

Tuning Selective Transport of Biomolecules Through Site-Mutated Nucleoporin-Like Protein (NLP) Hydrogels

Yun Jung Yang^{†,‡}, Danielle J. Mai^{†,§}, Shuaili Li[†], Melody A. Morris[†], and Bradley D. Olsen^{*,†}

[†]Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

[‡]Department of Biological Engineering, Inha University, 100 Inha-ro, Nam-gu, Incheon, 22212, Republic of Korea

[§]Department of Chemical Engineering, Stanford University, Stanford, CA 94305, USA

*Correspondence: Professor Bradley D. Olsen, Department of Chemical Engineering, Massachusetts Institute of Technology, 77 Massachusetts Ave 66-558A, Cambridge, MA, 02139, USA.

Tel: +1 617-715-4548

E-mail address: bdolsen@mit.edu

Abstract

Natural selective filtering systems (e.g. the extracellular matrix, nuclear pores, and mucus) separate molecules selectively and efficiently, and the detailed understanding of transport mechanisms exploited in these systems provides important bioinspired design principles for selective filters. In particular, nucleoporins consist of consensus repeat sequences that are readily utilized for engineering repeat proteins. Here, the consensus repeat sequence of Nsp1, a yeast nucleoporin, is polymerized to form a nucleoporin-like protein (NLP) and mutated to understand the effect of sequence on selective transport. The hydrophilic spacers of the NLPs were re-designed considering net charge, charge distribution, and polarity. Mutations were made in, near, and far from the FSFG interacting domain to explore the role of highly conserved residues as a function of spatial proximity. A nuclear transport receptor-cargo complex, nuclear transport factor 2-green fluorescent protein (NTF2-GFP), was used as a model for changes in transport. For mutations of the charged spacer, some mutations of highly conserved charged residues were possible without knocking out selective transport of the NTF2, but the formation of regions of clustered negative charge has an unfavorable effect on nuclear transporter permeation. Thus, positive net charge and alternating positive and negative charge within the hydrophilic spacer are advantageous for recognition and selective transport. In the polarity panel, mutations that increased the interaction between NTF2-GFP and the gel led to decreased permeation of the NTF2-GFP due to blocking of the interface and inability of the NTF2-GFP to transport into the gel. Therefore, these results provide a strategy for tuning selective permeability of biomolecules using the artificially designed consensus repeat-based hydrogels.

Keywords: selective transport, charge, site-mutation, hydrogel, nucleoporin, purification

Introduction

Modular peptides have been widely engineered into artificial proteins to mimic the outstanding properties of natural materials. Durable silk,¹ thermoresponsive elastin,^{2,3} cell adhesive integrin,⁴ and mineralized dentin⁵ are all good examples. Since the properties of protein-based materials are derived from sequence and structure, rationally designed sequences are one of the ways to achieve changes in structural, mechanical, and biophysical properties that generate novel properties.⁴⁻⁸ Sometimes changes of even a single amino acid per consensus repeat result in drastic changes in material properties.⁹ For example, replacement of the first glycine in the elastin-like polypeptide (ELP) repeat VPGVG with L-alanine to yield VPAVG changes the mechanical properties to become more similar to a plastic than to an elastomer.¹⁰ Many other proteins such as those derived from antimicrobial peptides,¹¹ enzymes and antibodies,^{6,7,14} or antigens and antivirals¹⁵ have also been re-designed to enhance activity, selectivity, or stability over natural counterparts.

In protein materials, selective permeation is a valuable property for applications such as pharmaceutical purification,¹⁶ toxin removal,¹⁷ or molecular delivery.¹⁸ Industrial separations such as the separation of monoclonal antibodies are often challenging because of the similarity between the different molecules to be separated (for example antibodies that differ only in the variable region). However, natural selective filtering systems (the extracellular matrix,¹⁹ nuclear pores,^{16,20} and mucus²¹) perform these challenging separations with both high flux and selectivity.²² Among them, nucleoporin is a promising candidate to engineer into synthetic protein materials due to its consensus repeat sequence. Nucleoporins reside in nuclear pore complex (NPC) of the nuclear envelope, allowing around 1,000 specifically interacting molecules to traffic through each NPC per second.²³ The dense nucleoporin sieve non-specifically blocks large molecules, and specific binding interactions between phenylalanine–glycine (FG) repeats of nucleoporin and nuclear transport receptors (NTRs) allow selective passage of only NTRs and their cargo. Example NTRs include importin β and nuclear transport factor 2 (NTF2).²⁴ Many nucleoporins

have FG repeats occurring across large regions of the protein sequence spaced apart by sequences of hydrophilic amino acids. The highly conserved FG sequences are known as the docking site of all transport proteins, and the loosely conserved charged hydrophilic spacers are believed to promote transporters' permeation.¹⁸⁻²⁷

Although nucleoporins show remarkable selective permeability of biomolecules, it is very difficult to produce full nucleoporins recombinantly. Moreover, NPCs consist of more than 30 types of nucleoporins with different sequence and architecture.²⁸ This low recombinant yield and system complexity motivated the need for a simplified sequence design to produce engineered materials that mimic the selective filtration by nucleoporins. In a previous study, Olsen's lab demonstrated the ability to extract a minimal consensus repeat from the nucleoporin Nsp1. A single amino acid change (Asp to Ser) per 19-mer consensus repeat changes both the rate of diffusion and accumulation of the importin β within gels formed from the high molar mass consensus repeat polymers.²⁰ In another study, Chen et al. developed a nucleoporin-inspired peptide containing two FG repeats linked by a charged spacer where the charge type and distribution in the midblock determined self-assembly and binding with other charged peptides.²⁹ These findings suggest that the electrostatics of the consensus repeat design may affect its interaction with permeant molecules and their transport rates through the material. While these initial studies show the potential for mutations to simplified nucleoporin sequences to understand transport properties, a thorough study on the effect of mutations on the diffusion of a model NTR has not yet been performed.

In this present paper, site-mutated nucleoporin-like proteins (NLPs) were designed considering charge distribution and polarity across a variety of positions in the nucleoporin's consensus repeat. These NLPs were polymerized into protein hydrogels with a dense NLP meshwork capable of non-specifically repelling proteins due to entropic effects while permitting passage of NTRs due to specific binding interactions.^{16,30} Since the role of spacer sequences has not been resolved due to the inherently unfolded FG peptides,³¹ this site-mutated NLP approach is meaningful to explore how the NLP's sequence changes affect

selective diffusion of nuclear transporters. Mutations of highly conserved residues proximate to the FG region were also explored to understand their specific role in transport. The observed variations in permeation using a model transporter nuclear transport factor 2 (NTF2) provides a method to tune the selective permeability of biomolecules in artificially designed protein gels.

Materials and Methods

DNA cloning

Two consensus sequences derived from yeast Nsp1 were originally designed for recombinant nucleoporin-like protein (P-1NLP-P and P-2NLP-P)²⁰ synthesis. Additional site-mutated consensus repeats were generated by replacing amino acids in 1NLP or 2NLP to modify net charge, charge distributions, and polarity of highly conserved residues near the FSFG region (**Table 1**). All 19-mer NLP peptides were repeated 8 times and chemically synthesized (GenScript, USA) with *NheI* and *SpeI* sites at 5' and 3', respectively. The cloning vector was modified to produce a sequence with the following order: *BamHI*-adenine-pentameric coiled-coil associative domains³²-*NheI*-(19-mer NLP peptide)₁₆-*SpeI*-pentameric coiled-coil associative domains-*HindIII*, in the pET22b(+) plasmid (EMD Millipore, USA), which contains a C-terminal 6x His tag (**Figure S1**). Adenine was added for the correct frame for expression, and 8-mer NLP peptides were doubled in between *NheI*-*SpeI* for the 16 repeated NLP midblock to match the length of the C-terminal region of the native nucleoporin Nsp1 from which they were derived.

