by highly-dynamic hydrophobic FG-motifs. (2022). doi: https://doi.org/10.1101/2022.09.20.508740
56. Weisblum, B. & Haenssler, E. Fluorometric properties of the bibenzimidazole derivative Hoechst 33258, a fluorescent probe specific for AT concentration in chromosomal DNA. *Chromosoma* **46**, 255-260 (1974).
57. Yarmoluk, S. M., Kovalska, V. B. & Volkova, K. D. in *Advanced Fluorescence Reporters in Chemistry and Biology III* (ed Demchenko, A.) 161-199 (Springer, 2011).
58. Sung, Y. et al. Optical diffraction tomography for high resolution live cell imaging. *Opt Express* **17**, 266-277 (2009).
59. Kim, K. & Guck, J. The relative densities of cytoplasm and nuclear compartments are robust against strong perturbation. *Biophys J* **119**, 1946-1957 (2020).
60. Biswas, A., Kim, K., Cojoc, G., Guck, J. & Reber, S. The Xenopus spindle is as dense as the surrounding cytoplasm. *Dev Cell* **56**, 967-975.e5 (2021).
61. Ng, S. C. & Görlich, D. A simple thermodynamic description of phase separation of Nup98 FG domains. *Nat Commun* **13**, 6172 (2022).
62. Chan, H. S. & Dill, K. A. The effects of internal constraints on the configurations of chain molecules. *The Journal of Chemical Physics* **92**, 3118-3135 (1990).
63. Song, J., Ng, S. C., Tompa, P., Lee, K. A. & Chan, H. S. Polycation-π interactions are a driving force for molecular recognition by an intrinsically disordered oncoprotein family. *PLoS Comput Biol* **9**, e1003239 (2013).
64. Pollard, V. W. et al. A novel receptor-mediated nuclear protein import pathway. *Cell* **86**, 985-994 (1996).
65. Görlich, D., Henklein, P., Laskey, R. & Hartmann, E. A 41 amino acid motif in importin alpha confers binding to importin beta and hence transit into the nucleus. *EMBO J.* **15**, 1810-1817 (1996).
66. Wen, W., Meinkotht, J. L., Tsien, R. Y. & Taylor, S. S. Identification of a signal for rapid export of proteins from the nucleus. *Cell* **82**, 463-473 (1995).
67. Patel, S. S., Belmont, B. J., Sante, J. M. & Rexach, M. F. Natively unfolded nucleoporins gate protein diffusion across the nuclear pore complex. *Cell* **129**, 83-96 (2007).
68. Ader, C. et al. Amyloid-like interactions within nucleoporin FG hydrogels. *Proc Natl Acad Sci USA* **107**, 6281-6285 (2010).
69. Sinnokrot, M. O. & Sherrill, C. D. Substituent effects in pi-pi interactions: sandwich and T-shaped configurations. *J Am Chem Soc* **126**, 7690-7697 (2004).

70. Ringer, A. L., Figgs, M. S., Sinnokrot, M. O. & Sherrill, C. D. Aliphatic C-H/pi interactions: Methane-benzene, methane-phenol, and methane-indole complexes. *J Phys Chem A* **110**, 10822-10828 (2006).
71. Lin, Y., Currie, S. L. & Rosen, M. K. Intrinsically disordered sequences enable modulation of protein phase separation through distributed tyrosine motifs. *J Biol Chem* **292**, 19110-19120 (2017).
72. Bremer, A. et al. Deciphering how naturally occurring sequence features impact the phase behaviours of disordered prion-like domains. *Nat Chem* **14**, 196-207 (2022).
73. Bayliss, R., Littlewood, T. & Stewart, M. Structural Basis for the Interaction between FxFG Nucleoporin Repeats and Importin-β in Nuclear Trafficking. *Cell* **102**, 99-108 (2000).
74. Bayliss, R. et al. Structural basis for the interaction between NTF2 and nucleoporin FxFG repeats. *EMBO J* **21**, 2843-2853 (2002).
75. Isgro, T. A. & Schulten, K. Binding dynamics of isolated nucleoporin repeat regions to importin-β. *Structure* **13**, 1869-1879 (2005).
76. Port, S. A. et al. Structural and Functional Characterization of CRM1-Nup214 Interactions Reveals Multiple FG-Binding Sites Involved in Nuclear Export. *Cell Rep* **13**, 690-702 (2015).
77. Crowley, P. B. & Golovin, A. Cation-pi interactions in protein-protein interfaces. *Proteins* **59**, 231-239 (2005).
78. Nott, T. et al. Phase transition of a disordered nuage protein generates environmentally responsive membraneless organelles. *Molecular Cell* **57**, 936-947 (2015).
79. Elvin, C. M. et al. Synthesis and properties of crosslinked recombinant pro-resilin. *Nature* **437**, 999-1002 (2005).
80. DePace, A. H., Santoso, A., Hillner, P. & Weissman, J. S. A critical role for amino-terminal glutamine/asparagine repeats in the formation and propagation of a yeast prion. *Cell* **93**, 1241-1252 (1998).
81. Serio, T. R. et al. Nucleated conformational conversion and the replication of conformational information by a prion determinant. *Science* **289**, 1317-1321 (2000).
82. Andersen, K. R. et al. Scaffold nucleoporins Nup188 and Nup192 share structural and functional properties with nuclear transport receptors. *Elife* **2**, e00745 (2013).
83. Frey, S. & Görlich, D. A new set of highly efficient, tag-cleaving proteases for purifying recombinant proteins. *J Chromatogr A* **1337**, 95-105 (2014).
84. Frey, S. & Görlich, D. Purification of protein complexes of defined subunit stoichiometry using a set of orthogonal, tag-cleaving proteases. *J Chromatogr A* **1337**, 106-115 (2014).
85. Barer, R. Determination of dry mass, thickness, solid and water concentration in living cells. *Nature* **172**, 1097-1098 (1953).
86. Park, Y., Depeursinge, C. & Popescu, G. Quantitative phase imaging in biomedicine. *Nature Photonics* **12**, 578-589 (2018).
87. Wolf, E. Three-dimensional structure determination of semi-transparent objects from holographic data. *Optics Communications* **1**, 153-156 (1969).
88. Kim, K. et al. High-resolution three-dimensional imaging of red blood cells parasitized by Plasmodium falciparum and in situ hemozoin crystals using optical diffraction tomography. *J Biomed Opt* **19**, 011005 (2014).
89. Lim, J., Ayoub, A. B., Antoine, E. E. & Psaltis, D. High-fidelity optical diffraction tomography of multiple scattering samples. *Light Sci Appl* **8**, 82 (2019).
90. Lee, M., Hugonnet, H. & Park, Y. Inverse problem solver for multiple light scattering using modified Born series. *Optica* **9**, 177 (2022).

91. Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**, 676-682 (2012).
92. Zhao, H., Brown, P. H. & Schuck, P. On the distribution of protein refractive index increments. *Biophys J* **100**, 2309-2317 (2011).

Acknowledgments

S.C.N. and D.G. thank Waltraud Taxer and Susanne Brandfass for technical help, Prof. Tom A. Rapoport for critical reading, as well as the Max-Planck-Gesellschaft and the Deutsche Forschungsgemeinschaft (SFB 860 and SFB 1190 to D.G.) for funding. A.B. and S.R. thank Dr. Kyoohyun Kim and Prof. Jochen Guck for their support. A.B. was also supported by an Add-On Fellowship from the Joachim-Herz-Stiftung.

