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 – consistent with a sieve-like phase. Phase separation tolerated mutations of Nup98-typical GLFG motifs, provided the domain-hydrophobicity remained preserved. NTR-entry, however, was sensitive to (certain) deviations from canonical FG motifs, suggesting a co-evolutionary adaptation. Unexpectedly, we found that arginines promote efficient FG-phase entry also by means other than cation-π interactions. Although incompatible with NTR·cargo complexes, a YG phase displayed remarkable transport selectivity, particularly for evolved GFP^{NTR}-variants. GLFG to FSFG mutations made the FG phase hypercohesive, precluding NTR-entry. Longer spacers relieved this hypercohesive phenotype. The antagonism between cohesion and NTR·FG interactions appears thus key to transport selectivity.

Introduction

The nuclear envelope (NE) separates the nucleus from the cytoplasm, confining an intercompartmental 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 in intact cells 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'. 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 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 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 intrinsically disordered[23]. FG domains are anchored to the inner scaffold of the NPCs[2,3] and bind NTRs during facilitated translocation[24,25]. Moreover, they can engage in multivalent, cohesive interactions to form a 'selective FG phase'[14,26–32]. Such FG phases 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[27,28,31,33,34]. 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. A binding of NTRs to FG motifs competes with FG-FG interactions and allows NTRs to 'melt' (along with their cargoes) through the meshes.

A state of matter similar to the FG phase was later described also as biomolecular condensates or liquid-liquid phase separation (LLPS)[36–42], e.g., in germline granules. At least some FG repeats bind to the NPC scaffold[35]. Such adhesive interactions probably seal the barrier towards the inner walls of the NPC.

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[33,43]. 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) speeding up. Based on these principles, GFP has been evolved to pass NPCs either very slowly (super-inert GFPs) or very rapidly (GFP^{NTR} variants), with the extreme versions (of identical size) differing more than 10 000-fold in their NPC passage rates[33].

High transport selectivity of an FG phase requires the local FG repeat concentration to exceed a threshold of ~200 mg/ml[27]. Such high local concentration can only partially be imposed by anchoring the ~15 MDa of FG domain mass to the NPC scaffold. The other key contribution is localized phase separation, in particular, of Nup98 FG domains[31].

FG domains from Nup98 or its homologs[22,31,44] (e.g., Nup116, Nup100 and Nup145 in yeast) are the most cohesive[31,45] and crucial ones for maintaining the permeability barrier of

NPCs[30,43,46,47]. 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[31]. Such self-assembled Nup98 FG phases recapitulate the transport selectivity of NPCs particularly well[31,33,34]. 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. Recently, we reported that RanGTPase-controlled cargo import and export scenarios can be recapitulated faithfully with a single Nup98 FG phase[34]. 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 evolved 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 FG-phobic cargoes. 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 and Discussion

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[31,34]. This setup was used to determine saturation concentrations (C_{sat}) by measuring the fractions that remained soluble after pelleting the formed phases. In separated 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[45]. For conceptual clarity, however, we used a derivative that had been simplified[34], 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 GLEBS domain 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.[34]), suggesting that $GLFG_{52x12}$ captures the barrier properties of the original domain indeed very well.

A negative net charge antagonizes FG phase separation

Introducing one negative charge (as an aspartate) per repeat unit, however, had a very drastic effect and abolished phase separation of the corresponding variant, $GLFG//D_{52x12}$ (Fig. 1c). This can be rationalized by electrostatic repulsion and thus an increase in water solubility of the spacers and also the entire domain[48,49]. It is consistent with the extreme selection against Asp and Glu in native Nup98 FG domains and other cohesive FG domains[50,51], and against a negative net charge in FG domains in general.

Shorter inter FG spacers enhance cohesive interactions

In a next step, we varied the length of the inter GLFG-spacers while keeping their amino acid composition and 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[52,53] but also stains FG phases in a way that depends on the presence of cohesive interactions[34].

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.[54–56]). 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[31], but this difference might be explained by some fluorescence quenching within such a dense phase. The protein concentration of 470 mg/ml (380 mM FG motifs) 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. The concentration ratio is also linearly related to the free energy change of phase separation[57]:

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

where R is the gas constant (8.31 $J/mol \cdot K$) 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[58,59] and/or enhance the local cooperativity of such cohesive contacts[48,49]. 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 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, ref.[33]) as probes, whose surfaces ranged from very FG-phobic to weakly FG-philic (mCherry, EGFP, efGFP_8Q, efGFP_8R). 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, namely to 0.12 (for mCherry) or even to 1.5 (for efGFP_8R) within the $GLFG_{52x15}$ phase. 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 the latter 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 NTF2 immerses into the phase while Ran still faces the bulk solvent. Similar "rim-staining" patterns were observed for the transportin·M9-EGFP, Xpo1·RanGTP·NES-EGFP, and importin β·IBB-EGFP complexes.

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[33]. Remarkably, partitioning of the importin β·IBB-sffrGFP7 complex into the $GLFG_{52x10}$ phase (and the other FG phases) was even stronger than that of importin β·IBB-Atto488, where the GFP moiety has been removed completely[31], indicating that increased surface FG-philicity is sufficient to overcome the size factor. Likewise, the tetrameric (110 kDa) GFP^{NTR}-3B7C, which was also engineered for FG-philicity and rapid NPC passage[33], 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 (Supp. Fig. 1; refs.[57,60]). Increasing the salt concentration had an 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 (Fig. 4). A reduction in the motif hydrophobicity essentially abolished phase assembly[26,30,50,60,61]. 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 at similar saturation concentrations as the standard sequence. Thus, phase separation tolerates deviations from the GLFG motif sequence as long as the overall hydrophobicity is not reduced. A lowered

hydrophobicity, however, impedes phase assembly. 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.[45,50,51,61,62]).

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 accumulation of 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 what hydrophobic motif confers the cohesive interaction. The evolved GFP^{NTR} 3B7C variant appears here to be a particularly promiscuous client of these phases.

Why then is the NPC-barrier not formed by, e.g., YG repeats? YG repeat proteins do exist in nature, with extracellular resilins being the best studied examples[63]. These elastomers store mechanical energy and allow insects e.g., 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), however, would lead to an irreversible clogging of the barrier. Thus, the propensity of the phenolic tyrosine side chains to covalent crosslinking could explain why tyrosines had been so strongly selected against within Nup FG repeat domains. Other reasons might be an evolutionary bias (see below), that tyrosines are more promiscuous than phenylalanines in their interactions[33], and that YG-rich proteins are prone to form amyloid-like structures. The latter is exemplified by the yeast prion Sup35[64,65].