P-NLP-P expression and purification

All mutant NLPs were expressed in *E.coli* OverExpress C41(DE3) cells (Lucigen, USA) in terrific broth (TB) media with 100 µg mL⁻¹ ampicillin at 37 °C. Isopropyl β-D-1-thiogalactopyranoside (IPTG) induction was performed at O.D.₆₀₀ ~ 1.0 with the addition of 1 mM IPTG. After induction, cells were cultured at 24 °C overnight. Cells containing P-NLP-Ps were lysed in denaturing lysis buffer (10 mM Tris, 100 mM

NaH₂PO₄, 10 mM imidazole and 8 M urea, pH 8.0) and frozen at -80 °C. Thawed lysates were completely dissolved, sonicated, and clarified by centrifugation (14,000 g for 60 min at 4 °C). Ni-NTA affinity chromatography was performed under denaturing conditions, and 250 mM imidazole was used for elution. NLP elutants were dialyzed against TEN buffer (10 mM Tris, 1 mM EDTA, 50 mM NaCl, pH 8.0) three times followed by distilled water seven times. 20 mM Tris and 6 M urea were added into the elutant and pH was adjusted to 8.0. Anion exchange chromatography with a HiTrap Q HP column (ÄKTA pure, GE Healthcare, Sweden) was carried out by increasing the sodium chloride concentration to 0.4 M. Following chromatography, dialysis was performed again three times with TEN buffer and seven times with distilled water. The dialyzed protein solutions were centrifuged to precipitate impurities, and the supernatant was lyophilized. P-C₃₀-P (Table 1) was also prepared as reported previously^{20,33}. The protein purities were determined by 12% sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and visualized by Coomassie brilliant blue staining (**Figure 1c**).

NTF2-GFP expression

NTF2-GFP was expressed in strain SG13009(pREP4) using 2XYT media with 50 µg mL⁻¹ kanamycin and 100 µg mL⁻¹ ampicillin at 37 °C.³³ IPTG induction was performed at O.D.₆₀₀ ~ 0.5 with the addition of 0.5 mM IPTG. After induction, cells were cultured at 24 °C overnight. Cells were frozen at -80 °C and thawed in non-denaturing lysis buffer (50 mM NaH₂PO₄, 300 mM sodium chloride, 10 mM imidazole, 10 mM β-mercaptoethanol, pH 8.0). After sonication, DNase, RNase, and lysozyme (100 mg L⁻¹) were mixed and incubated at 4 °C for an hour. Supernatants were obtained by centrifugation (14,000 g for 60 min at 4 °C). Ni- NTA affinity chromatography was performed with non-denaturing lysis buffer, and target protein was eluted with 250 mM imidazole. Anion exchange chromatography was performed using a HiTrap Q HP column (GE Healthcare, Sweden) in an ÄKTA pure FPLC under non-denaturing condition, with sodium chloride increase to 0.4 M. Final protein solutions were concentrated by Amicon® Ultra-15 centrifugal filter

(10 kDa molecular weight cutoff, MilliporeSigma) with phosphate-buffered saline (PBS, Corning Inc., USA).

Fluorescent tag labeling

NTF2-GFP was labeled by fluorescein isothiocyanate (FITC) with a molar ratio of 1 : 10 in PBS (pH 8.8) under 4 °C overnight. FITC labeled NTF2-GFP was purified using ion exchange chromatography under non-denaturing condition, with lysis buffer (50 mM NaH₂PO₄, 300 mM sodium chloride, 10 mM imidazole, 10 mM β-mercaptoethanol, pH 8.0) and sodium chloride increase to 0.4 M. Excess fluorescent dye was removed after conjugation using an illustra NAPTM 10 column (GE Healthcare). The samples were dialyzed against 20 mM Tris (pH 8) and 100 mM sodium chloride, then the buffer was exchanged to PBS (pH 8.0) by using a 10 kDa centrifugal filter (catalog # UFC901008, EMD Millipore). The final concentration and degree of labeling were determined by measuring absorbance at 280 and 495 nm. The final degree of labeling was 0.52 dye molecules per NTF2-GFP. Bovine serum albumin (BSA, New England Biolabs, USA) was fluorescently tagged using Alexa Fluor® 594 NHS-ester (Thermo Fisher Scientific) with the same manner as NTF-GFP except at a protein:dye molar ratio of 1 : 100. The final concentration and degree of labeling of BSA were determined by measuring absorbance at 280 and 590 nm. The final degree of labeling was 5.1 dye molecules per BSA.

Material characterization

20 w/v% of NLP hydrogels were prepared by dissolving lyophilized NLP in 50 mM Tris/HCl and 200 mM NaCl (pH 7.5) buffer. Gelation occurred at 4°C overnight. Oscillatory shear measurements were performed on an Anton Paar MCR 702 TwinDrive rheometer (Anton Paar, Austria). A cone-and-plate geometry with 10 mm diameter and 1.975° cone angle was used. NLP hydrogels were loaded onto the bottom plate, and the cone was lowered to the final test position. Subsequently, the outer edge of the gap was sealed with a mineral oil barrier to prevent dehydration. Frequency sweep experiments were performed between 0.01

and 50 rad s^{-1} at 1% strain, which was confirmed to be in the linear viscoelastic regime by strain sweep experiments (0.1-10% strain at 1 rad s^{-1}) (**Figure S2**). All measurements were taken at 25°C using Peltier temperature control, and three replicates of each NLP hydrogel sample were tested.

To perform transport assays, NLP hydrogels were loaded into one side of rectangular borosilicate capillaries with 0.9 mm inner diameter (Vitrocom, USA). The capillaries were then centrifuged at 5000 rpm for 10 min to flatten the gel interface. The other side of the capillary was filled with transport solutions of NTF2-GFP and BSA in PBS buffer, mixed to final concentrations of $5 \mu\text{M}$ each. Loading of the capillaries requires that the NLP hydrogels display some degree of shear-thinning behavior, as they are injected with a syringe. The capillary ends were sealed with a 1:1:1 mixture of Vaseline, lanolin, and paraffin wax. Images were taken with a Zeiss Axioplan fluorescence microscope equipped with an Axiocam 503 mono camera, FluoArc control unit, and HBO 100 mercury plasma lamp (Carl Zeiss, Germany). Images were acquired 5 and 60 minutes after the transport solutions were added. Fluorescence intensity profiles were obtained by ImageJ (National Institutes of Health). A background subtraction for uneven lighting was developed by fitting the bath region of the 5 minute sample of the capillary to a quadratic and altering the constant term to be the minimum intensity far into the gel. The relative intensity was then calculated by subtracting the background from the intensity and normalizing by the difference in the bath and gel. The maximum intensity of NTF2-GFP and BSA transported in hydrogel were measured by obtaining the fluorescence intensity profiles from the gel interface to about $500 \mu\text{m}$ inside the gel. The integration was determined by a numerical integration through the entire gel region.

Results and Discussion

Design of site-mutated nucleoporin-like proteins (NLPs)

On the basis of the nucleoporin repeat previously demonstrated,²⁰ a panel of mutated repeat sequences was prepared by adjusting spatial charge distribution and polarity. Due to the low expression

level of proteins having 3NLP, 7NLP, 8NLP, 12NLP and 13NLP midblocks even after multiple rounds of optimization, the experiments focused on the other 8 P-NLP-P proteins listed in **Table 1**. The isoelectric point (pI), molecular weight (M.W.), and charge at pH 7.5 are derived using PROTEIN CALCULATOR v3.4.³⁴ The first group of proteins (**Table 1a**) varies the charge of highly conserved midblock residues because it is well known that the midblock charge affects protein's interactions with transport proteins.^{29,35,36} The original two sequences 1NLP and 2NLP were designed with the mutation S15D. 4NLP was designed with the mutation K8E. 5NLP and 6NLP have mutations K19S and K19E, respectively. The 8th and 19th positions were selected because they are highly conserved and proximate to the conserved FSFG repeat, providing a test of how bordering electrostatic residues contribute to selectivity. For this group of mutations, the calculated charge on the protein varied from -37.8 to -5.8 at pH 7.5 because each mutation was repeated 16 times throughout the P-NLP-P. In all cases, mutations were prepared in pairs of varying strength, although both members of all pairs could not always be expressed, and the more aggressive electrostatic mutations typically did not express well.

The second group (**Table 1b**) involves mutations to change the polarity of conserved residues. From the base sequence 2NLP, 9NLP contains the mutation P1S, 10NLP contains the mutation A2S, and 11NLP contains the mutation A7S. All of these residues are highly conserved (**Figure 1**) and bordering the FSFG repeat. The net charge remained the same at -5.8 for all proteins in this group. A number of other mutations were attempted; however, suitable quantities of the protein could not be expressed for materials characterization.