Author Contributions

S.C.N. planned and conducted most of the experiments. A.B. conducted ODT measurements and data analysis. T.H. designed the TetraGFP^{NTR} variant. J.S. conducted preliminary experiments with YG phases. D.G. conceived the overall concepts of the study and wrote the manuscript. S.C.N., S.R., and D.G. contributed to experimental design, data analysis, and interpretation.

Competing Interests Statement

The authors declare no competing interests.

Figure legends

Fig. 1 Inter-GLFG spacer length determines the phase separation propensity.

(a & b) Sequences (partial) of the wild-type *Tetrahymena thermophila* MacNup98A ("Mac98A") FG domain **(a)** and a sequence-regularized variant, $GLFG_{52x12}$ **(b)**. The latter served as a template for variant design in this study. $GLFG_{52x12}$ is composed of 52 repeat units, each containing one GLFG motif and one eight-amino-acid spacer (i.e. length of each repeat unit=12 AA). For space economy, only the N-terminal ≈130 residues of each are shown (see Supplementary Note 1 for complete sequences).

(c) Based on $GLFG_{52x12}$, a variant with one residue in each inter-GLFG spacer replaced by Asp was constructed ($GLFG//D_{52x12}$). The sequence is listed with that of Mac98A and $GLFG_{52x12}$. For each, the N-terminal sequence up to the fourth FG motif is shown. For $GLFG_{52x12}$ and $GLFG//D_{52x12}$, the C-terminal sequences follow the same design strategy as the N-terminal sequences shown (see Supplementary Note 1 for complete sequences). Each FG domain or variant was expressed, purified, and dissolved at a concentration of 1 mM in 4 M guanidinium hydrochloride. As a test of phase separation property, this stock of each was diluted 100-fold quickly with assay buffer (50 mM Tris/HCl pH 7.5, 150 mM NaCl, 5 mM DTT) to allow phase separation. The dilution was centrifuged. SDS samples of the obtained pellets (FG phase), if there were, and supernatants (soluble content) were loaded for SDS-PAGE at an equal ratio (6%), followed by Coomassie blue-staining for quantification. Saturation concentration, C_{sat}, for phase separation of each was taken as the concentration of the supernatant. For $GLFG//D_{52x12}$, no phase separation was detected at the assay protein concentration (= 10 μM), and the exact C_{sat} was not determined.

(d) Based on $GLFG_{52x12}$, variants with the same number (= 52) of GLFG motifs but varying inter-GLFG spacer lengths were designed. N-terminal sequences up to the fourth FG motif are shown (see Supplementary Note 1 for complete sequences). For each variant, a 1 mM stock in 4 M guanidinium hydrochloride was prepared, and phase separation was analysed as described above. **(c & d)** Each of the assays was performed twice on independent samples with similar results, and representative images are shown. The same applies to Figs. 4a and 9d.

Fig. 2 Variation of inter-GLFG spacer length impacts intra-FG phase density.

(a) Mac98A FG domain and variants with varying spacer length, as indicated (the former has irregular spacer length with an average of about 8 AA), were dissolved at a concentration of 1 mM (except that $GLFG_{52x15}$ was dissolved at a concentration of 2 mM, because of its lower saturation concentration of 7.5 μM in the assay buffer) in 4 M guanidinium hydrochloride, and phase separation was initiated by a rapid 50-fold dilution with assay buffer. Each of the dilutions was further diluted fourfold in 12 μM

Hoechst 33342. Samples were analysed by confocal laser-scanning microscopy (CLSM). The numbers in orange indicate the fluorescence intensities of the Hoechst dye inside the FG phases. The fluorescence intensities were relative to that of $GLFG_{52x12}$ (arbitrarily set to 100). The above assay was performed twice on independent samples with similar results, and representative images are shown. The same applies to Fig. 10b.

(b) Phase separation of Mac98A FG domain and its variants was initiated similarly as in **(a)** and the samples were analysed by optical diffraction tomography (ODT) independently (in the absence of Hoechst). Panels show maps of refractive index (RI). For each FG domain/ variant, ten independent FG particles were analysed. Representative images and mean values (±S.D.) of RI, mass density (C_{dense}), FG motif concentration and solvent % are shown. *Note: Mac98A FG domain has <52 FG motifs but also contains FG-like motifs/ hydrophobic residues in spacers (Supp. Table 2), which are not counted here.

(c) C_{sat} shown in Fig. 1d is plotted against the spacer length (L) of the variants. Mean values of C_{sat} (n=2) are plotted and fitted to a simple exponential function (dashed line) with the R-squared value indicated.

(d) C_{dense} obtained from ODT is plotted against L. For each variant, ten independent FG particles were analysed, and the data are presented as mean ± S.D. (in molar concentration).

(e) Gibbs free energy for phase separation (ΔG) of each variant was calculated by the equation in red (i.e. Eq.1 in the main text; with T = 298 K) and is plotted against L. The data are presented as mean values. Note: C_{sat} of $GLFG_{52x10}$ was extrapolated from the equation derived in **(c)** (=0.008 μM), and the corresponding ΔG was computed accordingly.

Fig. 3 Variation of inter-GLFG spacer length impacts the transport selectivity of barrier.

FG phases assembled from the Mac98A FG domain and indicated variants were prepared as described in Fig. 2. They were challenged with the indicated permeation probes (for probe descriptions, see main text). In each case, the calculated molecular weight of the complex is indicated. Scanning settings/ image brightness were adjusted individually to cover the large range of signals. The numbers in white refer to the partition coefficients of the fluorescent species into the FG phases (fluorescence ratios in the central regions of the particles to that in the surrounding buffer). N.D.: Not determined due to difficulties in defining the central region. Each of the above assays was performed twice on independent samples with similar results, and representative images/mean values are shown. The same applies to Figs. 5-10.

Fig. 4 Phase separation property is dependent on the overall hydrophobicity.

(a) $GLFG_{52x12}$ serves as a template for variant design in this experiment. Note that the inter-GLFG spacers of $GLFG_{52x12}$ do not contain Phe and Leu. All the variants contain 52 repeat units, and each unit is 12-amino-acid long. The N-terminal sequence of each variant up to the fourth FG motif is shown; the rest of the sequence (~620 residues/ 48 repeat units) follows the same design strategy as shown (see Supplementary Note 1 for complete sequences). Phase separation of the variants was analysed as in Fig. 1c at [Variant] = 10 μM. No phase separation was observed for $GAFG_{52x12}$, $GLLG_{52x12}$, and $GAYG_{52x12}$ under these conditions, and the exact saturation concentrations were not determined.

(b) Phase separation of $GLFG_{52x12}$ and its variants was initiated and the resulting dense phases were analysed by optical diffraction tomography (ODT). Panels show maps of refractive index (RI). For each, ten independent FG particles were analysed. Representative images and mean values (±S.D.) of RI, mass density (C_{dense}) and motif concentration are shown.

(c) C_{dense} for each FG phase variant determined by ODT. For each, ten independent particles were analysed, and the data are presented as mean ± S.D. (in molar concentration).

