Charge influences substrate recognition and self-assembly of hydrophobic FG sequences

Wesley G. Chen[1], Jacob Witten[1,2], Scott C. Grindy[3], Niels Holten-Andersen[3], Katharina Ribbeck[1,a]

[1]Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139

[2]Computational and Systems Biology Initiative, Massachusetts Institute of Technology, Cambridge, MA 02139

[3]Department of Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139

[a]Corresponding author:

Katharina Ribbeck

Email: ribbeck@mit.edu

Telephone: +1 (617) 715-4575

MIT, Department of Biological Engineering

77 Massachusetts Avenue

Rm. 56-341

Cambridge, MA 02139

USA

Abstract:

The nuclear pore complex controls the passage of molecules via hydrophobic phenylalanine-glycine (FG) domains on nucleoporins. Such FG-domains consist of repeating units of FxFG, FG, or GLFG sequences, which can be interspersed with highly charged amino acid sequences. Despite the high density of charge exhibited in certain FG-domains, if and how charge influences FG-domain self-assembly and selective binding of nuclear transport receptors is largely unexplored. Studying how individual charged amino acids contribute to nuclear pore selectivity is challenging with modern *in vivo* and *in vitro* techniques due to the complexity of nucleoporin sequences. Here, we present a rationally designed approach to deconstruct essential components of nucleoporins down to 14 amino acid sequences. With these nucleoporin-based peptides, we systematically dissect how charge type and placement of charge influences self-assembly and selective binding of FG-containing gels. Specifically, we find that charge type determines which hydrophobic substrates FG sequences recognize while spatial localization of charge tunes hydrophobic self-assembly and receptor selectivity of FG sequences.

Introduction

The nuclear pore complex (NPC) is a megadalton structure that controls the exchange of material between the nucleus and cytoplasm (1-3) through a combination of passive and facilitated diffusion. Above a ~30 kDa cutoff, proteins require complexation with nuclear transport receptors (NTRs) to efficiently translocate at rates of nearly one thousand molecules per second (4-7). This fast translocation rate relies on transient interactions between hydrophobic phenylalanine-glycine (FG) domains on intrinsically disordered FG-containing nucleoporins (FG-nups) and hydrophobic patches on NTRs (1,5,8-20). Without such assistance of NTRs, proteins remain excluded from the NPC. Despite the necessity of FG-domains (which contain repeating units of FG sequences such as FxFG or GLFG) for facilitated diffusion and self-assembly of the selective matrix, how hydrophobic FG-domains exhibit such selectivity to specific hydrophobic domains on NTRs and what parameters can tune such molecular recognition necessary for facilitated transport remain open questions. Sequence analysis and molecular dynamic simulations predict that the biochemistry and charge surrounding individual FG sequences should play an essential role in FG-mediated molecular recognition and therefore determine the organization and selectivity of the nuclear pore (4,6,21-24). However, due to the complexity and redundancy of FG-nups within NPCs, a systematic biochemical dissection of the amino acid space surrounding individual FG sequences has remained an experimental challenge.

In this study, we investigate if and how electrostatic interactions surrounding FG sequences tailor both the self-assembly of FG sequences and the selective recognition of hydrophobic substrates. For systematic dissection of how charge type and localization influence FG-mediated selectivity, full FG-domains are too degenerate and complex to provide insight into how single amino acids

contribute to selectivity. To overcome this challenge, we established a novel approach that uses short peptides. The use of peptides and polypeptides has previously been essential in design of other minimalistic polymer-based biomaterials such as extracellular matrices (25), silk proteins (26,27), and elastin-like polypeptides (28-30). While peptides may not recapitulate all properties of the original protein, they allow for mechanistic dissection of structure and function at the single amino acid level. Here, peptides can be designed to represent the essential components of charged FG-nups, allow for single amino acid substitutions, and can be engineered to self-assemble into hydrogels mimicking the high density of FG-domains found in certain FG-nups. With this peptide system, we show how the amino acids surrounding FG sequences has a significant impact on hydrophobic interactions necessary for molecular recognition of substrates and self-assembly. Choice of anionic and cationic residues is able to reverse the selective recognition of the FG-based gels. This allows for FG sequences to discriminate between substrates exhibiting various types of hydrophobic domains using a combination of hydrophobic and electrostatic interactions. Moreover, spatial localization of charge determines the degree to which electrostatic interactions influence hydrophobic interactions with FG sequences. Lastly, we find that the choice of charged residues such as lysine, arginine, glutamic acid, and aspartic acid have different contributions to the self-assembly of FG sequences while electrostatic properties dominate for molecular recognition.