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 had also 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 Nup98 FG domain from *C. elegans* (Supp. Table 1) and that the slightly weaker cohesion of this motif (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 surprisingly little effects: 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[66–69] being probably more constrained than purely superficial sites.

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

Lysine and arginine have long aliphatic/ hydrophobic sidechains with terminal positive charges; they engage in salt bridges and are highly soluble in water. However, despite these

similarities, the two have profoundly different effects on NPC passage: 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 large factors[33].

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) down to a partition coefficient of 0.1. The arginine effect can be explained by the positively charged guanidinium group engaging in cation-π interactions with the π electron cloud of phenylalanines[70]. The lysine amine group can engage in similar interactions, however, the planar guanidinium is favoured here by better shape complementarity to the planar phenylalanine ring and by being more easily released from the hydrogen-bonding network of water. The multivalency of arginines amplifies the effect to the ~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 (partition coefficient of 8) than the 25R→K variant (of 0.1). This points to arginine interactions that are distinct from cation-π. We suspect that these are of two kinds. First, arginine might engage more easily also in general hydrophobic interactions (e.g., with the three leucines per repeat unit) – again because its planar guanidium group interacts less favourably with the hydrogen bonding network of water than the tetragonal ammonium group of lysine. 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.

TetraGFP^{NTR}: an engineered FG phase marker

We previously described the GFP^{NTR} 3B7C variant, which was evolved 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[33]. 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 signal with less background than its progenitor 3B7C (Fig. 8b, c). TetraGFP^{NTR} also showed preference of GLFG over FSFG phase (see 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[62].

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$_{52x12}$ recombinant 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 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 had 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. In fact, 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[57], and even when reducing the number of $FSFG_{x12}$ repeats from 52 to 29 ($FSFG_{29x12}$, Fig. 9d), C_{sat} was just above the detection limit with 0.2 μM.

In the next step, we asked if a transport-selective FSFG phase can also be generated with regular repeats. For this, we extended the inter-FSFG spacer length of the $FSFG_{52x12}$ domain initially from 8 to 11 residues (Fig. 10). The resulting $FSFG_{52x15}$ phase allowed NTF2 and 3B7C to accumulate to partition coefficients of 2600 and 1300 respectively, while still being rather non-permissive towards the IBB-EGFP·importin β complex. Extending the spacer length to 14 residues ($FSFG_{52x18}$) then allowed also the importin β·cargo 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 through entropic effects. In any case, this set of experiments illustrates nicely the antagonism between inter FG-cohesion and FG·NTR interactions and suggest that selective NPC passage indeed happens by NTRs transiently disengaging cohesive contacts within the FG phase.

Methods

Nomenclature:

All sequence-regularized variants of 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 Supp. 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 encoding the variants in a codon-optimized form 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 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 as a histidine-tagged form in bacteria and purified as described[34]. The proteins were expressed in *E. coli* NEB Express at 30°C for 4 h and induction by 0.4 mM IPTG. The expressed 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. For each, inclusion bodies containing the FG domain/variant 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 3 hr at room temperature

to a Ni(II) chelate column. The column was washed in extraction buffer, with 200 mM ammonium acetate pH 7.5. The target protein was eluted with 30% acetonitrile, 265 mM formic acid, 10 mM ammonium formate, and directly lyophilized. The lyophilized powder was weighted for quantification. Unless specified elsewhere, 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 as a fusion form containing a His$_{14}$-ZZ-*sc*SUMO group, which improved the expression, in *E. coli* NEB Express at 30 °C for 4 h and induction by 0.4 mM IPTG. Cells were resuspended and lysed as described above. However, the expressed target protein 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 cleared by ultracentrifugation was 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 bdSENP1 in protease buffer was applied for overnight on-column cleavage. The cleaved target protein was soluble and eluted with the protease buffer and then was re-buffered to 30% acetonitrile by a PD10 Sephadex column (GE Healthcare), and 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[31,33]. Elution was performed by on-column protease cleavage[71,72].

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 100-fold with assay buffer (50 mM Tris/HCl, 5 mM DTT, 150 mM NaCl, unless specified), to the concentration 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 ratio of the pellet (condensed FG phase) and supernatant was 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 = 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 independently at least two times with similar results, and the representative gel images are shown. Full scans of gels with molecular weight markers are provided in the Source Data file.

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) was rapidly diluted with assay buffer (600-fold for Mac98A, $GLFG_{52x10}$, $GLFG_{52x11}$, $GLFG_{52x12}$ and $FSFG_{52x15}$, 500-fold for $GLFG_{52x13}$; 300-fold for $GLFG_{52x14}$ and 100-fold for $GLFG_{52x15}$). The resulting mixture was placed on µ-slide 18-well (IBIDI, Germany) which had been passivated with 0.1 mg/ml MBP. FG particles were allowed to sediment under gravity for 15 mins. The three-dimensional (3D) refractive index (RI) of FG particles was measured using a custom-built optical diffraction tomography (ODT) microscope employing Mach-Zehnder interferometry similar to the one described[55,56]. 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.[54,73,74]. The image acquisition, field retrieval, and RI tomogram reconstruction were performed using custom-written MATLAB scripts (R2020a). 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[75]. To calculate the mass density of each ROI, we used the following relationship:

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

where n = measured RI in the ROI, n_{medium} = RI of the assay medium (1.336, measured with Abbe refractometer: ABBE-2WAJ from Arcarda), α = RI increment (0.190 ml/g for proteins[76]) 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. Means values were fitted.

FG phase preparation for permeation assays:

For each, 1 µl of a fresh stock of FG domain or variant (1 mM 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[34], 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 maintained Ran in 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.

Confocal laser scanning microscopy and quantification of partition coefficients:

Hoechst 33342, mCherry and GFP/Alexa488 signals were acquired (as described[33,34]) 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. To determine the partition coefficients, raw signals in the centres of 3-5 FG particles (IN), which are in-focus, and 3 reference areas within the aqueous region (OUT) were quantified using Leica Application Suite X 3.3.0. If necessary, XZ scans were acquired to determine the signals in the centres of FG particles. Partition coefficients (Part. Coeff.) for each particle were calculated according to: Part. Coeff.= IN / Mean (OUT). As reported previously[33] standard deviations were typically <10% between individual particles and between experiments. Mean values were shown in the figures. All measurements were performed at least two times independently with similar results, and the representative images are shown in the figures.

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

These were performed as described[33] and were performed two times independently with similar results, and the representative images are shown in Fig. 8b.

Data Availability

A source data file is provided with this article. The data that support the findings of this study are provided in the Supplementary Information or Source Data file.