Fig. 5 FG domain variants with non-canonical Fx motifs phase-separate into barriers that allow entry of NTR-cargo complexes.

FG or FG-like phases assembled from the indicated variants were challenged with the indicated probes. Scanning settings/ image brightness were adjusted individually due to the large range of signals. The numbers in white refer to the partition coefficients of the fluorescent species into the phases.

Fig. 6 Spatial arrangement of Phe and Leu in FG domain is crucial for NTR-binding but not for engineered GFP^{NTR}-variants.

FG or FG-like phases assembled from the indicated variants were challenged with the indicated probes. Scanning settings/ image brightness were adjusted individually due to the large range of signals. The numbers in white refer to the partition coefficients of the fluorescent species into the phases.

Fig. 7 Arginines promote FG phase entry not just through cation-π interactions.

Indicated condensed phases were challenged with sffrGFP4, which is an engineered FG-philic GFP variant, and "sffrGFP4 25R→K", in which all the surface FG-philic arginine residues were replaced

by FG-phobic lysines. Note that the GLLG//L phase, which lacks aromatic residues, still allows the partition of the Arg-rich sffrGFP4, indicating interactions other than cation-π promoted the entry.

Fig. 8 An engineered GFP variant that binds specifically to GLFG motifs and NPCs.

(a) A rational mutation (M225F) was introduced to GFP^{NTR}_3B7C, and the resulting variant was named $TetraGFP^{NTR}$. Indicated FG or FG-like phases were challenged with GFP^{NTR}_3B7C and $TetraGFP^{NTR}$. Note that except for the canonical GLFG phase, the mutation reduces the partition coefficients into all other phases.

(b) Panels show confocal scans of digitonin-permeabilized HeLa cells that had been incubated for 5 min with 250 nM of either GFP^{NTR}_3B7C or $TetraGFP^{NTR}$. Live images were taken without fixation or washing steps, i.e., they show directly an NPC-binding in relation to the free concentration and non-specific aggregation with nuclear or cytoplasmic structures. While 3B7C still showed some weak binding to nuclear structures other than NPCs, $TetraGFP^{NTR}$ stains NPCs more crisply, illustrating that $TetraGFP^{NTR}$ is a more specific NPC binder. This experiment was performed twice on independent samples with similar results, and representative images are shown.

(c) Plot profiles corresponding to lines across nuclei marked in **(b)**. Fluorescence (GFP) intensities are normalized to the maximum (arbitrarily set to 100) of each line.

Fig. 9 Exchanging GLFG to FSFG motifs results in a hypercohesive FG phase that precludes entry of NTR-cargo complexes.

(a) Based on a framework of $GLFG_{52x12}$, $FSFG_{52x12}$ was constructed as described above. The FG phase assembled from $FSFG_{52x12}$ was challenged with the indicated probes. Scanning settings/ image brightness were adjusted individually due to the large range of signals. N.D.: Not determined due to difficulties in defining the central region. Note that all probes only arrested at the periphery of the phase.

(b) Based on $FSFG_{52x12}$, two mutants with a reduced number (= 35 or 36) of FSFG motifs were constructed. “Mutant 1” corresponds to a sequence where every third FSFG motif in $FSFG_{52x12}$ was mutated to SSSG (FG-phobic). “Mutant 2” is similar, but every fifth and sixth FSFG motif were mutated. For space economy, only the N-terminal ≈144 residues of each are shown. The C-terminal sequences follow the same design strategy as the region shown (see Supplementary Note 1 for complete sequences). Corresponding regions of $GLFG_{52x12}$ and $FSFG_{52x12}$ are shown for comparison.

(c) FG phases assembled from the above variants were challenged with the indicated probes as in **(a)**.

(d) $GLFG_{29x12}$ and $FSFG_{29x12}$: corresponding to the N-terminal 29 repeats of $GLFG_{52x12}$ and $FSFG_{52x12}$, respectively, were tested for phase separation at a concentration of 20 μM. Note that the FSFG motifs lead to stronger phase separation propensity than the GLFG motifs (see also Supp. Fig. 2).

Fig. 10 Hypercohesion of FSFG motifs can be balanced by longer inter-FG spacers.

(a) Based on $FSFG_{52x12}$, two variants with the same number (52) of FSFG motifs but longer inter-FSFG spacers were constructed ($FSFG_{52x15}$ and $FSFG_{52x18}$).

(b) Upper panels: condensed phases assembled from the indicated variants were stained with Hoechst 33342. The numbers in orange indicate the fluorescence intensities of the Hoechst dye inside the FG phases. The fluorescence intensities are relative to that of $GLFG_{52x12}$ (arbitrarily set to 100). Lower panels: the condensed phases were analysed by optical diffraction tomography (ODT) independently in the absence of Hoechst. Panels show maps of refractive index (RI). For each, ten independent FG particles were analysed. Representative images and mean values (±S.D.) of RI, mass density (C_{dense}), and FG motif concentration are shown.

(c) Condensed phases assembled from $FSFG_{52x12}$ and its two variants with longer spacers were challenged with the indicated probes. Note that lengthening the spacers relaxes the permeation selectivity.

Figure 1

a ***Tt*MacNup98A (Mac98A) FG domain**

Irregular repeats

```
-MFGNTGGG
GLFGNTQTQQTGG
GLFGQPQQ
TQFGQTGATGG
GLFGGAT
NTFGGGGGG
GLFGGNNNQQTNPTAGG
GIFGQGTTGLGGAPAQTGG
GLFGAPQNNQGG
GLFGGGTTTGG
GMFGNQANTQTGGG
.....(537 AA not shown)
```

b **GLFG$_{52x12}$**

Each repeat : 12 AA

Spacer: 8 AA

```
GLFGNTGGAPAG
GLFGNTQTQQGG
GLFGQPQQTQGG
GLFGQTGATTGG
GLFGGATNTAPG
GLFGGGGGNPTG
GLFGGNNNQQTG
GLFGQGTTQTGG
GLFGAPQNNQGG
GLFGGGTTTTGG
GLFGANTQTGGG
.....(536 AA not shown)
```

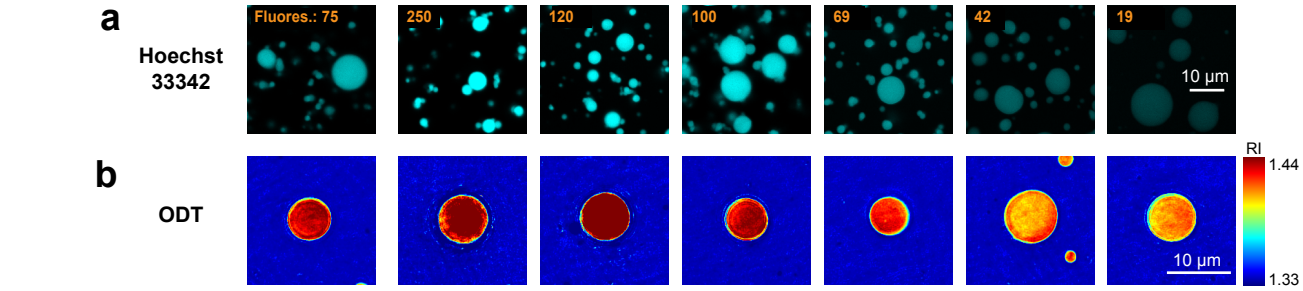
c

		Phase separation test with 10 μM protein: Pellet / Supern.	Saturation conc. (C_{sat})
Mac98A	MFG -NTGGG- GLFG -NTQTQQTGG- GLFG -QPQQTQ- FGQTGAT...		0.3 μM
GLFG$_{52x12}$	GLFG -NTGGAPAG- GLFG -NTQTQQGG- GLFG -QPQQTQGG- GLFG...		0.1 μM
GLFG//D$_{52x12}$	GLFG -NTGDAPAG- GLFG -NTQDQQGG- GLFG -QPQDTQGG- GLFG...		>10 μM