The mutated consensus sequence was repeated 16 times and fused with coiled-coil domains (P) at each end (**Figure 1a**) to enable gel formation through P-domain association.^{33,37} This midblock length matches the number of consensus repeats in the C-terminal region of native Nsp1. As a non-NLP control, P-C₃₀-P protein was utilized, which consists of a structureless midblock of similar molar mass to the NLP₁₆ repeats (**Figure 1b**) while maintaining the same overall triblock gel-forming structure.^{20,33} Production and

purification of all proteins was visualized by SDS-PAGE (**Figure 1c**). P-C₃₀-P appeared at a molar mass of 58 kDa on the SDS-PAGE; this shift in apparent molar mass has been observed in a previous study due to the highly unusual distribution of amino acids in the protein.³³

Gelation of NLP proteins

The P domain facilitates gelation of the P-NLP-P constructs in only a few minutes, driven by association of the coiled-coils into pentameric bundles. This gelation time is much faster than the natural nucleoporin, which requires several hours to form a gel (**Figure 2a**)^{20,38}. Since all NLP mutants have a molar mass of approximately 45 kDa, with the same 19 amino acid consensus repeat length, all NLP hydrogels are expected to have a similar pore size. This mesh-like structure (**Figure 2b**) acts as a size-dependent barrier which expels non-specific molecules larger than size cut-off. The P-C₃₀-P hydrogel was used a negative control, as it shows no specific binding interaction with target molecule (NTF2, nuclear transport factor 2).²²

Gel formation is demonstrated by linear oscillatory shear rheology, where all gels show a similar linear mechanical properties. Across all NLP mutant hydrogels, the storage modulus (G') was independent of frequency and larger than the loss modulus (G'') in the high frequency region (1-50 rad s⁻¹). The crossover frequency of the storage and loss moduli, indicative of the relaxation time of the dynamic crosslinks between P domains, appeared between 0.03-0.06 rad s⁻¹ for all gels (**Figure S3**). The plateau storage moduli (G') of all P-NLPs-P hydrogels range from 4.7±0.7 kPa to 6.4±0.9 kPa. The P-C₃₀-P gel showed a similar G' of 5.1±0.3 kPa (**Figure 2c and Figure S3**). This finding is consistent with molecular theory predictions that the modulus is governed by the number of elastically active chains, which should be approximately equal due to the equivalent molecular topology between all of the different constructs. Importantly, this rheological data shows that the macroscopic linear mechanical properties, both relaxation time and modulus, of the gels are minimally affected by the sequence mutations performed; it should be noted that some site mutations

can cause changes to the rheological behavior of gels. This sequence-independent mechanical behavior implies that factors relating to network structure and dynamics that may affect protein transport through the gels remain unchanged across the mutant panel, and the effect of the mutations is therefore restricted to specific physicochemical interactions with the target proteins of interest.

Specific binding interactions control transport

To measure selective and non-selective transport, two different model transport probes were developed with approximately equal molar mass and physical size. Alexa 594-tagged bovine serum albumin (BSA) has ~66 kDa molar mass and 3.5 nm Stokes' radius and was chosen as a non-specific model protein (**Figure 3a**).^{39,40} Nuclear transport factor 2 (NTF2, ~14 kDa, 2.5 nm of Stokes' radius)⁴¹ was chosen as a model nuclear transport protein due to its size and previously established transport characteristics. NTF2 is an essential small homodimeric transport protein in yeast, and it presents as dimers at concentrations above 5 μ M.⁴² NTF2 permeates the nuclear pore complex quickly due to its relatively weak binding to FG repeats compared to other transporters.^{42,43} NTF2 was recombinantly fused with green fluorescent protein (GFP, ~28 kDa, 2.35 nm Stokes' radius)⁴¹ (**Figure 3b**) to produce a protein with a higher molar mass than BSA. Since NTF2 forms a homodimer of two subunits in its active form, the final molar mass of NTF2-GFP is ~86 kDa, which is large enough to be rejected from NLP hydrogels based on size alone.

Protein transport within NLP materials was strongly dependent upon the sequence of the NLP. Repulsion of the non-specific large probe BSA (**Figure 4a, 5a**) and selective permeation of the nuclear transporter NTF2-GFP fusion (**Figure 4b, 5b**) were evaluated via 1-dimensional transport assays. Hydrogels were filled in borosilicate capillaries, and the headspace was filled by a mixed solution of red fluorescent BSA and green fluorescent NTF2-GFP. The capillary was sealed and laid horizontally to minimize flow and gravitational effects during the diffusion assay (**Figure 3c, d**). Increases in fluorescence signal between the buffer, buffer-hydrogel interface, and hydrogel were depicted as images (**Figures 4c,**

5c) and graphs (**Figures S2 and S3**). As **Figures 4a and 5a** show, the red fluorescent BSA was almost entirely rejected from the P-NLP-P and P-C₃₀-P hydrogels. After 60 min, there was a small amount of BSA within the gel due to gel swelling. This limited diffusion is negligible compared to the diffusion of NTF2-GFP, further discussed below. It shows that the molecules of ~ 66 kDa or 3.5 nm hydrodynamic radius are large enough to be blocked by the polymer network formed by P-NLP-P and P-C₃₀-P hydrogels. The NTF2-GFP did not diffuse into the P-C₃₀-P hydrogel, exhibiting this entropic repulsion effect (**Figure 5c, P-C₃₀-P hydrogel**) characteristic of non-specifically interacting molecules.¹⁶

On the contrary, when there is a specific interaction between the transport protein and the NLP hydrogel, the entropic repulsion was overcome to allow transport through the gel. NTF2-GFP diffusion across the hydrogel interface was observed in NLP hydrogels, with a different degree of fluorescence intensity (**Figures 4b and 5b**) depending on the NLP sequence. The time-dependence of this signal is illustrated in **Figure S5**. **Figures 4c and 5c** show selective permeability of NTF2-GFP against BSA at 60 min. The permeation is analyzed in two ways (**Figure 6**): the maximum intensity (typically close to the gel/liquid interface) and the total integration of the intensity within the gel. The maximum intensity correlates to the diffusion across the liquid/gel interface, whereas the integration correlates to diffusion within the gel. In relation to previously published theory for transport in NLP gels,⁹ the maximum intensity and total integration of the intensity represent distinct physical properties; the maximum intensity correlates with the recognition site availability (β) and the equilibrium binding coefficient (K_{eq}) whereas, assuming the protein is able to permeate into the gel, the integration correlates to δ , the diffusivity ratio, and γ , the bound diffusivity ratio. These data clearly indicate that the specific binding interaction between NTF2-GFP and NLP hydrogels enables the transporting molecules to overcome the entropic repulsive barrier and enter the gels at a much higher rate.

Despite being located relatively far from the FSFG interacting domain, mutations of charged residues within the hydrophilic midblock have a substantial impact on transport through the NLP gels. In

the case of the electrostatic mutant series (**Table 1a**), all NLP hydrogels share the same 1st -7th amino acids including the FSFG binding site and its neighboring residues. In this group, P-2NLP-P showed the highest fluorescence intensity after an hour. P-6NLP-P also presented relatively high NTF2-GFP permeation and accumulation. The maximum intensity of NTF2-GFP at 60 min of P-2NLP-P and P-6NLP-P were 2.89 ± 0.05 , and 2.02 ± 0.11 , respectively, in comparison to 1.05 ± 0.02 and 1.04 ± 0.06 for the BSA. These differences represent 2.75 and 1.95-fold increases values compared to the maximum BSA intensity, indicating that the NTF2-GFP selectively permeates through the gel over BSA. The integration of NTF2-GFP at 60 min of P-2NLP-P and P-6NLP-P were 1.82 ± 0.36 and 1.65 ± 0.17 fold increases in comparison to the P-C₃₀-P controls. In contrast, P-1NLP-P, P-4NLP-P, and P-5NLP-P exhibited statistically insignificant NTF2 fluorescence accumulation near the gel interface at 60 min (**Figure 4c**, **Table S2**). These three NLP hydrogels did not show selective permeation of NTF2-GFP when compared to BSA (**Figure 4**, **Table S2**).