Code Availability

Custom MATLAB scripts for ODT reconstruction is 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., Grünwald, D., Hoekstra, A., Rohleder, D., Kues, T., Siebrasse, J. P. & Peters, R. 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., Krüger, A. M., Dörner, K., Bhatia, R., Wirths, S., Poetes, D., Kutay, U., Siebrasse, J. P. & Kubitscheck, U. 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. 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).
25. Bayliss, R., Ribbeck, K., Akin, D., Kent, H. M., Feldherr, C. M., Görlich, D. & Stewart, M. Interaction between NTF2 and xFxFG-containing nucleoporins is required to mediate nuclear import of RanGDP. *J Mol Biol* **293**, 579-593 (1999).
26. 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).
27. Frey, S. & Görlich, D. A saturated FG-repeat hydrogel can reproduce the permeability properties of nuclear pore complexes. *Cell* **130**, 512-523 (2007).
28. 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).
29. Milles, S. & Lemke, E. A. Single molecule study of the intrinsically disordered FG-repeat nucleoporin 153. *Biophys J* **101**, 1710-1719 (2011).
30. 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).
31. 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).
32. Schmidt, H. B. & Görlich, D. Transport Selectivity of Nuclear Pores, Phase Separation, and Membraneless Organelles. *Trends Biochem Sci* **41**, 46-61 (2016).
33. Frey, S., Rees, R., Schünemann, J., Ng, S. C., Fünfgeld, K., Huyton, T. & Görlich, D. Surface properties determining passage rates of proteins through nuclear pores. *Cell* **174**, 202-217.e9 (2018).
34. 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).
35. Andersen, K. R., Onischenko, E., Tang, J. H., Kumar, P., Chen, J. Z., Ulrich, A., Liphardt, J. T., Weis, K. & Schwartz, T. U. Scaffold nucleoporins Nup188 and Nup192 share structural and functional properties with nuclear transport receptors. *Elife* **2**, e00745 (2013).
36. Hyman, A. A., Weber, C. A. & Jülicher, F. Liquid-liquid phase separation in biology. *Annu Rev Cell Dev Biol* **30**, 39-58 (2014).

37. 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).
38. Shin, Y. & Brangwynne, C. P. Liquid phase condensation in cell physiology and disease. *Science* **357**, eaaf4382 (2017).
39. Hondele, M., Sachdev, R., Heinrich, S., Wang, J., Vallotton, P., Fontoura, B. M. A. & Weis, K. DEAD-box ATPases are global regulators of phase-separated organelles. *Nature* **573**, 144-148 (2019).
40. 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).
41. Martin, E. W., Holehouse, A. S., Peran, I., Farag, M., Incicco, J. J., Bremer, A., Grace, C. R., Soranno, A., Pappu, R. V. & Mittag, T. Valence and patterning of aromatic residues determine the phase behavior of prion-like domains. *Science* **367**, 694-699 (2020).
42. Musacchio, A. On the role of phase separation in the biogenesis of membraneless compartments. *EMBO J* **41**, e109952 (2022).
43. 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).
44. 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).
45. Labokha, A. A., Gradmann, S., Frey, S., Hülsmann, B. B., Urlaub, H., Baldus, M. & Görlich, D. Systematic analysis of barrier-forming FG hydrogels from Xenopus nuclear pore complexes. *EMBO J* **32**, 204-218 (2013).
46. 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).
47. 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).
48. 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).

49. 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).
50. 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).
51. 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). http://dx.doi.org/10.1101/2022.09.20.508740
52. 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).
53. 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).
54. Sung, Y., Choi, W., Fang-Yen, C., Badizadegan, K., Dasari, R. R. & Feld, M. S. Optical diffraction tomography for high resolution live cell imaging. *Opt Express* **17**, 266-277 (2009).
55. Kim, K. & Guck, J. The relative densities of cytoplasm and nuclear compartments are robust against strong perturbation. *Biophys J* **119**, 1946-1957 (2020).
56. 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).
57. Ng, S. C. & Görlich, D. A simple thermodynamic description of phase separation of Nup98 FG domains. (2022). http://dx.doi.org/10.1101/2022.07.26.501556
58. 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).
59. 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).
60. 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**, (2022).
61. 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).

62. Ader, C., Frey, S., Maas, W., Schmidt, H. B., Görlich, D. & Baldus, M. Amyloid-like interactions within nucleoporin FG hydrogels. *Proc Natl Acad Sci USA* **107**, 6281-6285 (2010).
63. Elvin, C. M., Carr, A. G., Huson, M. G., Maxwell, J. M., Pearson, R. D., Vuocolo, T., Liyou, N. E., Wong, D. C., Merritt, D. J. & Dixon, N. E. Synthesis and properties of crosslinked recombinant pro-resilin. *Nature* **437**, 999-1002 (2005).
64. 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).
65. Serio, T. R., Cashikar, A. G., Kowal, A. S., Sawicki, G. J., Moslehi, J. J., Serpell, L., Arnsdorf, M. F. & Lindquist, S. L. Nucleated conformational conversion and the replication of conformational information by a prion determinant. *Science* **289**, 1317-1321 (2000).
66. 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).
67. Bayliss, R., Leung, S. W., Baker, R. P., Quimby, B. B., Corbett, A. H. & Stewart, M. Structural basis for the interaction between NTF2 and nucleoporin FxFG repeats. *EMBO J* **21**, 2843-2853 (2002).
68. Isgro, T. A. & Schulten, K. Binding dynamics of isolated nucleoporin repeat regions to importin-β. *Structure* **13**, 1869-1879 (2005).
69. Port, S. A., Monecke, T., Dickmanns, A., Spillner, C., Hofele, R., Urlaub, H., Ficner, R. & Kehlenbach, R. H. Structural and Functional Characterization of CRM1-Nup214 Interactions Reveals Multiple FG-Binding Sites Involved in Nuclear Export. *Cell Rep* **13**, 690-702 (2015).
70. Crowley, P. B. & Golovin, A. Cation-pi interactions in protein-protein interfaces. *Proteins* **59**, 231-239 (2005).
71. Frey, S. & Görlich, D. A new set of highly efficient, tag-cleaving proteases for purifying recombinant proteins. *J Chromatogr A* **1337**, 95-105 (2014).
72. 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).
73. Wolf, E. Three-dimensional structure determination of semi-transparent objects from holographic data. *Optics Communications* **1**, 153-156 (1969).

74. Kim, K., Yoon, H., Diez-Silva, M., Dao, M., Dasari, R. R. & Park, Y. 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).
75. Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**, 676-682 (2012).
76. Zhao, H., Brown, P. H. & Schuck, P. On the distribution of protein refractive index increments. *Biophys J* **100**, 2309-2317 (2011).