d

		Variation of inter-GLFG spacer length	Pellet / Supern.	C_{sat}
Shorter spacers	**GLFG$_{52x10}$**	GLFG-NTGGAP-GLFG-TQTQQG-GLFG-QPQQQG-GLFG...		≤0.04 μM
	GLFG$_{52x11}$	GLFG-NTGGAPG-GLFG-TQTQQGG-GLFG-QPQQQGG-GLFG...		0.04 μM
standard	**GLFG$_{52x12}$**	GLFG-NTGGAPAG-GLFG-NTQTQQGG-GLFG-QPQQTQGG-GLFG...		0.1 μM
Longer spacers	**GLFG$_{52x13}$**	GLFG-NTGGTAPAG-GLFG-NTQTQQQGG-GLFG-QPQQSTQGG-GLFG...		0.5 μM
	GLFG$_{52x14}$	GLFG-NTGGTGAPAG-GLFG-NTQTQNQQGG-GLFG-QPQQSPTQGG-GLFG...		2.5 μM
	GLFG$_{52x15}$	GLFG-NTGGTGAAPAG-GLFG-NTQTQNTQQGG-GLFG-QPQQSPNTQGG-GLFG...		7.5 μM

Figure 2



	Mac98A	$GLFG_{52x10}$	$GLFG_{52x11}$	$GLFG_{52x12}$	$GLFG_{52x13}$	$GLFG_{52x14}$	$GLFG_{52x15}$
Spacer length (AA):	≈ 8	6	7	8 (standard)	9	10	11
RI	1.426 ± 0.006	1.451 ± 0.012	1.449 ± 0.008	1.426 ± 0.004	1.417 ± 0.007	1.397 ± 0.011	1.396 ± 0.006
C_{dense} (mg/ml)	**474** ± 32	**603** ± 64	**595** ± 40	**475** ± 22	**426** ± 39	**323** ± 58	**314** ± 33
FG motif conc. (mM)	**305** ± 26*	**548** ± 58	**510** ± 34	**379** ± 18	**317** ± 29	**226** ± 41	**207** ± 22
Solvent %	66.8 ± 2.2	57.8 ± 4.5	58.4 ± 2.8	66.8 ± 1.6	70.2 ± 2.7	77.4 ± 4.1	78.0 ± 2.3

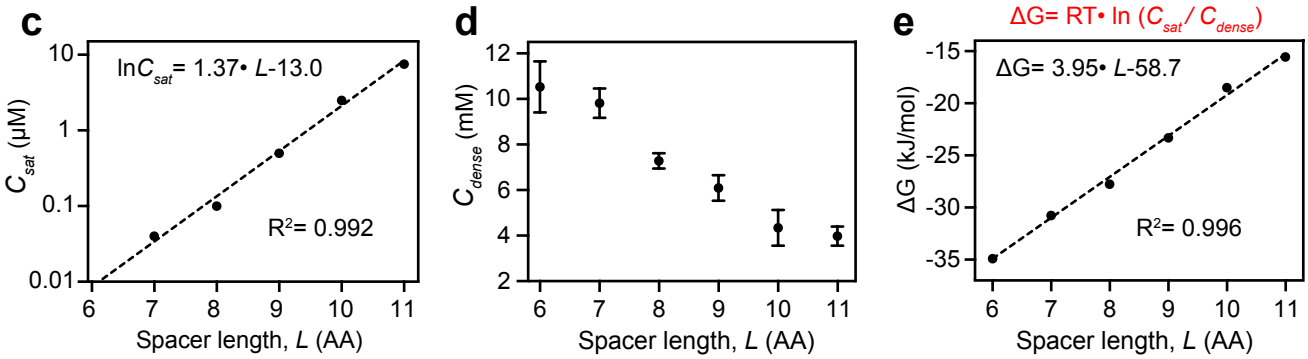


Figure 3

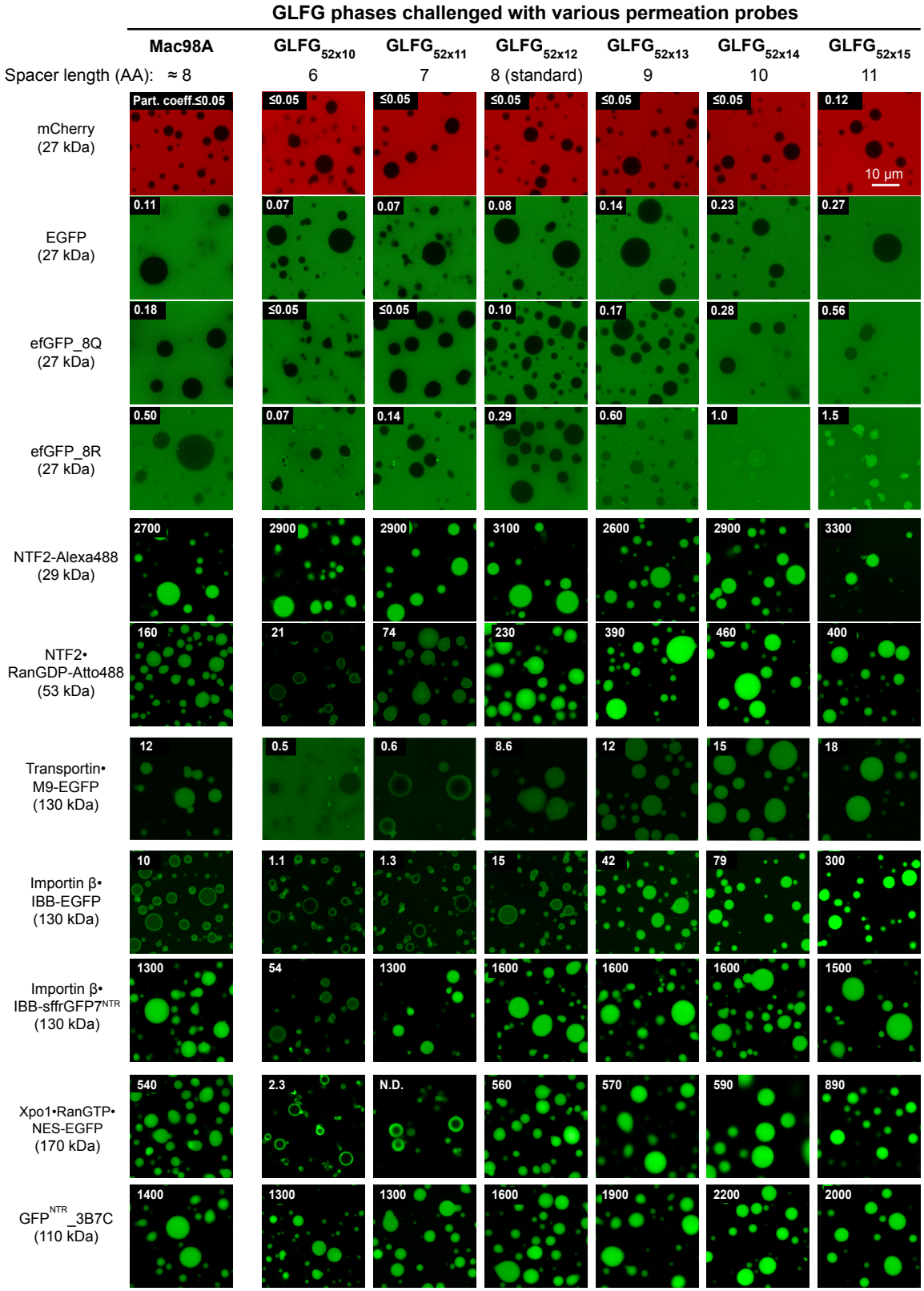
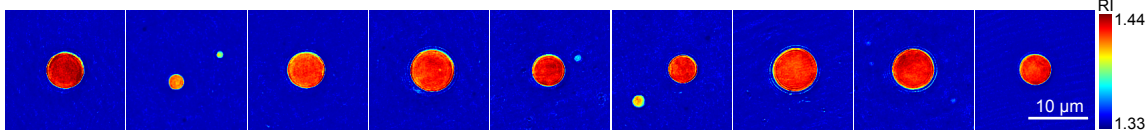