Results

As a model FG-nup to base our FG-domain mimicking peptides on, we chose the yeast nucleoporin Nsp1 as it is essential in *Saccharomyces cerevisieae* (31) and contains a repeating subsequence (284-553) with a high density of charge (Figure 1A). Sequence analysis of 15-

repeats of FG-domains in Nsp1$^{284-553}$ reveals a non-uniform distribution of charge in the sequence space separating FG sequences (here, specifically, we use the FSFG sequence). Cationic residues can be found near the center and edges of the repeat whereas anionic amino acids reside at least three positions away from FSFG sequences (Figure 1B). Moreover, we find several highly conserved lysines (K) situated two to three amino acids away from the FG-domain at positions 6 and 17 and a conserved glutamic acid (E) at position 9 (Figure 1A). The non-uniformity in conserved charge distributions and amino acid identity suggest that charge and biochemical properties of amino acid sidechains may play an essential role in governing how neighboring FG sequences respond to environmental substrates and that significant function can be encoded in the non self-assembling domains of FG-nups.

For testing each of the parameters of charge type, placement, and sidechain biochemistry and how they affect FG function, we synthesized the consensus Nsp1 peptide sequence and also a simplified 14 amino acid version. The deconstructed peptides consist of two terminal FSFG sequences with neighboring charged or neutral amino acids. To establish how the presence of charge affects FG-mediated selectivity, we designed the peptide sequence FSFGAXAAXAFSFG, where X is either a lysine (K), glutamic acid (E) or a neutral serine (S) as seen in Figure 1C. Since lysine overall is the most highly conserved residue in this wildtype Nsp1 subsequence, we use the lysine-containing peptide as the reference for all other experimental comparisons (termed FGAK peptide). To determine if the selectivity properties of FG-domains could be reversed by switching neighboring charge types, we synthesized the anionic peptide FGAE. As a control of how the presence of charge influences FG function, we synthesized the neutral FGAS-peptide in which the lysines have been converted to serines. To

test how the spatial localization of lysines affect FG function, we designed two variants of FGAK in which the lysines have been placed directly adjacent to the FG-domain or separated by two alanines (termed FGK and FGA_2K, respectively, Figure 1D). Last, it appears that only certain charged amino acids such as lysines and glutamic acids (E) exhibit 100% conservation at positions 6, 9, and 17 (Figure 1A), suggesting amino acid biochemistries may also regulate FG-based molecular recognition. To test how chemical structures of the amino acid sidechains affects structure and function of FG-domains, we designed a third class of peptides in which the lysines are replaced with cationic arginines (R, FGAR) or aspartic acids (D, FGAD) as seen in Figure 1E. With these 14-amino acid nucleoporin-based peptides, we now analyze how single amino acid substitutions can alter the selective binding and self-assembling capabilities of individual FG-domains.

Characterization of charge presence on FG-mediated self-assembly

Before studying the selective recognition of the FG-based peptides, we first establish if the peptides are capable of forming gels at a concentration relevant to intact nuclear pore complexes. By measuring the mechanical properties of the peptide solutions, we can further understand how charge affects FG-domain interactions. We used a concentration of 2% (w/v) for each of the peptides, which corresponds to 28 mM concentrations of FG sequences, a value that is well within estimations of FG sequences found in densely packed NPCs (18,32,33). To quantify gelation, we measure the stiffness of the resulting material using small amplitude oscillatory frequency sweeps. We report the storage (G') and loss (G") modulus of the peptide solution and specify that a gel forms when G' is greater than G", which indicates successful self-assembly of a stable network of peptides. The consensus peptide sequence of Nsp1 was unable to form a gel