Acknowledgements

S.C.N. and D.G. thank Waltraud Taxer for technical help 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 the data analysis. T.H. designed the TetraGFP^{NTR} variant. J.S. conducted preliminary experiments of the 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 experiment design, data analysis and interpretation.

Competing Interests Statement

The authors declare no competing interests.

Figure legends

Figure 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 motifs 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 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 length were designed. N-terminal sequences up to the fourth FG motifs are shown (see Supplementary Note 1 for complete sequences). For each variant, a 1 mM stock in 4 M guanidinium hydrochloride was prepared and the 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. Full scans of gels with molecular weight markers are provided in the Source Data file. The same applies to Figs. 4 and 9d.

Figure 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 dilution 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). Panels show maps of refractive index (RI). For each FG domain/ variant, ten independent FG particles were analysed (see **(d)** for the variations between the particles). Representative images and mean values of RI, mass density (C_{dense}), FG motif density and solvent % are shown. *Note: Mac98A FG domain has <52 FG motifs, but also contain 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} 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. Source data for **(a-e)** is provided in the Source Data file.

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

FG phases assembled from Mac98A FG domain and indicated variants were prepared as described in Fig. 2 legend, and were challenged with indicated permeation probes. Transportin·M9-EGFP: a complex of transportin and (FG-phobic) EGFP containing an M9 domain for transportin-binding. Importin β (Impβ) in complex with three distinct species was tested: either with EGFP containing an IBB domain for Impβ-binding, or with an FG-philic GFP variant ($sffrGFP7^{NTR}$) containing the IBB

domain, or with only the small (6 kDa) IBB domain labelled with Atto488. Xpo1·RanGTP·NES-EGFP: an "export complex" composed of Exportin 1 (Xpo1) + RanQ69L (a mutant of Ran locked in the GTP form) + EGFP fused to a nuclear export signal (NES) for Xpo1-binding. 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, 6, 7, 8, 9a, 9c and 10c.

Figure 4: Phase separation property is dependent on the overall hydrophobicity.

$GLFG_{52x12}$ 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 motifs is shown, while the rest of the sequence (~620 residues/ 48 repeat units) follow 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.

Figure 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.

Figure 6: Spatial arrangement of Phe and Leu in FG domain is crucial for NTR-binding but not for evolved 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.

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

Indicated 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 phase, which lacks aromatic residues, still allows the partition of the Arg-rich sffrGFP4, indicating interactions other than cation-π.

Figure 8: An engineered GFP variant which binds specifically to GLFG motifs and NPCs.

(a) A rational mutation (M225F) was introduced to GFP^{NTR}_3B7C and the corresponding variant is 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 aggregates with nuclear structures other than NPCs, $TetraGFP^{NTR}$ stains NPCs sharply, illustrating that $TetraGFP^{NTR}$ is a more specific NPC-binder.

(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.

Figure 9: Exchanging GLFG to FSFG motifs results in a hypercohesive FG phase that disallows 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 reduced number (= 35 or 36) of FSFG motifs were constructed. "Mutant 1" corresponds to a sequence where one in every third FSFG motif in $FSFG_{52x12}$ is mutated to SSSG (FG-phobic). "Mutant 2" corresponds to another sequence where every fifth and sixth FSFG motifs in $FSFG_{52x12}$ are mutated to SSSG. 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) Phases assembled from the above variants were challenged with the indicated probes as in **(a)**.

(d) $GLFG_{29x12}$ and $FSFG_{29x12}$ correspond to the N-terminal 29 repeats of $GLFG_{52x12}$ and $FSFG_{52x12}$ respectively and were tested for phase separation. Phase separation of these two variants was analysed at concentration = 20 μM. Note that the FSFG motifs lead to stronger phase separation propensity than the GLFG motifs (see also Supp. Fig. 2).

Figure 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 spacer length were constructed ($FSFG_{52x15}$ and $FSFG_{52x18}$).

(b) 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).