GLFG phases challenged with various permeation probes
Mac98A
GLFG52x10
GLFG52x11
GLFG52x12
GLFG52x13
GLFG52x14
GLFG52x15
Spacer length (AA): ≈ 8
6
7
8 (standard)
9
10
11
mCherry (27 kDa)
Part. coeff.≤0.05
≤0.05
≤0.05
≤0.05
≤0.05
≤0.05
0.12
10 µm
EGFP (27 kDa)
0.11
0.07
0.07
0.08
0.14
0.23
0.27
efGFP_8Q (27 kDa)
0.18
≤0.05
≤0.05
0.10
0.17
0.28
0.56
efGFP_8R (27 kDa)
0.50
0.07
0.14
0.29
0.60
1.0
1.5
NTF2-Alexa488 (29 kDa)
2700
2900
2900
3100
2600
2900
3300
NTF2• RanGDP-Atto488 (53 kDa)
160
21
74
230
390
460
400
Transportin• M9-EGFP (130 kDa)
12
0.5
0.6
8.6
12
15
18
Importin β• IBB-EGFP (130 kDa)
10
1.1
1.3
15
42
79
300
Importin β• IBB-sffrGFP7NTR (130 kDa)
1300
54
1300
1600
1600
1600
1500
Xpo1•RanGTP• NES-EGFP (170 kDa)
540
2.3
N.D.
560
570
590
890
GFPNTR_3B7C (110 kDa)
1400
1300
1300
1600
1900
2200
2000

Figure 4

a

Variations of FG motifs on a framework of 52x12mer repeats		Phase separation test (C_{sat})
$GLFG_{52x12}$	GLFG -NTGGAPAG- GLFG -NTQTQQGG- GLFG -QPQQTQGG-GLFG ...	0.1 μM
$GAFG_{52x12}$	GAFG -NTGGAPAG- GAFG -NTQTQQGG- GAFG -QPQQTQGG-GAFG ...	>10 μM
$GLLG_{52x12}$	GLLG -NTGGAPAG- GLLG -NTQTQQGG- GLLG -QPQQTQGG-GLLG ...	>10 μM
$GLLG//L_{52x12}$	GLLG -NTGLAPAG- GLLG -NTQLQQGG- GLLG -QPQLTQGG-GLLG ...	0.8 μM
$GXFG//L_{52x12}$	GGFG -NTGLAPAG- GTFG -NTQLQQGG- GQFG -QPQLTQGG-GAFG ...	0.8 μM
$GFLG_{52x12}$	GFLG -NTGGAPAG- GFLG -NTQTQQGG- GFLG -QPQQTQGG- GFLG ...	0.3 μM
$GLYG_{52x12}$	GLYG -NTGGAPAG- GLYG -NTQTQQGG- GLYG -QPQQTQGG-GLYG ...	≤0.04 μM
$GAYG_{52x12}$	GAYG -NTGGAPAG- GAYG -NTQTQQGG- GAYG -QPQQTQGG-GAYG ...	>10 μM
$GIFG_{52x12}$	GIFG -NTGGAPAG- GIFG -NTQTQQGG- GIFG -QPQQTQGG- GIFG ...	0.1 μM
$GLFA_{52x12}$	GLFA -NTGGAPAG- GLFA -NTQTQQGG- GLFA -QPQQTQGG- GLFA ...	0.2 μM
$SLFG_{52x12}$	SLFG -NTGGAPAG- SLFG -NTQTQQGG- SLFG -QPQQTQGG- SLFG ...	0.2 μM
$GLFS_{52x12}$	GLFS -NTGGAPAG- GLFS -NTQTQQGG- GLFS -QPQQTQGG- GLFS ...	0.5 μM

b



	GLFG	GLLG//L	GXFG//L	GFLG	GLYG	GIFG	GLFA	SLFG	GLFS
RI	1.426 ±0.004	1.404 ±0.006	1.414 ± 0.009	1.420 ± 0.005	1.425 ± 0.003	1.421 ± 0.005	1.427 ± 0.007	1.423 ± 0.004	1.425 ± 0.004
C_{dense} (mg/ml)	**475** ± 22	**360** ± 30	**413** ± 47	**440** ± 29	**466** ± 14	**448** ± 27	**478** ± 34	**457** ± 20	**467** ± 22
Motif conc. (mM)	**379** ± 18	**292** ± 24	**329** ± 38	**351** ± 23	**367** ± 11	**358** ± 21	**377** ± 27	**356** ± 16	**364** ± 17