and flowed when inverted, so it was not suitable for further analysis. For the simplified peptides, we first establish that the reference peptide FGAK is able to form hydrogels with a stiffness of 10^4 Pa across the frequencies tested (Figure 2A). To ensure that the FG sequences are responsible for the self-assembly process, we converted the hydrophobic phenylalanines to serines (SGAK). With this substitution, the peptide solution remains in an aqueous phase and no longer exhibits a dominant storage modulus (Supplementary Figure 1), suggesting that phenylalanine provides the necessary hydrophobic interactions for self-assembly. Reversing the charge to glutamic acids (FGAE peptides) reveals that the gel-forming properties are maintained with stiffness between 5-9 KPa at the frequencies tested (Figure 2B). To determine if the charge is responsible for maintaining the hydrated state of the gel in the self-assembly process, we compared the material to the solution of neutral FGAS peptides. Without the charge, FGAS precipitates out of solution (data not shown). Together, these results indicate that the presence of charge is essential in maintaining a hydrated network of FG sequences in self-assembly. However, too much charge increases the solubility and prevents gelation, as seen with the consensus sequence peptide. Without the presence of electrostatic repulsion within the gels, the network collapses and forms a precipitate. Moreover, for this class of designed peptides, FG sequences are essential for gelation, which is the same as for intact FG-nups (9,10).

Transport of NTF2 is asymmetric in cationic and anionic FG-based gels

To understand if the FG-based peptide gels are able to recapitulate the selective recognition properties of the native FG-nups, we test for their ability to enrich and select for intact nuclear transport receptors (NTRs) in an FG-dependent manner. As a model NTR, we choose NTF2 as it contains an essential tryptophan (W7) that is required for binding to the FG-domains. By

159 substituting the tryptophan for an alanine (W7A) to ablate FG-binding capabilities(14), we can
160 determine if the FG-domains are available in the gel for NTR binding. First, we determine if the
161 positively charged FGAK gel is able to preferentially select for NTF2 compared to the W7A
162 mutant. To prepare gels for selective transport assays, we dissolve FGAK peptides at 2% (w/v)
163 as before and load the material into capillaries. Fluorescently labeled NTF2 or W7A is then
164 injected into the capillary and sealed to create a 1-D diffusion chamber where the diffusion
165 profile is monitored for up to 5 hours at one-minute intervals (Figure 2C). To quantify the
166 selective binding of the various transport receptors into gels, we estimate the effective diffusion
167 coefficient of the receptors in the gel using the time-evolving fluorescent profiles and the
168 maximum accumulation of the reporters at the interface as metrics for determining how well the
169 gel is able to differentiate between native NTF2 and W7A. Lower effective diffusivities and
170 higher accumulation indicate stronger interactions with the gel.
171
172 As seen in Figure 2D and F, the FGAK gel accumulates NTF2 at higher concentrations at the gel
173 interface compared to the W7A mutant. While accumulation at the interface is different, we find
174 the effective diffusivities to be similar (Figure 2F), suggesting that the favorable hydrophobic
175 interactions are weak and reversible, which is consistent with previous models of nuclear pore
176 transport. These experiments show that FGAK has a selective binding preference for NTF2 over
177 the W7A mutant, indicating that FG-domains are available and functional as a binding domain. It
178 also appears that the short peptide gels cannot recapitulate the unique facilitated diffusion
179 properties of the intact nuclear pore. For instance, native NTRs that are overall anionic and
180 hydrophobic are predicted to bind to and transport into overall cationic nucleoporin gels (21,22).
181 Since the peptides form fibers and sheets (Figure 2A-B) that are not observed in native

nucleoporin assemblies (24), they likely do not form the requisite saturated disordered matrix that allows for exclusion of inert molecules that typically cannot transport through NPCs (9). Nevertheless, the designed FGAK peptide has the sensitivity to recognize single hydrophobic amino acid differences between NTF2 and W7A, which is a necessary first step for selective hydrophobic transport.

To test if binding to NTRs is sensitive to charged residues neighboring the FG domain, we determine if FGAE is able to recapitulate the same selective properties of FGAK, despite the reversal of charge. As seen in Figure 2E-F, FGAE selects for NTF2 and W7A equally at the interface and there is no significant difference in diffusion coefficients. These results indicate FGAE is unable to differentiate between the native and mutant form of NTF2. Since FGAK and FGAE exhibit different selective properties for NTF2 and W7A, we further investigate if the asymmetry results from a structural issue or a charge based selective recognition. Figures 2A-B show that FGAK and FGAE gels exhibit structural differences at the microscopic level (sheets vs. fiber assembly) and macroscopic level (inset). It is possible that due to structural differences, the FG sequences are not exposed and unavailable for binding by NTF2 in the FGAE gel. To test if hydrophobic domains are accessible, we use Nile Red dye, which fluoresces within hydrophobic environments. Supplementary Figure 2 shows that Nile Red fluorescence is detected in both FGAK and FGAE gels, indicating that the uncharged hydrophobic dye is able to interact with the FG-domains within the peptide-based gels despite the structural variation. These results suggest that the lysine and glutamic acid are important in regulating the selective properties of the FG sequences from a charge interaction perspective, as opposed to simply altering the gel's microstructure.