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

Figure 1

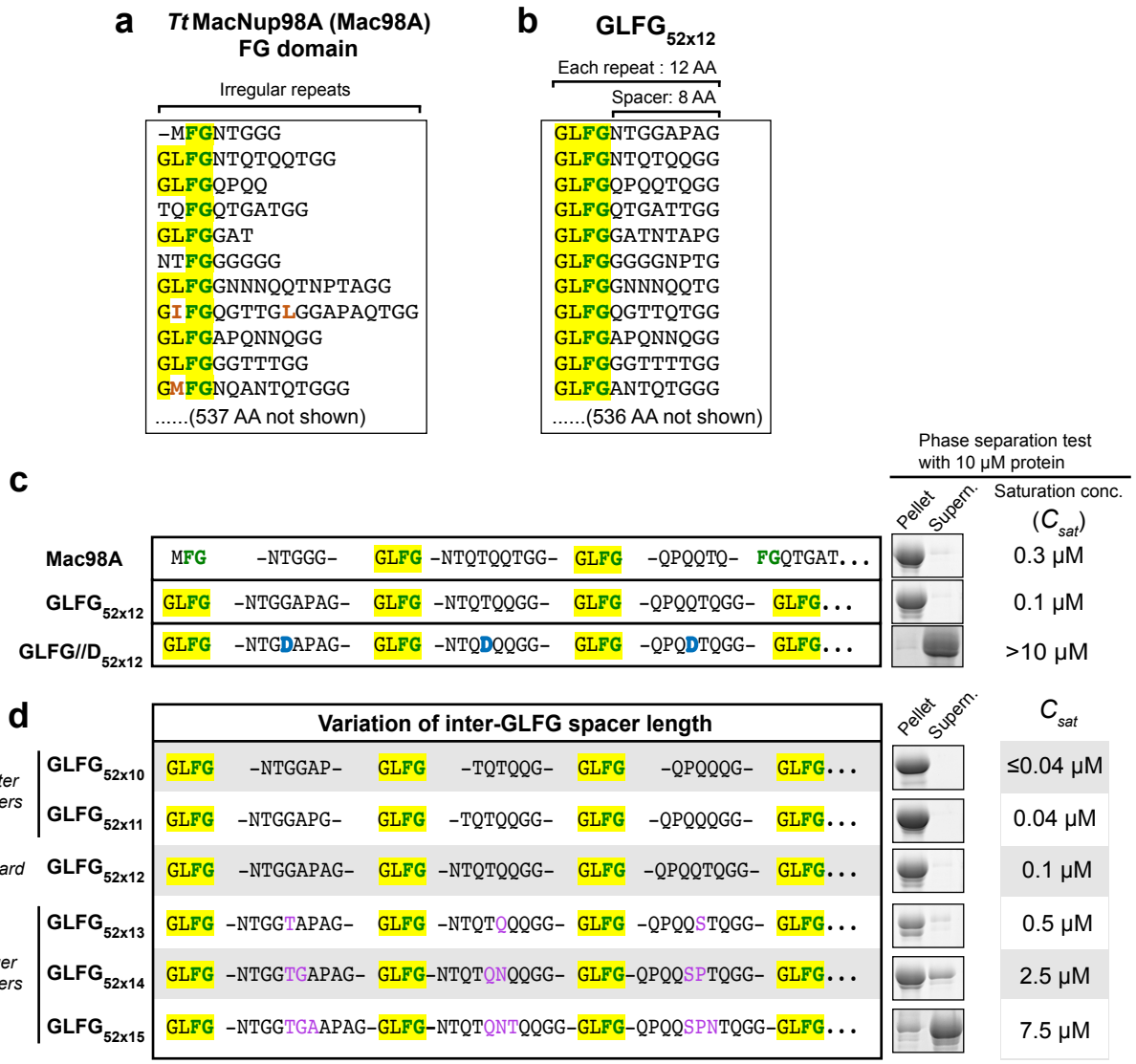

Figure 2

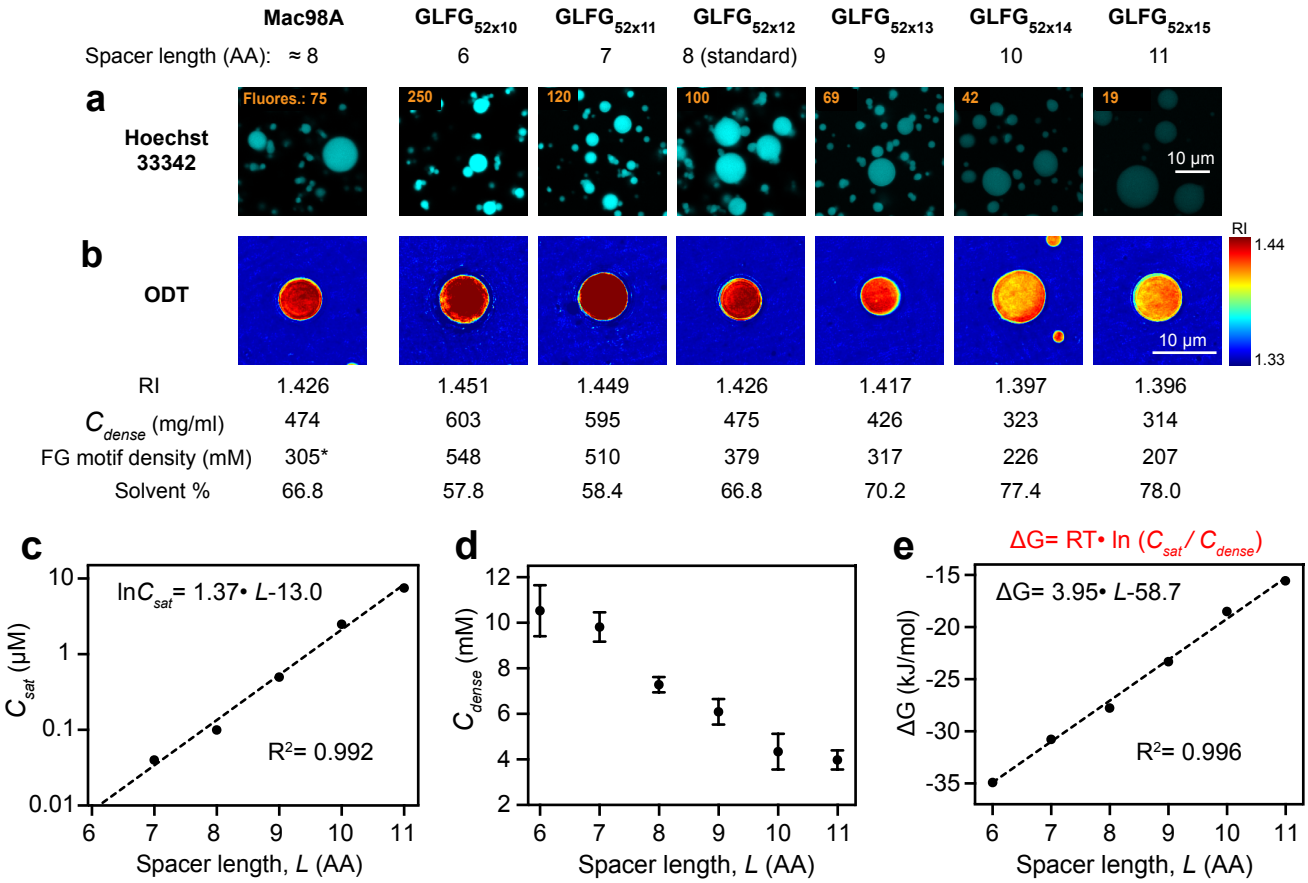