c

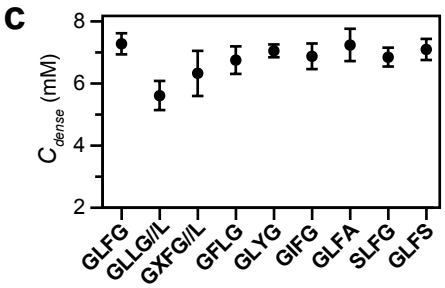


Figure 5

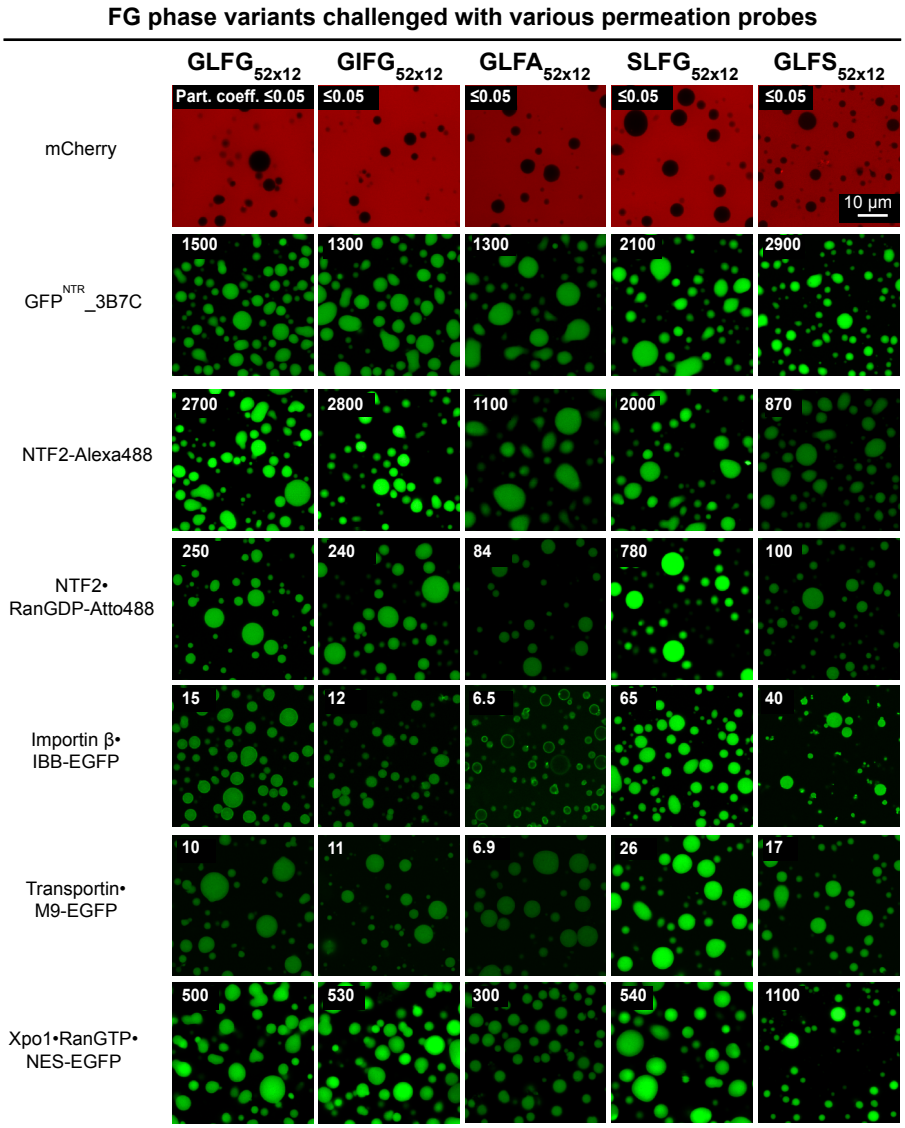

FG phase variants challenged with various permeation probes
GLFG52x12
GIFG52x12
GLFA52x12
SLFG52x12
GLFS52x12
mCherry
Part. coeff. ≤0.05
≤0.05
≤0.05
≤0.05
≤0.05
10 µm
GFPNTR_3B7C
1500
1300
1300
2100
2900
NTF2-Alexa488
2700
2800
1100
2000
870
NTF2•
RanGDP-Atto488
250
240
84
780
100
Importin β•
IBB-EGFP
15
12
6.5
65
40
Transportin•
M9-EGFP
10
11
6.9
26
17
Xpo1•RanGTP•
NES-EGFP
500
530
300
540
1100

Figure 6

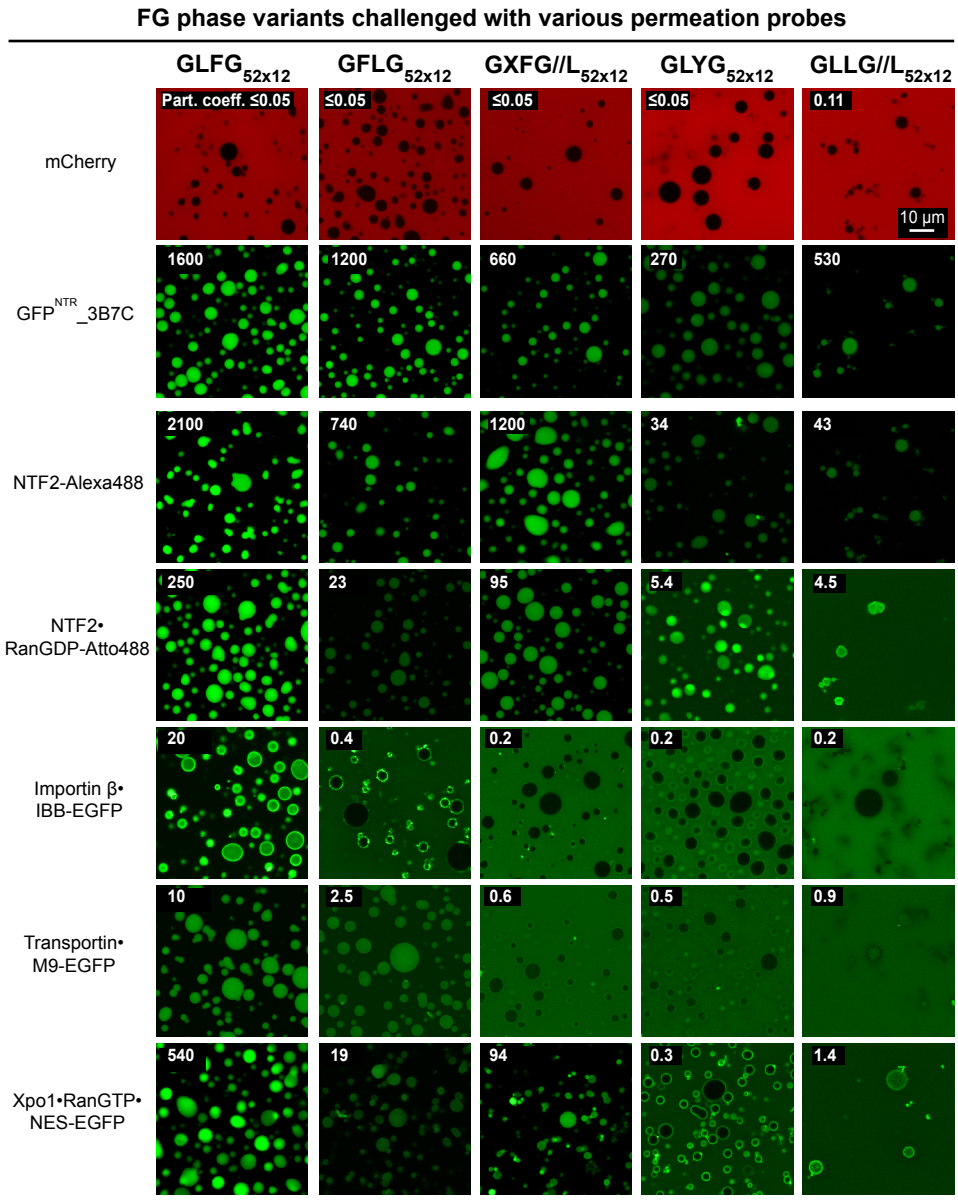

FG phase variants challenged with various permeation probes
GLFG52x12
GFLG52x12
GXFG//L52x12
GLYG52x12
GLLG//L52x12
mCherry
Part. coeff. ≤0.05
≤0.05
≤0.05
≤0.05
0.11
10 µm
GFPNTR_3B7C
1600
1200
660
270
530
NTF2-Alexa488
2100
740
1200
34
43
NTF2•
RanGDP-Atto488
250
23
95
5.4
4.5
Importin β•
IBB-EGFP
20
0.4
0.2
0.2
0.2
Transportin•
M9-EGFP
10
2.5
0.6
0.5
0.9
Xpo1•RanGTP•
NES-EGFP
540
19
94
0.3
1.4