The construct P-2NLP-P has the most positive charge of any of the constructs (-5.8 at pH 7.5), which may explain its highest transport and permeability for NTF2-GFP which has a charge of -12.0 at pH 7.5. In contrast, the rest of the P-NLP-P constructs that are electrostatically mutated have negative charges of -21.8 or -37.8. In natural Nsp1, the positively charged nucleoporin domains are believed to promote the specific translocation of negatively charged transport receptors more readily.³⁵ However, P-6NLP-P represents an exception to this rule; it had strong permeability with a negative charge of -37.8. This clearly indicates that the spatial position of the charged residues has a key impact on the transport properties. In the case of both P-6NLP-P and P-4NLP-P, the charge mutations occur to highly conserved lysine residues that border the FSFG repeat region. Although the mutated residues are highly conserved for both proteins, the results suggest that mutation of the lysine that is N-terminal to the FSFG is well tolerated, while mutation of the lysine that is C-terminal to the FSFG is not well tolerated. A key difference between the two mutation sites is that the mutation to P-4NLP-P creates a negative charge cluster within the protein

over the sequence from the 8th to 11th position where the mutated sequence reads EPDE. A similar negative charge cluster of three aspartic acids (DDD) is formed in P-1NLP-P which also shows no specific transport. Therefore, it can be concluded that some mutations of highly conserved charged residues are possible, but the formation of regions of clustered negative charge has an unfavorable effect on nuclear transporter permeation. Thus, positive net charge and alternating positive and negative charge within the hydrophilic spacer are advantageous for selective permeability of NTF2-GFP over BSA.

To complement the electrostatic panel, a second series of experiments was performed to systematically explore the role of three highly conserved residues proximate to the FSFG repeat region in the NLP sequence design. For this series, all NLP hydrogels have the same net charge and charge distribution from the 8th to the 19th position. Mutations were made to highly conserved amino acids in the 1st, 2nd, or 7th positions surrounding the FSFG repeat (mutations in the FSFG repeat have been shown to eliminate binding⁴⁴). In this group, P-10NLP-P (mutation A2S) showed the best selective permeation and NTF2-GFP accumulation, indicating that not all mutations of conserved residues are detrimental to transport properties in NLPs. The maximum fluorescence intensity of P-10NLP-P near the gel interface was 3.36 ± 0.32 (**Figure 6**). P-9NLP-P also showed high selectivity with a maximum intensity of 1.97 ± 0.24 . Additionally, P-10NLP-P had statistically higher integrated intensity through the entire gel compared to the control, whereas P-9NLP-P was statistically similar. These differences indicate that P-10NLP-P is able to better transport NTF2-GFP through the gel, whereas P-9NLP-P tends to simply accumulate nuclear transporter in the interfacial region. The maximum fluorescence intensity of P-11NLP-P was 1.07 ± 0.02 , indicating that it does not present selective permeation of target molecules. This behavior is confirmed by the absence of a peak in concentration near the interface, which would indicate interaction between the gel and nuclear transporter. As expected, there was no specific binding between NTF2-GFP and the P-C₃₀-P hydrogel (1.05 ± 0.02).

Since all three mutations modified highly conserved residues near the FSFG binding site by

increasing the polarity of the changed residue, it was expected that selective permeation might be affected by changes in the binding interactions of NTF2 with the NLP. However, only mutation of the alanine in position 7 after the FSFG repeat knocked out selective transport. It is possible that the mutation of certain residues proximate to the FSFG repeat can affect NTF2 binding to the NLP, either strengthening or weakening the binding constant away from the optimum for selective permeation predicted with simple continuum transport modeling of diffusion.¹⁶ A competing hypothesis is that FSFG repeats form dynamic crosslinks that modulate the mesh size of the gel; the mutations considered here could also affect this process.^{45,46} As it is expected that P-11NLP-P has strong binding interactions with NTF2-GFP, it is possible that the NTF2-GFP was accumulated mainly at the liquid-gel interface, restricting further NTF2-GFP diffusion. From these studies, it is clear that there are a number of synergistic effects between charge (and resulting charge clusters) and polarity that provide opportunities for further experimentation. Neither overall charge nor polarity around the FSFG repeat can fully account for the selective permeability of NTF2-GFP in certain P-NLP-Ps. By balancing the strength of the interaction between the P-NLP-P and NTF2-GFP, the NTF2-GFP can both interact with the gel and also diffuse through the network to allow higher permeation.

Conclusions

This study explored the effect of various mutations on selective transport in nucleoporin-like proteins (NLPs), a family of repeat proteins extracted from the yeast nucleoporin Nsp1 that mimic the natural protein's selective transport properties. The amino acid sequences derived from nucleoporin's consensus repeats were site-mutated by considering net charge, charge distribution, and substitution of highly conserved residues proximate to the known FSFG binding site to provide a deeper understanding of selective permeation. The selective permeation of NTF2-GFP was stronger through proteins with a more positive charge, and clusters of negative charge had a particularly large impact in completely knocking out

selective transport. The position of polar groups near FSFG affected binding interactions and selective permeation. High NTF2-GFP binding interactions caused less accumulation of NTF2-GFP in the gel because transport was restricted to the interface and additional NTF2-GFP passage was blocked. Likewise, P-NLP-P hydrogels with lower binding interactions with NTF2-GFP were able to continuously transport NTF2-GFP past the interface, leading to higher accumulation. In summary, this series of mutations shows that while many highly conserved residues in the NLP sequence are tolerant to mutation, changes to others can knock out selective transport, and changes to the hydrophilic spacer can have a large impact on transport even if they are far from the FSFG repeat. These findings provide useful methodologies and tools for improving the design and engineering of NLPs to tune their transport properties for different applications in the purification of biomolecules.

Acknowledgements

This work was supported by the National Science Foundation and Defense Threat Reduction Agency under contracts 1705923 and HDTRA1-13-1-0038, respectively. We thank Minkyu Kim (Arizona State University) for sequence design.

Author Contributions

Y.J.Y, S.L., and B.D.O designed the experiments. Y.J.Y., S.L., and D.J.M. performed the experiments. Y.J.Y., D.J.M., M.A.M., and B.D.O. analyzed the data. Y.J.Y. and B.D.O. wrote the manuscript. B.D.O. is the principal investigator.

Author information

Corresponding Author

*E-mail: bdolsen@mit.edu

ORCID

Yun Jung Yang: 0000-0003-0684-5218

Danielle J. Mai: 0000-0001-5447-2845

Melody A. Morris: 0000-0001-5597-154X

Bradley D. Olsen: 0000-0002-7272-7140

Conflict of Interest

The authors declare no competing financial interest.

Supporting Information

Plasmid map of P-NLP-Ps, table of library of P-NLP-Ps, additional rheology data, additional capillary transport assay data and quantification, protein sequences, data processing code.

References

- (1) Vepari, C.; Kaplan, D. L. Silk as a Biomaterial. *Prog. Polym. Sci.* **2007**, *32*, 991–1007.
- (2) Bracalello, A.; Santopietro, V.; Vassalli, M.; Marletta, G.; Del Gaudio, R.; Bochicchio, B.; Pepe, A. Design and Production of a Chimeric Resilin-, Elastin-, and Collagen-Like Engineered Polypeptide. *Biomacromolecules* **2011**, *12*, 2957–2965.
- (3) Hyun, J.; Lee W. K.; Nath, N.; Chilkoti, A.; Zauscher, S. Capture and Release of Proteins on the Nanoscale by Stimuli-Responsive Elastin-Like Polypeptide “Switches.” *J. Am. Chem. Soc.* **2004**, *126*, 7330–7335.
- (4) Humphries, J.; Byron, A.; Humphries, M. J. Integrin ligands at a glance. *J Cell Sci.* **2006**, *119*, 3901-3903.
- (5) DiMarco, R. L.; Heilshorn, S. C. Multifunctional Materials through Modular Protein Engineering.