Figure 3

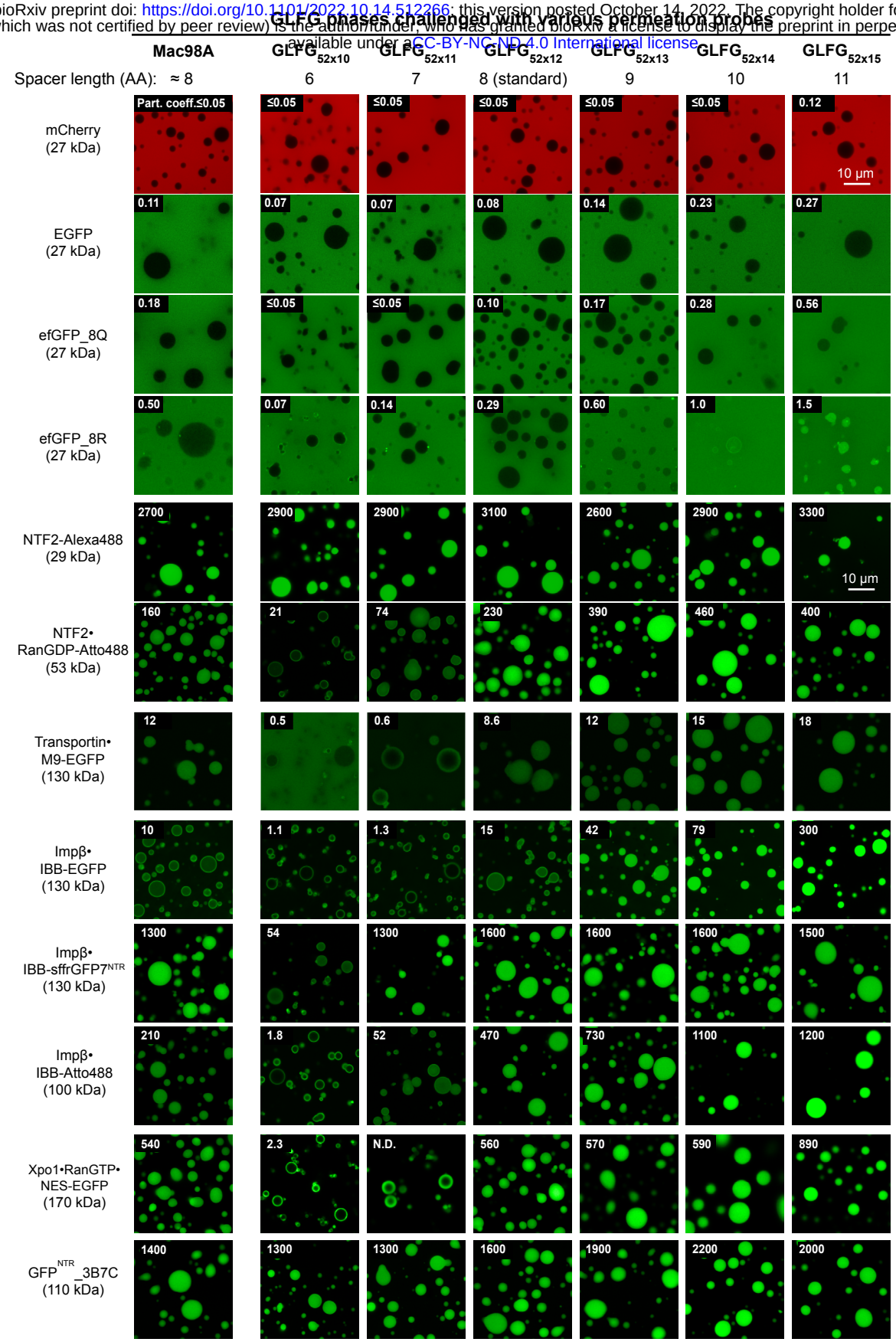

Figure 4

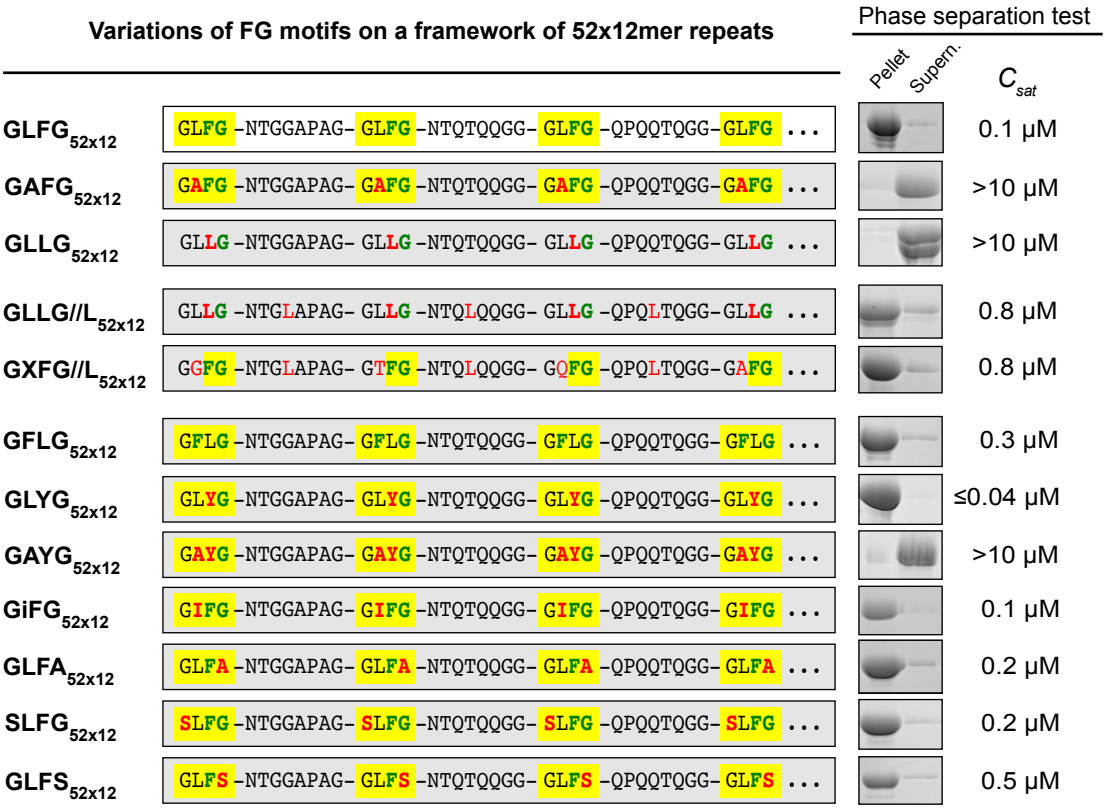

Figure 5

FG repeat variants