Figure 7

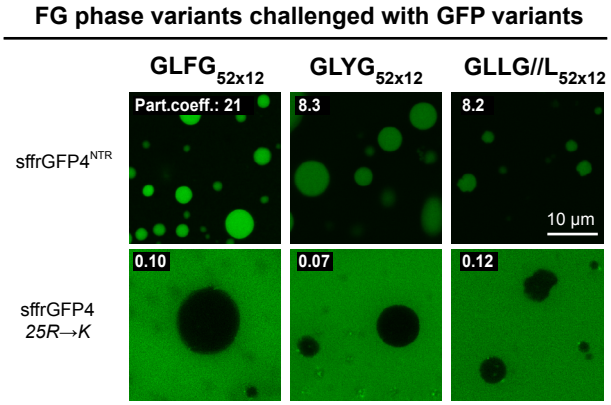

FG phase variants challenged with GFP variants
GLFG52x12
GLYG52x12
GLLG//L52x12
sffrGFP4NTR
sffrGFP4
25R→K
Part.coeff.: 21
8.3
8.2
10 µm
0.10
0.07
0.12

Figure 8

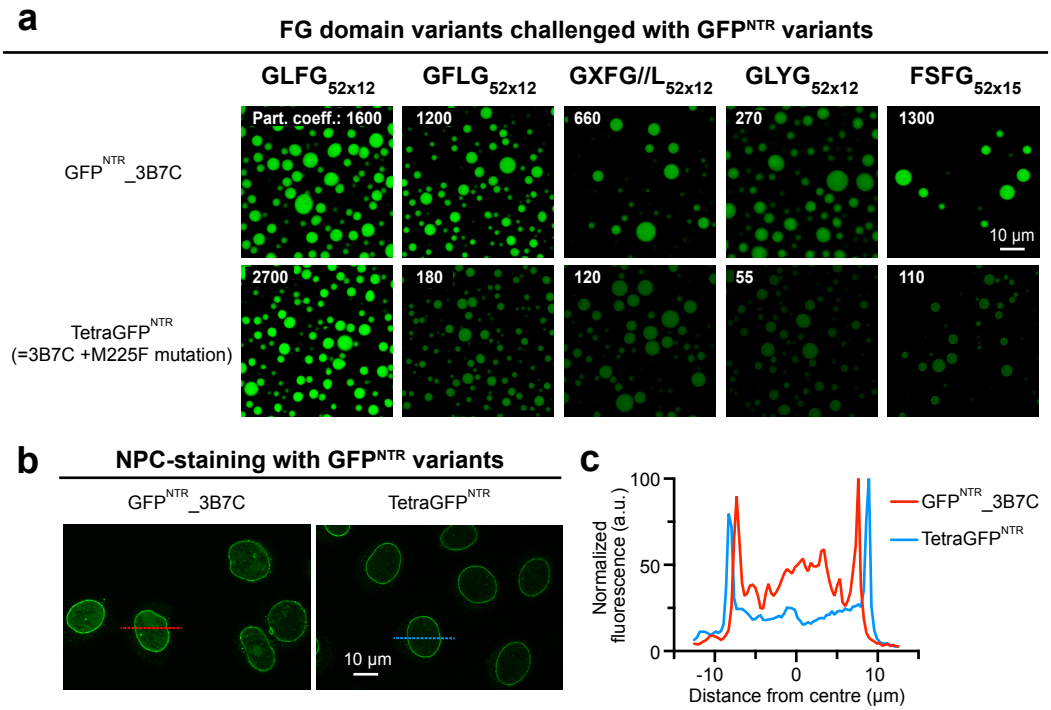
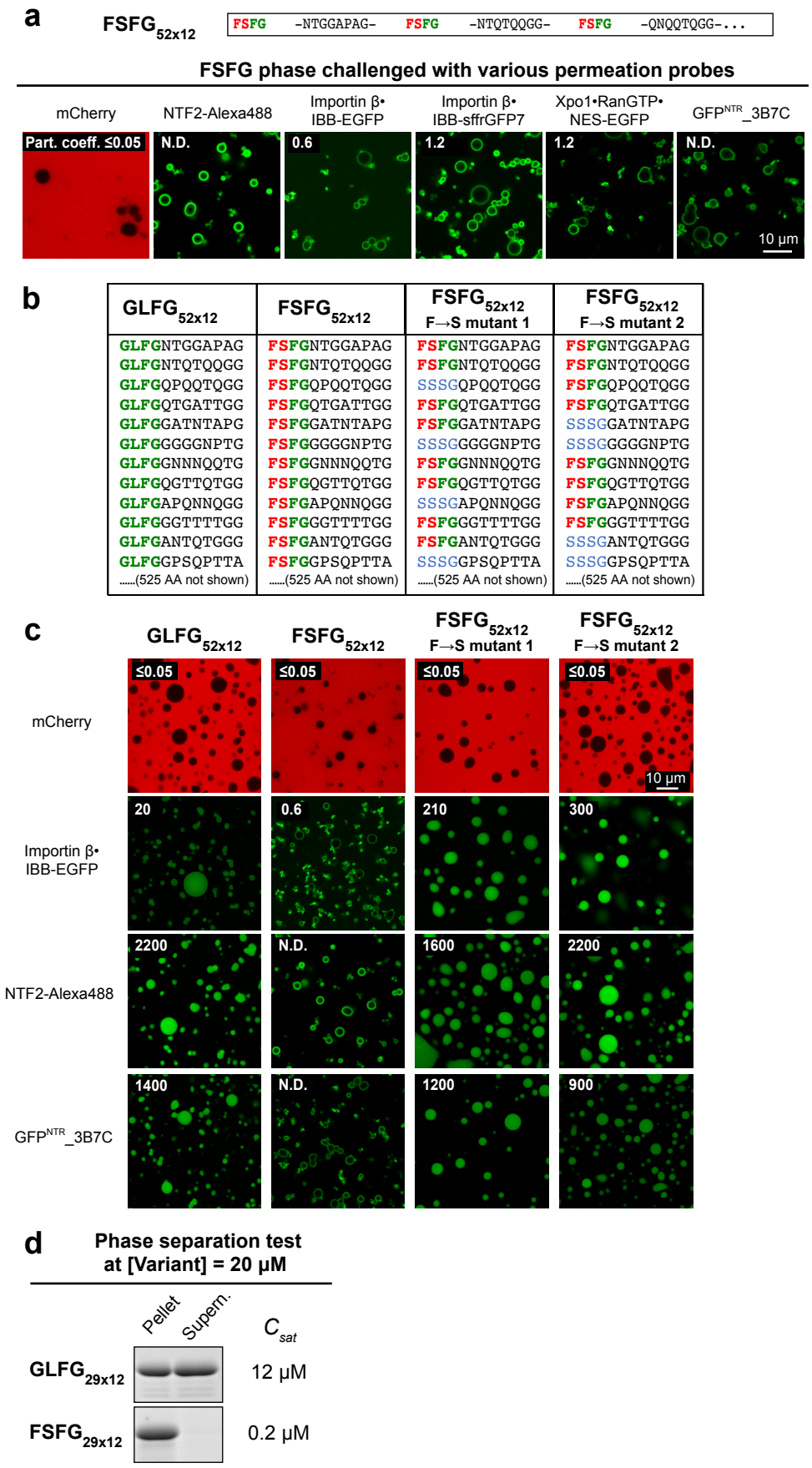