- Adv. Mater.* **2012**, *24*, 3923–3940.
- (6) Porebski, B. T.; Buckle, A. M. Consensus Protein Design. *Protein Eng. Des. Sel.* **2016**, *29*, 245–251.
 - (7) Yang, Y. J.; Holmberg, A. L.; Olsen, B. D. Artificially Engineered Protein Polymers. *Annu. Rev. Chem. Biomol. Eng.* **2017**, *8*, 549–575.
 - (8) Main, E. R. G.; Jackson, S. E.; Regan, L. The Folding and Design of Repeat Proteins: Reaching a Consensus. *Curr. Opin. Struct. Biol.* **2003**, *13*, 482–489.
 - (9) Ma, L.; Yamada, S.; Wirtz, D.; Coulombe, P. A. A “hot-Spot” Mutation Alters the Mechanical Properties of Keratin Filament Networks. *Nat. Cell Biol.* **2001**, *3*, 503–506.
 - (10) Ribeiro, A.; Arias, F. J.; Reguera, J.; Alonso, M.; Rodríguez-Cabello, J. C. Influence of the Amino-Acid Sequence on the Inverse Temperature Transition of Elastin-like Polymers. *Biophys. J.* **2009**, *97* (1), 312–320. <https://doi.org/10.1016/j.bpj.2009.03.030>.
 - (11) Almaaytah, A.; Ajingi, Y.; Abualhaijaa, A.; Tarazi, S.; Alshar'i, N.; Al-Balas, Q. Peptide Consensus Sequence Determination for the Enhancement of the Antimicrobial Activity and Selectivity of Antimicrobial Peptides. *Infect. Drug Resist.* **2017**, *10*, 1–17.
 - (12) Parmeggiani, F.; Pellarin, R.; Larsen, A. P.; Varadamsetty, G.; Stumpp, M. T.; Zerbe, O.; Caflisch, A.; Plückthun, A. Designed Armadillo Repeat Proteins as General Peptide-Binding Scaffolds: Consensus Design and Computational Optimization of the Hydrophobic Core. *J. Mol. Biol.* **2008**, *376*, 1282–1304.
 - (13) Knappik, A.; Ge, L.; Honegger, A.; Pack, P.; Fischer, M.; Wellenhofer, G.; Hoess, A.; Wolle, J.; Pluckthun, A.; Virnekas, B. Fully Synthetic Human Combinatorial Antibody Libraries (HuCAL) Based on Modular Consensus Frameworks and CDRs Randomized with Trinucleotides. *J. Mol. Biol.* **2000**, *296*, 57–86.
 - (14) Fosgerau, K.; Hoffmann, T. Peptide Therapeutics: Current Status and Future Directions. *Drug Discov. Today* **2015**, *20*, 122–128.

- (15) Schein, C. H.; Bowen, D. M.; Lewis, J. A.; Choi, K.; Paul, A.; van der Heden van Noort, G. J.; Lu, W.; Filippov, D. V. Physicochemical Property Consensus Sequences for Functional Analysis, Design of Multivalent Antigens and Targeted Antivirals. *BMC Bioinformatics* **2012**, *13*, S9.
- (16) Yang, Y. J.; Mai, D. J.; Dursch, T. J.; Olsen, B. D. Nucleopore-Inspired Polymer Hydrogels for Selective Biomolecular Transport. *Biomacromolecules* **2018**, *19*, 3905–3916.
- (17) Kim, M.; Chen, W. G.; Souza, B. S.; Olsen, B. D. Selective Biomolecular Separation System Inspired by the Nuclear Pore Complex and Nuclear Transport. *Mol. Syst. Des. Eng.* **2017**, *2*, 149–158.
- (18) Caspi, Y.; Zbaida, D.; Cohen, H.; Elbaum, M. Synthetic Mimic of Selective Transport Through the Nuclear Pore Complex. *Nano Lett.* **2008**, *8*, 3728–3734.
- (19) Lieleg, O.; Baumgärtel, R. M.; Bausch, A. R. Selective Filtering of Particles by the Extracellular Matrix: An Electrostatic Bandpass. *Biophys. J.* **2009**, *97*, 1569–1577.
- (20) Kim, M.; Chen, W. G.; Kang, J. W.; Glassman, M. J.; Ribbeck, K.; Olsen, B. D. Artificially Engineered Protein Hydrogels Adapted from the Nucleoporin Nsp1 for Selective Biomolecular Transport. *Adv. Mater.* **2015**, *27*, 4207–4212.
- (21) Petrou, G.; Crouzier, T. Mucins as Multifunctional Building Blocks of Biomaterials. *Biomater. Sci.* **2018**, *6*, 2282–2297.
- (22) Lieleg, O.; Ribbeck, K. Biological Hydrogels as Selective Diffusion Barriers. *Trends Cell Biol.* **2011**, *21*, 543–551.
- (23) D'Angelo, M. A.; Hetzer, M. W. Structure, Dynamics and Function of Nuclear Pore Complexes. *Trends Cell Biol.* **2008**, *18*, 456–466.
- (24) Ribbeck, K.; Görlich, D. The Permeability Barrier of Nuclear Pore Complexes Appears to Operate via Hydrophobic Exclusion. *EMBO J.* **2002**, *21*, 2664–2671.
- (25) Clarkson, W. D.; Kent, H. M.; Stewart, M. Separate Binding Sites on Nuclear Transport Factor 2

- (NTF2) for GDP-Ran and the Phenylalanine-Rich Repeat Regions of Nucleoporins P62 and Nsp1p. *J. Mol. Biol.* **1996**, *263*, 517–524.
- (26) Rout, M. P.; Wente, S. R. Pores for Thought: Nuclear Pore Complex Proteins. *Trends Cell Biol.* **1994**, *4*, 357–365.
 - (27) Powers, M. A.; Forbes, D. J. Nuclear Transport: Beginning to Gel? *Curr. Biol.* **2012**, *22*, R1006–R1009.
 - (28) Denning, D. P.; Patel, S. S.; Uversky, V.; Fink, A. L.; Rexach, M. Disorder in the Nuclear Pore Complex: The FG Repeat Regions of Nucleoporins Are Natively Unfolded. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 2450–2455.
 - (29) Chen, W. G.; Witten, J.; Grindy, S. C.; Holten-Andersen, N.; Ribbeck, K. Charge Influences Substrate Recognition and Self-Assembly of Hydrophobic FG Sequences. *Biophys. J.* **2016**, *113*, 2088–2099.
 - (30) Rout, M. P.; Aitchison, J. D.; Suprpto, A.; Hjertaas, K.; Zhao, Y.; Chait, B. T. The Yeast Nuclear Pore Complex. *J. Cell Biol.* **2000**, *148*, 635–652.
 - (31) Terry, L. J.; Wente, S. R. Flexible Gates: Dynamic Topologies and Functions for FG Nucleoporins in Nucleocytoplasmic Transport. *Eukaryot. Cell* **2009**, *8*, 1814–1827.
 - (32) Frey, S.; Görlich, D. FG/FxFG as Well as GLFG Repeats Form a Selective Permeability Barrier with Self-Healing Properties. *EMBO J.* **2009**, *28*, 2554–2567.
 - (33) Olsen, B. D.; Kornfield, J. A.; Tirrell, D. A. Yielding Behavior in Injectable Hydrogels from Telechelic Proteins. *Macromolecules* **2010**, *43*, 9094–9099.
 - (34) Protein Calculator v3.4, <http://protcalc.sourceforge.net/>, March 28, 2020.
 - (35) Colwell, L. J.; Brenner, M. P.; Ribbeck, K. Charge as a Selection Criterion for Translocation through the Nuclear Pore Complex. *PLoS Comput. Biol.* **2010**, *6*, e1000747.
 - (36) Li, L. D.; Crouzier, T.; Sarkar, A.; Dunphy, L.; Han, J.; Ribbeck, K. Spatial Configuration and