challenged with various permeation probes

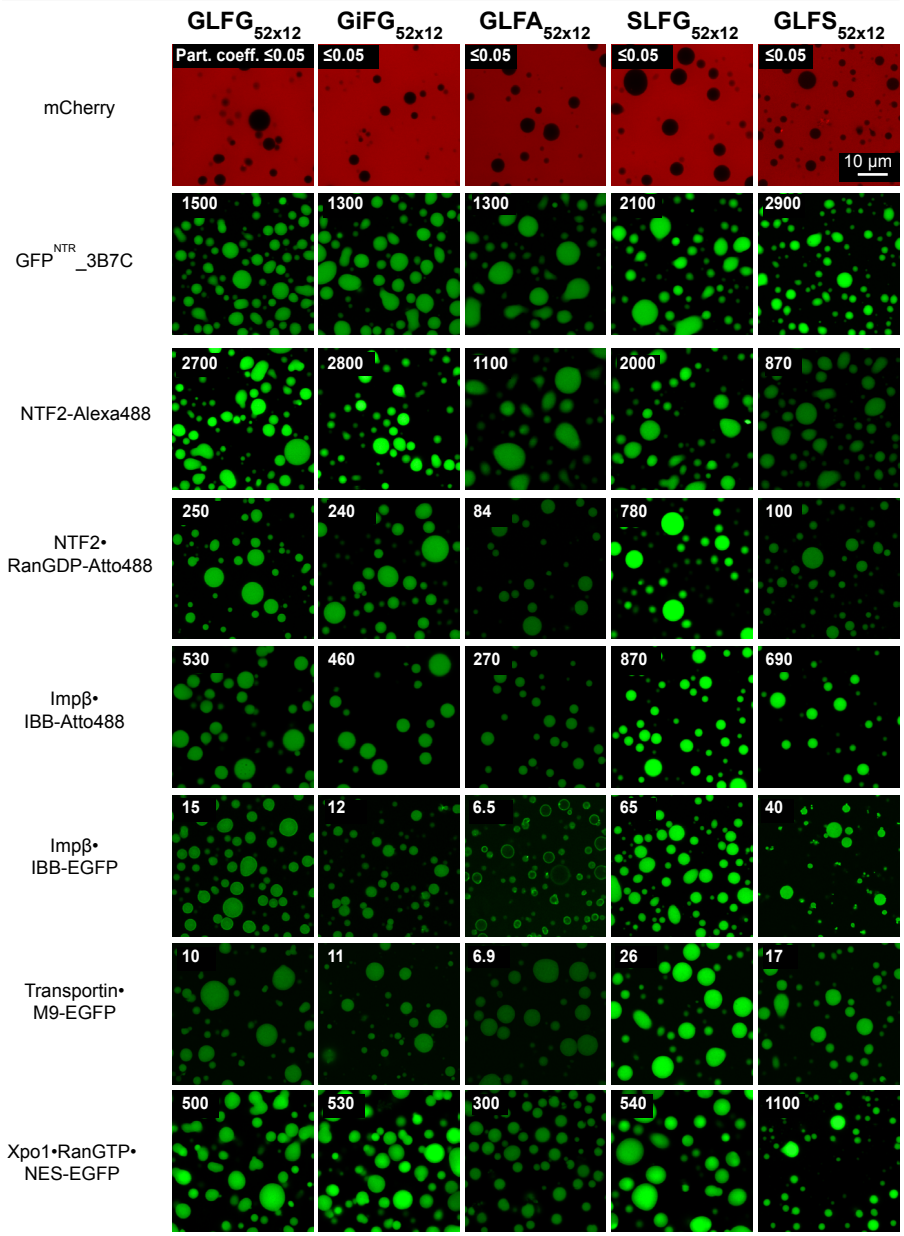

Figure 6

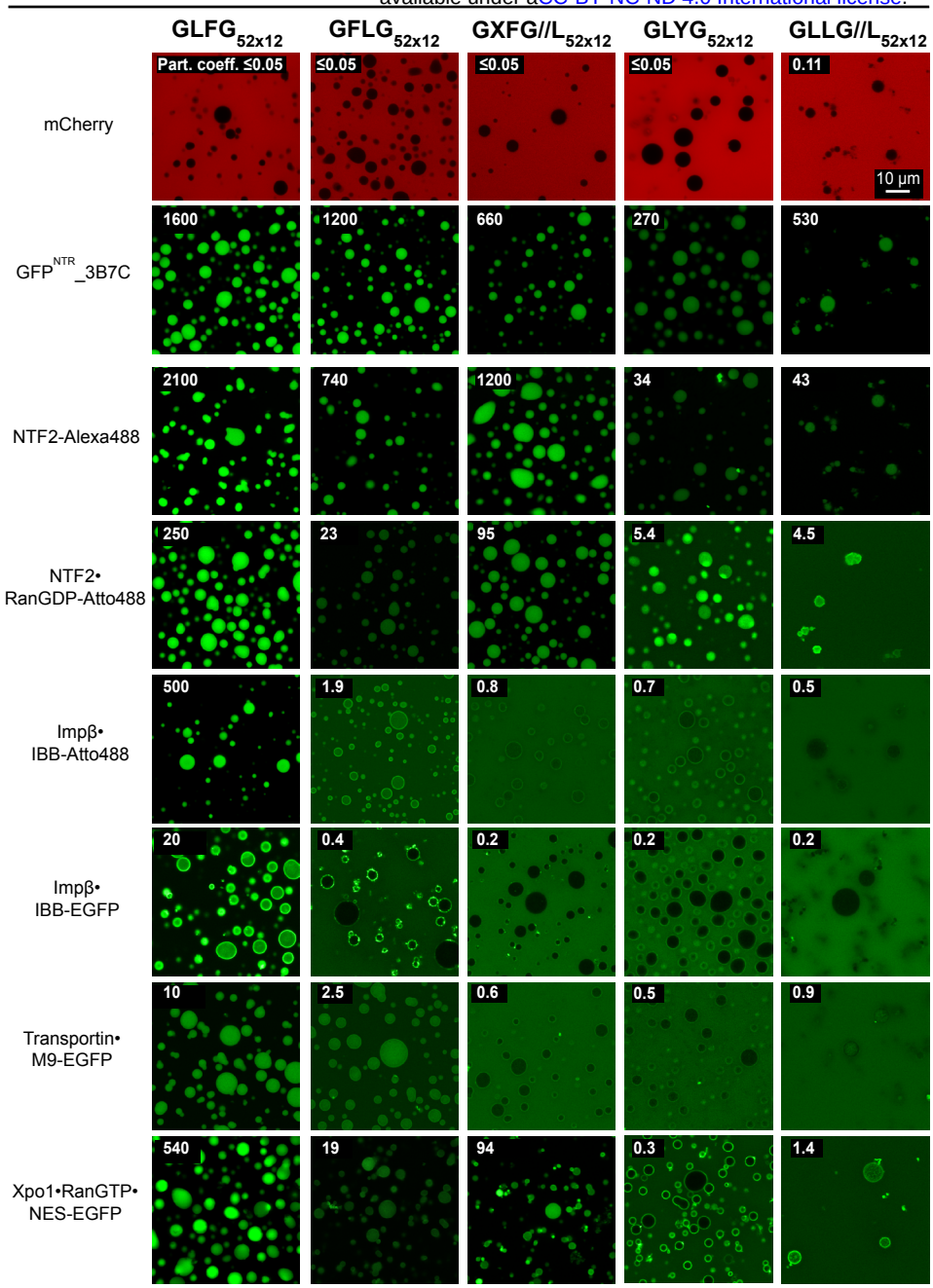

Figure 7