a
FG domain variants challenged with GFP^NTR variants
GLFG52x12
GFLG52x12
GXFG//L52x12
GLYG52x12
FSFG52x15
GFP^NTR_3B7C
TetraGFP^NTR
(=3B7C +M225F mutation)
Part. coeff.: 1600
1200
660
270
1300
10 µm
2700
180
120
55
110
b
NPC-staining with GFP^NTR variants
GFP^NTR_3B7C
TetraGFP^NTR
10 µm
c
Normalized fluorescence (a.u.)
100
50
0
-10
0
10
Distance from centre (µm)
GFP^NTR_3B7C
TetraGFP^NTR

Figure 9



a
FSFG52x12
FSFG -NTGGAPAG- FSFG -NTQTQQGG- FSFG -QNQQTQGG-...
FSFG phase challenged with various permeation probes
mCherry
NTF2-Alexa488
Importin β• IBB-EGFP
Importin β• IBB-sffrGFP7
Xpo1•RanGTP• NES-EGFP
GFPNTR_3B7C
Part. coeff. ≤0.05
N.D.
0.6
1.2
1.2
N.D.
10 µm
b
GLFG52x12
FSFG52x12
FSFG52x12 F→S mutant 1
FSFG52x12 F→S mutant 2
GLFGNTGGAPAG
GLFGNTQTQQGG
GLFGQPQQTQGG
GLFGQTGATTGG
GLFGGATNTAPG
GLFGGGGGNPTG
GLFGGNNNQQTG
GLFGQGTTQTGG
GLFGAPQNNQGG
GLFGGGTTTTGG
GLFGANTQTGGG
GLFGGPSQPTTA
.....(525 AA not shown)
FSFGNTGGAPAG
FSFGNTQTQQGG
FSFGQPQQTQGG
FSFGQTGATTGG
FSFGGATNTAPG
FSFGGGGGNPTG
FSFGGNNNQQTG
FSFGQGTTQTGG
FSFGAPQNNQGG
FSFGGGTTTTGG
FSFGANTQTGGG
FSFGGPSQPTTA
.....(525 AA not shown)
FSFGNTGGAPAG
FSFGNTQTQQGG
SSSGQPQQTQGG
FSFGQTGATTGG
FSFGGATNTAPG
SSSGGGGGNPTG
FSFGGNNNQQTG
FSFGQGTTQTGG
SSSGAPQNNQGG
FSFGGGTTTTGG
FSFGANTQTGGG
SSSGGPSQPTTA
.....(525 AA not shown)
FSFGNTGGAPAG
FSFGNTQTQQGG
FSFGQPQQTQGG
FSFGQTGATTGG
SSSGGATNTAPG
SSSGGGGGNPTG
FSFGGNNNQQTG
FSFGQGTTQTGG
FSFGAPQNNQGG
FSFGGGTTTTGG
SSSGANTQTGGG
SSSGGPSQPTTA
.....(525 AA not shown)
c
GLFG52x12
FSFG52x12
FSFG52x12 F→S mutant 1
FSFG52x12 F→S mutant 2
mCherry
Importin β• IBB-EGFP
NTF2-Alexa488
GFPNTR_3B7C
≤0.05
≤0.05
≤0.05
≤0.05
10 µm
20
0.6
210
300
2200
N.D.
1600
2200
1400
N.D.
1200
900
d
Phase separation test at [Variant] = 20 µM
Pellet
Supern.
Csat
GLFG29x12
12 µM
FSFG29x12
0.2 µM

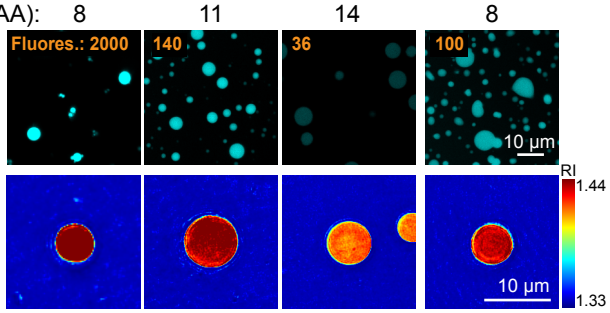
Figure 10

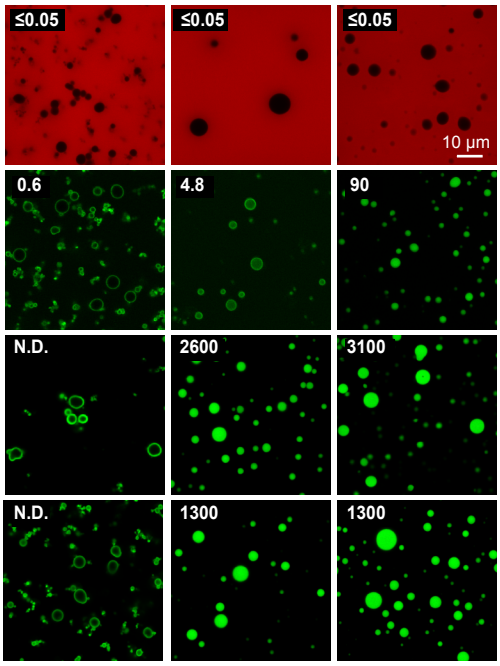
a

$FSFG_{52x12}$	FSFG -NTGGAPAG- FSFG -NTQTQQGG- FSFG -QNQQTQGG-...
$FSFG_{52x15}$	FSFG -NTGGTGAAPAG- FSFG -NTQTQNTQQGG- FSFG -QNQQSPNTQGG-...
$FSFG_{52x18}$	FSFG -NTGGTGAAGTAPAG- FSFG -NTQTQNTNPQQQGG- FSFG -QNQQSPNTTSTQGG-...

b

	$FSFG_{52x12}$	$FSFG_{52x15}$	$FSFG_{52x18}$	$GLFG_{52x12}$
Spacer length (AA):	8	11	14	8
Hoechst 33342	Fluores.: 2000	140	36	100
ODT				
RI	1.459 ± 0.008	1.436 ± 0.009	1.414 ± 0.004	1.426 ± 0.004
C_{dense} (mg/ml)	**649** ± 42	**529** ± 48	**412** ± 22	**475** ± 22
Motif conc.(mM)	**493** ± 32	**335** ± 30	**222** ± 12	**379** ± 18



c

FSFG phases challenged with various permeation probes