- Composition of Charge Modulates Transport into a Mucin Hydrogel Barrier. *Biophys J.* **2013**, *17*, 1357-1365.
- (37) Petka, W. A.; Harden, J. L.; McGrath, K. P.; Wirtz, D.; Tirrell, D. A. Reversible Hydrogels from Self-Assembling Artificial Proteins. *Science* **1998**, *281*, 389–392.
 - (38) Ader, C.; Frey, S.; Maas, W.; Schmidt, H. B.; Görlich, D.; Baldus, M. Amyloid-like Interactions within Nucleoporin FG Hydrogels. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 6281–6285.
 - (39) Cole, D. G.; Saxton, W. M.; Sheehan, K. B.; Scholey, J. M. A “Slow” Homotetrameric Kinesin-Related Motor Protein Purified from *Drosophila* Embryos. *J. Biol. Chem.* **1994**, *269*, 22913–22916.
 - (40) Hawe, A.; Hulse, W. L.; Jiskoot, W.; Forbes, R. T. Taylor Dispersion Analysis Compared to Dynamic Light Scattering for the Size Analysis of Therapeutic Peptides and Proteins and Their Aggregates. *Pharm. Res.* **2011**, *28*, 2302–2310.
 - (41) Ribbeck, K.; Görlich, D. Kinetic Analysis of Translocation through Nuclear Pore Complexes. *EMBO J.* **2001**, *20*, 1320–1330
 - (42) Chaillan-Huntington, C.; Butler, P. J. G.; Huntington, J. A.; Akin, D.; Feldherr, C.; Stewart, M. NTF2 Monomer-Dimer Equilibrium. *J. Mol. Biol.* **2001**, *314*, 465–477.
 - (43) Ben-Efraim, I.; Gerace, L. Gradient of Increasing Affinity of Importin Beta for Nucleoporins along the Pathway of Nuclear Import. *J. Cell Biol.* **2001**, *152*, 411–417.
 - (44) Hayama, R.; Sparks, S.; Hecht, L. M.; Dutta, K.; Karp, J. M.; Cabana, C. M.; Rout, M. P.; Cowburn, D. Thermodynamic Characterization of the Multivalent Interactions Underlying Rapid and Selective Translocation through the Nuclear Pore Complex. *J. Biol. Chem.* **2018**, *293*, 4555–4563.
 - (45) Marsh, J.A.; Forman-Kay, J.D. Sequence determinants of compaction in intrinsically disordered proteins. *Biophys J.* **2010**, *98*, 2383-2390.
 - (46) Frey, S.; Richter, R.P.; Görlich, D. FG-rich repeats of nuclear pore proteins form a three-dimensional

meshwork with hydrogel-like properties. *Science*. **2006**, 3, 815-817.

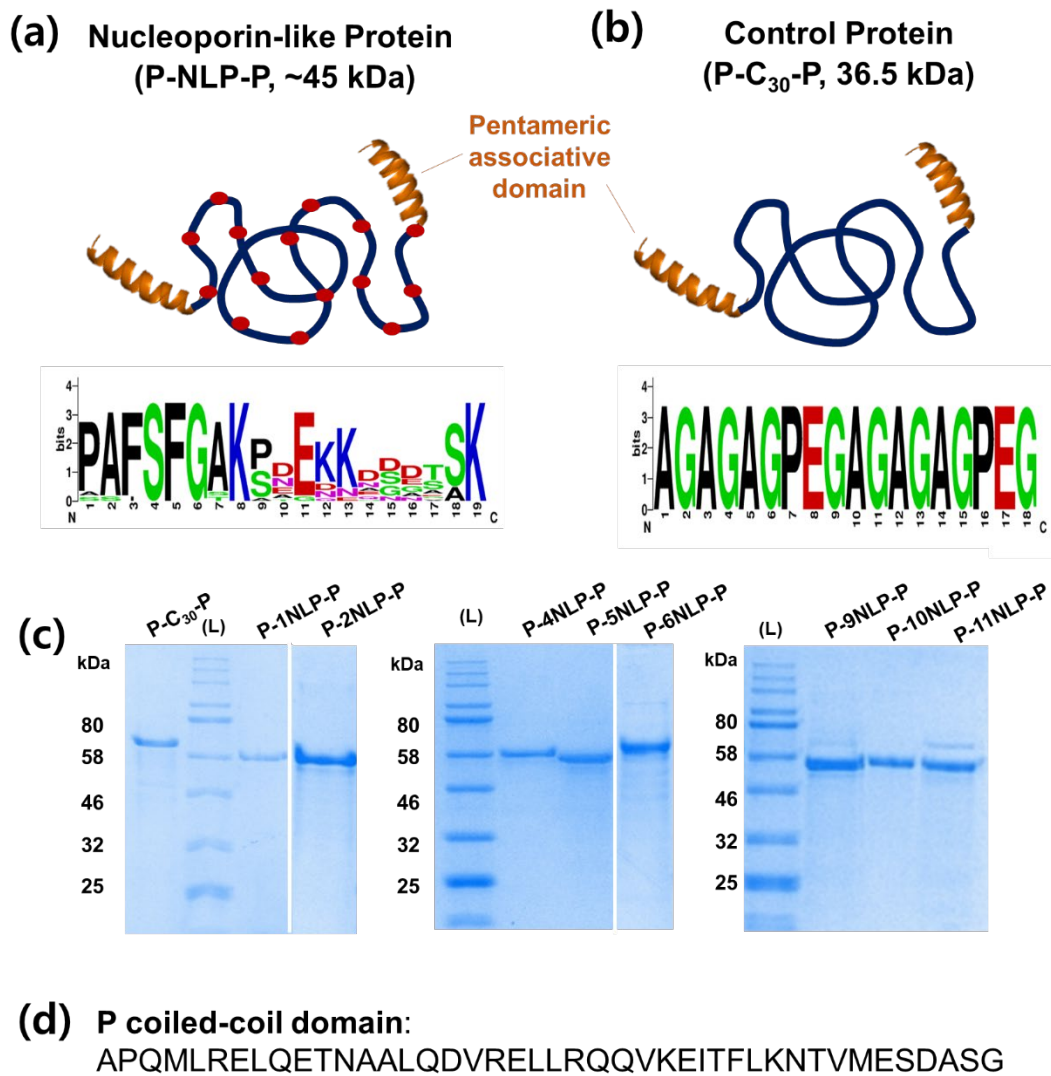


Figure 1. Sequence design of tri-block a) recombinant nucleoporin-like proteins (P-NLP-P) and b) control protein (P-C₃₀-P). The colors in sequence logo represent amino acids with hydrophobic side chains (black), polar side chain (green), negatively charged side chain (red), and positively charged side chain (blue). c) The purified proteins visualized by Coomassie brilliant blue staining. d) P coiled-coil domain sequence.

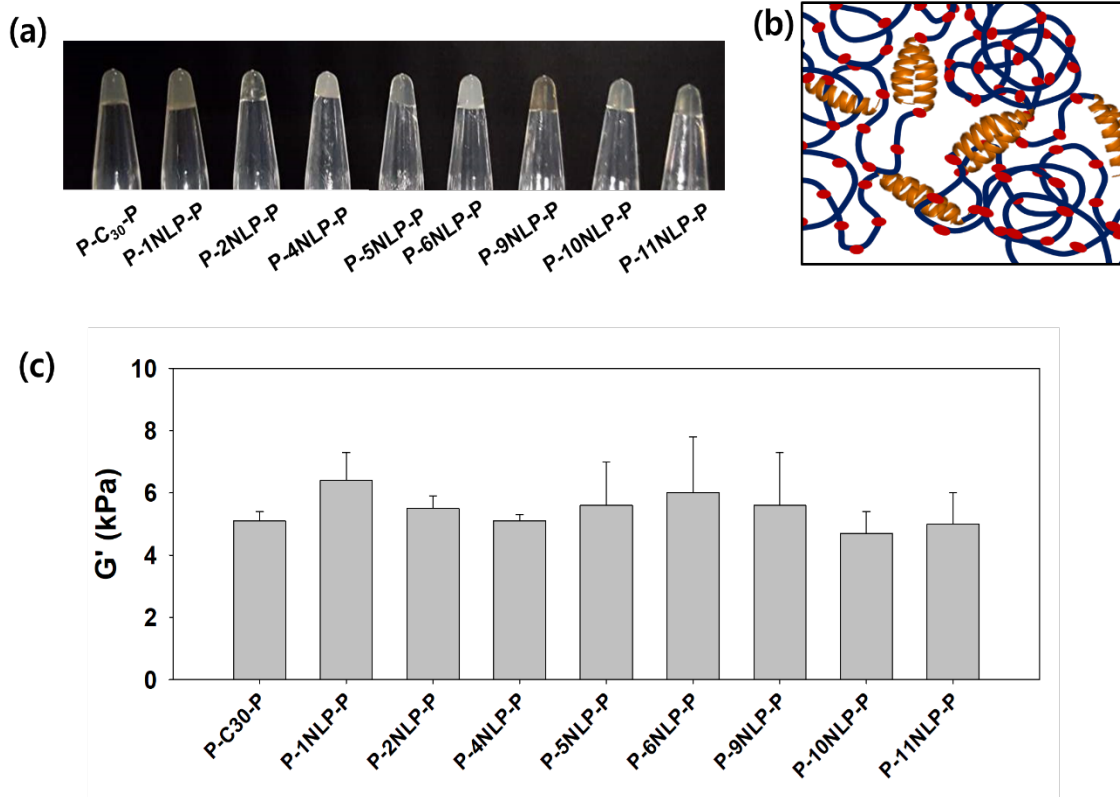
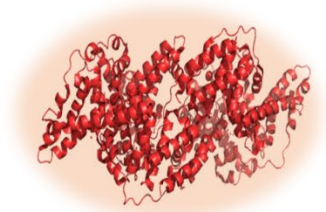
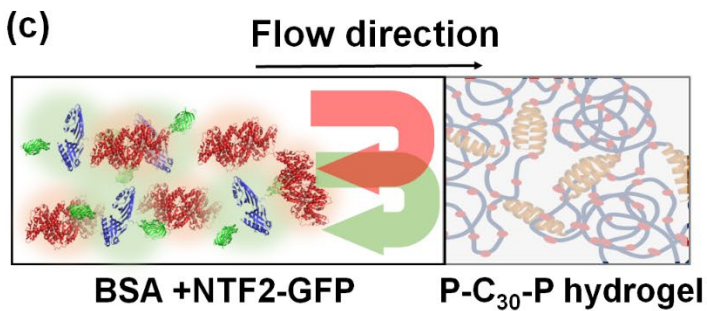