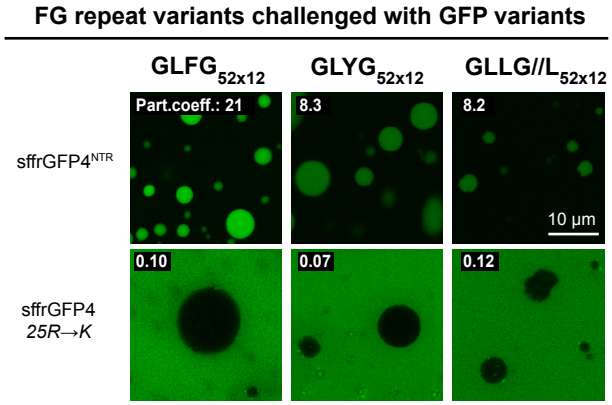

Figure 8

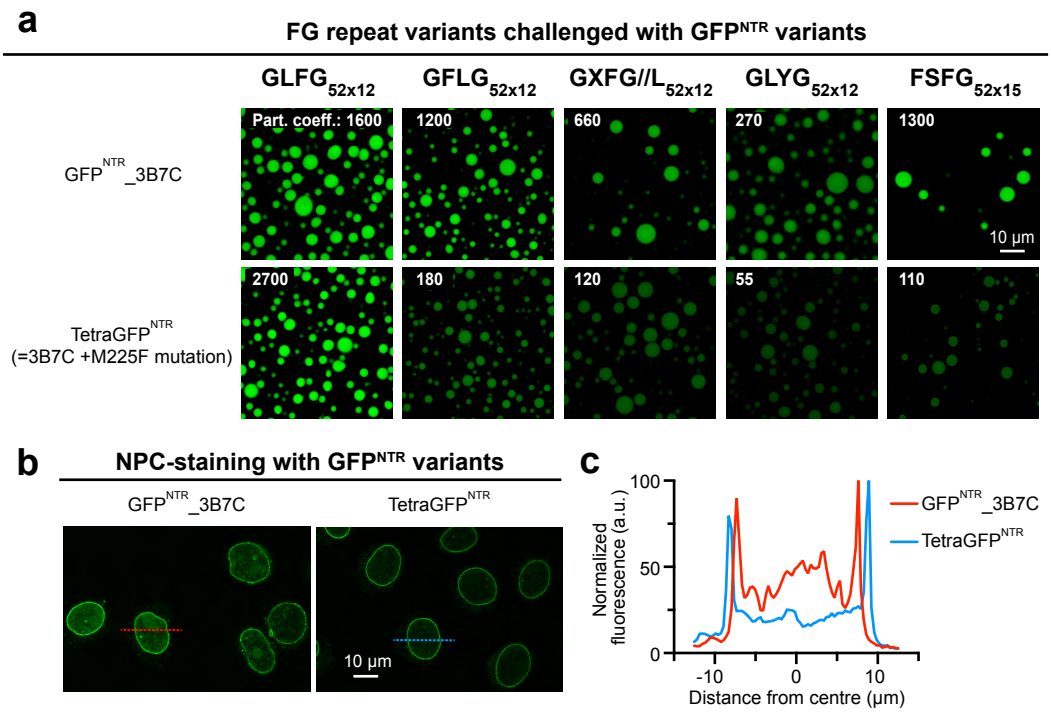

Figure 9

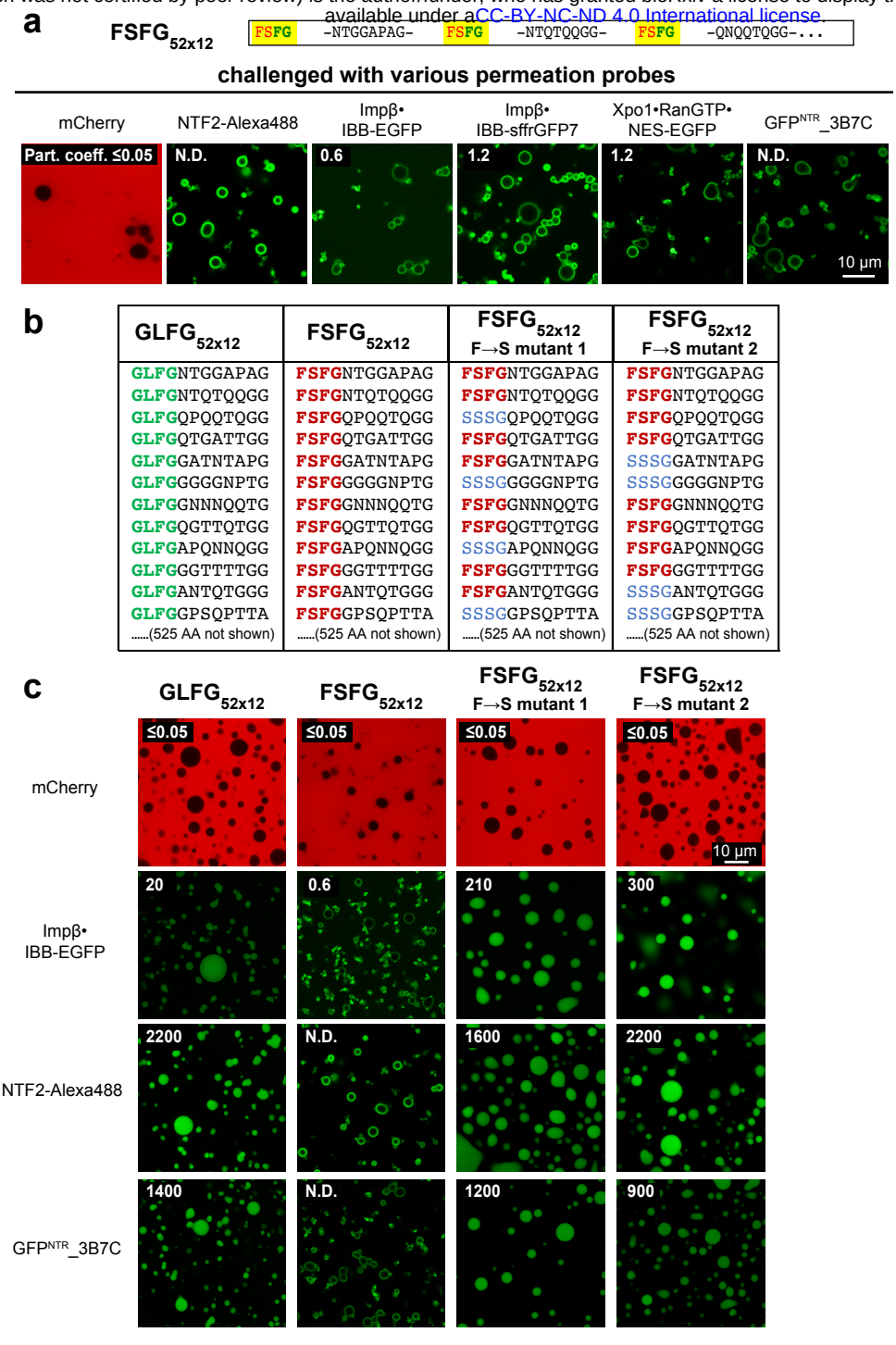

Figure 10