Figure 2. a) Physically crosslinked P-C₃₀-P and NLP hydrogels. b) Illustration of gel networks. Orange coils indicate physically associated P domains. c) Average plateau modulus of P-C₃₀-P and P-NLPs-P hydrogels.

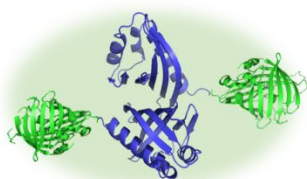
(a) BSA (~66 kDa)



(c)



(b) NTF2-GFP (~86 kDa)



(d)

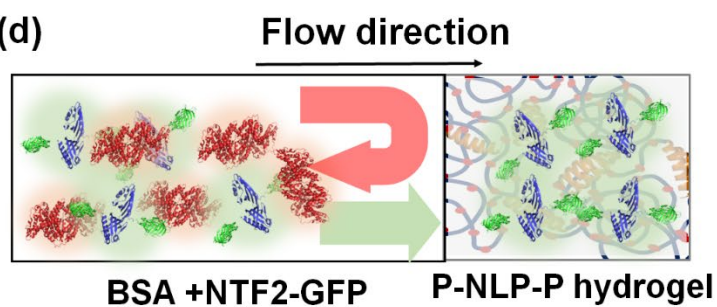
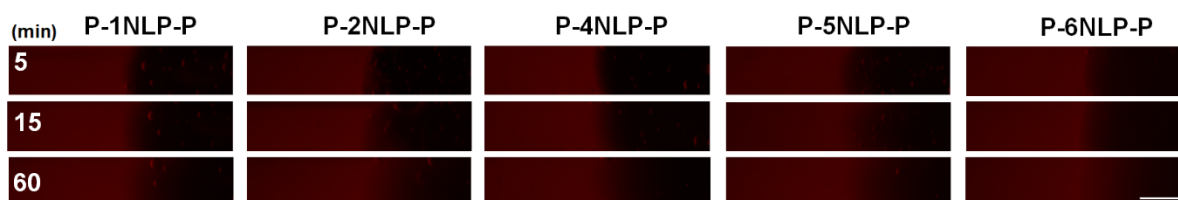


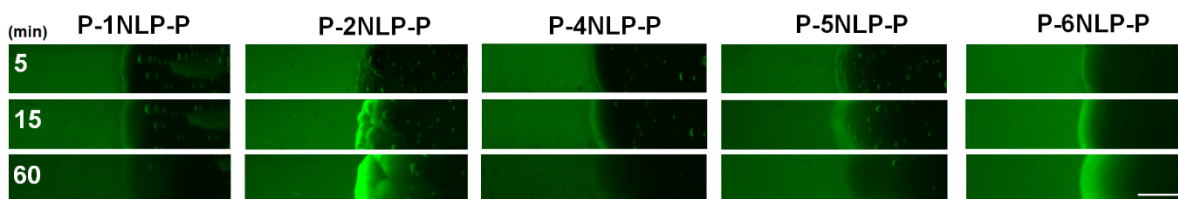
Figure 3. 3D structure of a) BSA and b) NTF2-GFP. An illustration of 1D transport assay for observing diffusion of molecules. c) P-C₃₀-P hydrogel rejects all molecules, while d) NLP hydrogel only allows NTF2-GFP.

Electrostatics of Conserved Residues

(a) BSA



(b) NTF2-GFP



(c) Selective Permeability of NTF2-GFP against BSA

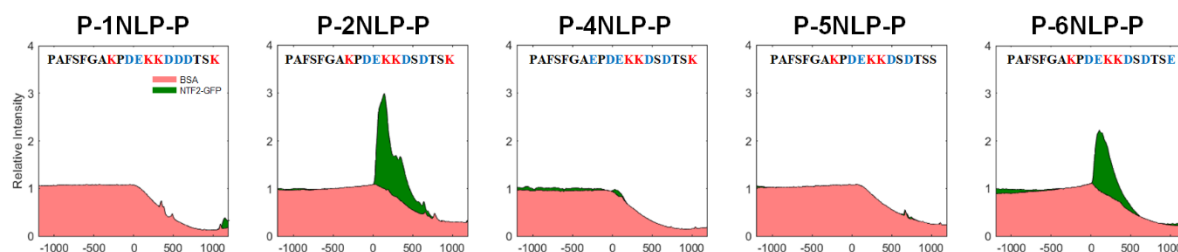
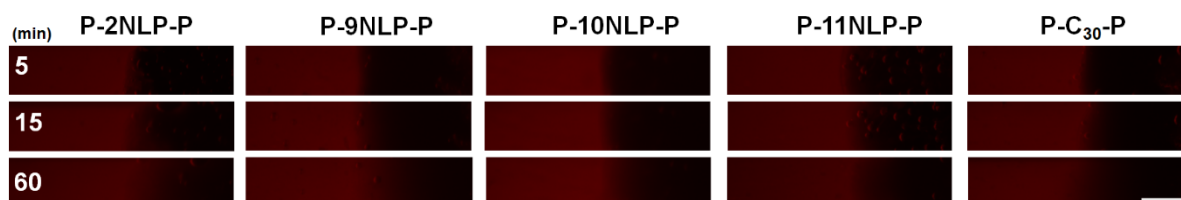


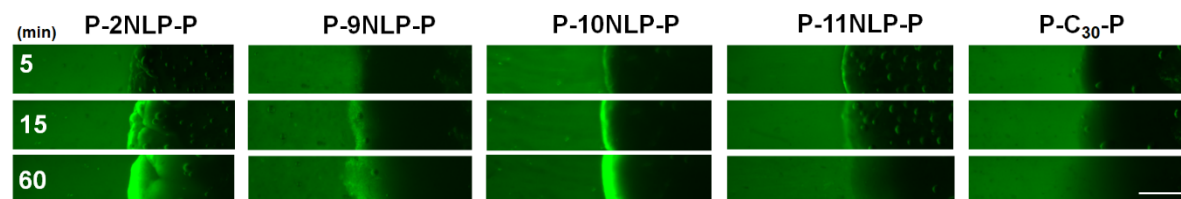
Figure 4. Images of a) BSA molecule diffusion and b) NTF2-GFP molecule diffusion across the buffer–hydrogel interface were acquired at 5, 15, and 60 min time points for the charge sequence. c) Plots representing fluorescence intensity of NTF2-GFP and BSA at 60 min time point. All scale bars are 500 μm .

Polarity of Conserved Residues

(a) *BSA*



(b) *NTF2-GFP*



(c) *Selective Permeability of NTF2-GFP against BSA*

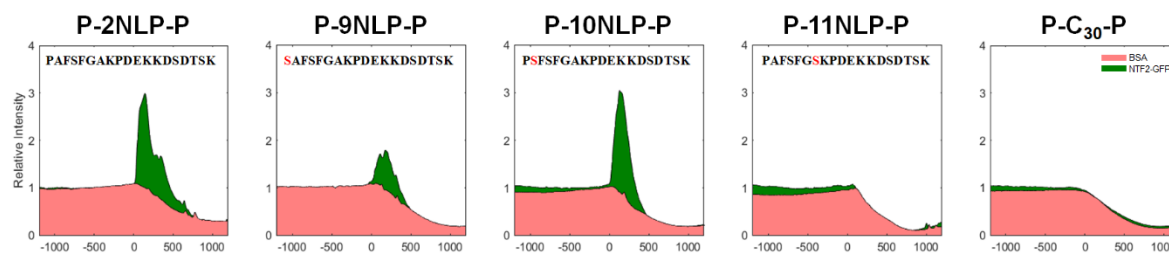


Figure 5. Images of a) BSA molecule diffusion and b) NTF2-GFP molecule diffusion across the buffer–hydrogel interface were acquired at 5, 15, and 60 min time points for the polarity sequence. c) Plots representing fluorescence intensity of NTF2-GFP and BSA at 60 min time point. All scale bars are 500 μ m.

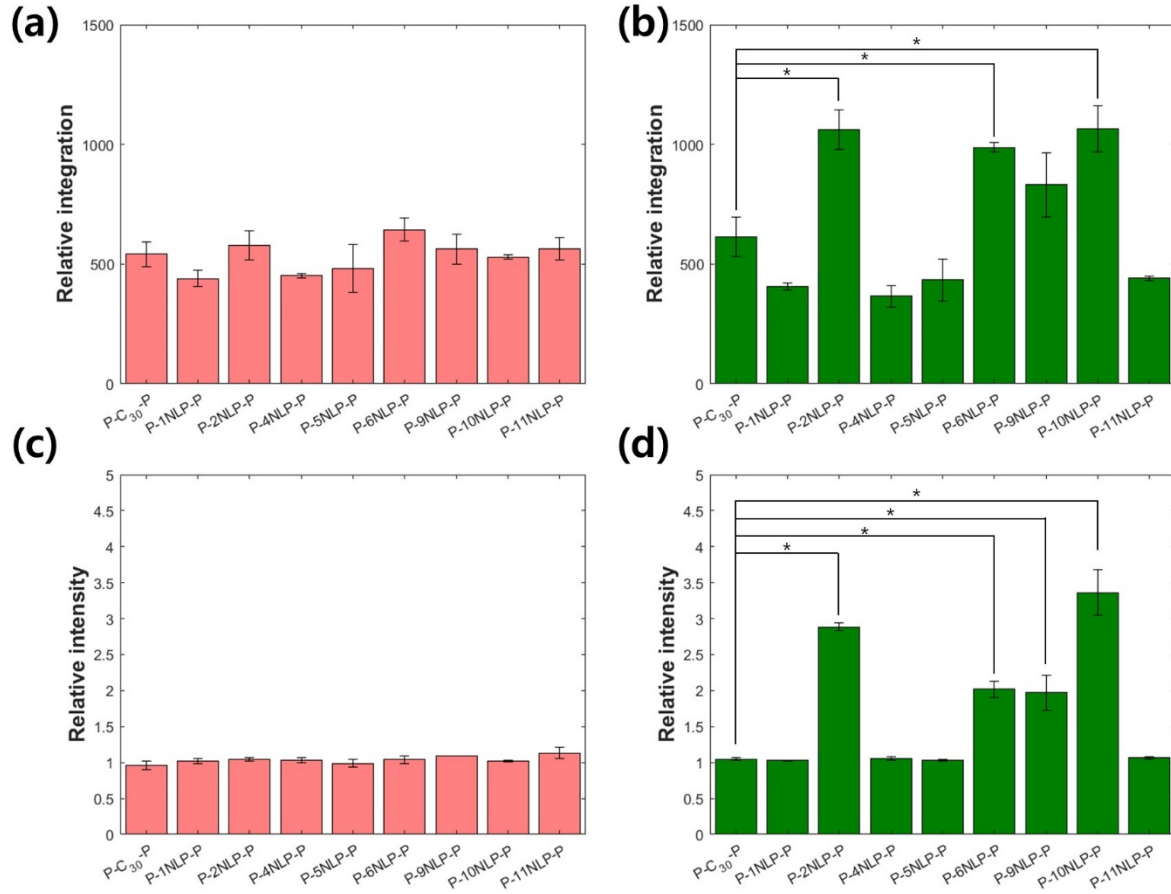


Figure 6. The relative integration (a,b) and maximum intensity (c,d) of accumulated a,c) BSA and b,d) NTF2-GFP in the gel at 60 min. The values were averaged from three different 1D transport assay batches, and the error bar shows standard deviation. (* indicates $p < 0.05$)

Table 1. Information of site-mutated nucleoporin-like proteins and P-C₃₀-P (negative control)

* Amino acid in red is the site-mutated one

Table 1a.

<i>Electrostatics of Conserved Residues</i>		pI	M.W.	Charge (pH 7.5)
P-1NLP-P	PAFSFGAKPDEKKDDDTSK	5.10	45.88	-21.8
P-2NLP-P	PAFSFGAKPDEKKDSDTSK	6.07	45.43	-5.8
P-3NLP-P	PAFSFGASPDEKKDSDTSK	5.00	44.77	-21.8
P-4NLP-P	PAFSFGAEPDEKKDSDTSK	4.71	45.45	-37.8
P-5NLP-P	PAFSFGAKPDEKKDSDTSS	5.00	44.77	-21.8
P-6NLP-P	PAFSFGAKPDEKKDSDTSE	4.71	45.45	-37.8
P-12NLP-P	PAFSFGAKPKKKKKSKTSK	10.98	46.04	121.9
P-13NLP-P	PAFSFGAKPDEDEDSDTSK	4.28	45.24	-69.7

Table 1b.

<i>Polarity of Conserved Residues</i>		pI	M.W.	Charge (pH 7.5)
P-2NLP-P	PAFSFGAKPDEKKDSDTSK	6.07	45.43	-5.8
P-9NLP-P	SAFSFGAKPDEKKDSDTSK	6.07	45.27	-5.8
P-10NLP-P	PSFSFGAKPDEKKDSDTSK	6.07	45.69	-5.8
P-11NLP-P	PAFSFGSKPDEKKDSDTSK	6.07	45.69	-5.8

Table 1c.

<i>Polarity of Conserved Residues</i>		pI	M.W.	Charge (pH 7.5)
P-7NLP-P	PAYSFGAKPDEKKDSDTSK	6.07	45.69	-5.9
P-8NLP-P	PAYSYGAKPDEKKDSDTSK	6.07	45.94	-5.9

Table 1d.

non-NLP control, P-C ₃₀ -P protein		pI	M.W.	Charge (pH 7.5)
P-C ₃₀ -P	AGAGAGPEGAGAGPEG	4.38	36.56	-30.7

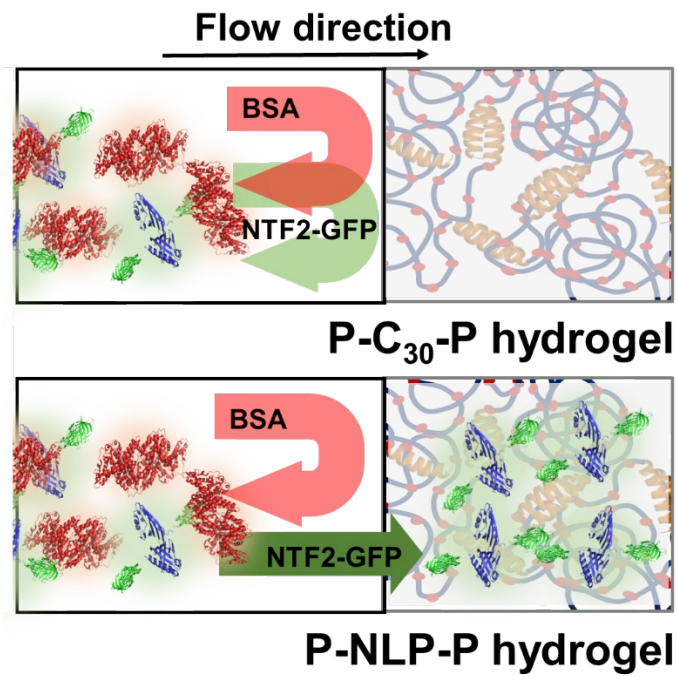


Table of Contents graphic