205

206 **Presence of charge regulates the selective recognition of FG-domains**

207 Although NTRs such as Importinβ and NTF2 have well-characterized specificity for particular
208 FG-nup sequences, intact receptors contain multiple binding pockets with varying affinities and
209 charge distributions, or require dimerization for function, which complicates systematic analysis
210 of charge contribution. Hence, as a replacement for complex NTRs, we designed fluorescent
211 peptide reporters with defined spatial arrangements of charged and non-polar amino acids to
212 systematically test the contribution of charge in hydrophobic selectivity (Figure 3B). The first
213 two reporter peptides contain three phenylalanines for a hydrophobic tail that enable binding to
214 the FG sequences, but one contains an adjacent anionic sequence composed of glutamic acids
215 (termed Hydrophobic (-)), while the other contains cationic lysines (Hydrophobic (+)). As
216 controls, we engineered two more reporters where phenylalanines were converted to hydrophilic
217 asparagines (N), and termed Hydrophilic (- or +) and should not interact with FG-domains. To
218 confirm that the synthetic hydrophobic reporters are able to interact with the aromatic phenyl
219 group in FG domains in the gels independently of their charged domains, we used phenyl-
220 sepharose columns to test for hydrophobic interactions. This assay has previously been
221 successful in isolating intact NTRs from cell lysate extracts (12). As seen in Supplementary
222 Figure 3, both hydrophobic reporters show an increase in retention time compared to their
223 hydrophilic counterpart, indicating that the hydrophobic interactions are occurring between the
224 hydrophobic reporters and the phenyl groups on the sepharose independent of the displayed
225 charge.

226

Using the FGAK gels, we first tested if the two hydrophobic fluorescent reporters demonstrate different selective uptake into the gel. FGAK was dissolved at 2% (w/v), loaded into glass capillaries, and bathed in a fluorescent reporter solution to test for selective binding. Reporters were loaded into capillaries at 10 μM concentrations and interactions were monitored for five hours at the buffer-gel interface. To quantify the selective binding of the various reporters into the respective gels, we again estimate the effective diffusion coefficient of the reporters in the gel and the maximum accumulation of the reporters at the interface as metrics for determining how well the gel is able to differentiate between hydrophobic and charged domains in the reporters.

Figure 3C shows that the cationic FGAK gel interacts with the Hydrophobic (-) reporters, with significant accumulation inside the gel. Quantifying the maximum concentration at the interface, we see that there is approximately a 2-fold increased enrichment compared to the Hydrophilic (-) reporter, which does not contain the hydrophobic tail. This result suggests that FG sequences in FGAK recognize reporter phenylalanines when the reporter exhibits negative charge. The contribution of hydrophobic interactions is further corroborated by the strong reduction in effective diffusion coefficient from 7 μm^2/min for Hydrophilic (-) down to below 1 μm^2/min for Hydrophobic (-), indicating that the hydrophobic interactions are able to induce tighter binding in the context of electrostatic attraction. In fact, the interactions are so strong that the Hydrophobic (-) reporters essentially do not diffuse over the five-hour period analyzed, suggesting that the binding is irreversible. In contrast, when the reporters contain a cationic tail, we see that Hydrophobic (+) and Hydrophilic (+) reporters diffuse into the gel with coefficients of 380 μm^2/min and 2000 μm^2/min, respectively. The Hydrophobic (+) reporters are able to bind 1.5 times above the original bath concentration at the interface whereas Hydrophilic (+) reporters

equilibrate with no discernible partitioning (Figure 3D). These results show that hydrophobic interactions can occur under electrostatically repelling environments, but the interactions are much weaker compared to the Hydrophobic (-) reporters. Moreover, by lowering the salt concentrations from 200 mM NaCl to 20 mM to decrease electrostatic screening and increase electrostatic repulsion, the accumulation at the interface decreases and the diffusion coefficient shows a trend toward an increase (Figure 3D), indicating that the repulsion weakens the overall strength of hydrophobic interactions with FG-domains. These results show that the lysines confer upon FGAK gels the striking ability to distinguish between two substrates that contain the same hydrophobic domain but with different surrounding charge types. In particular, the lysines allows for increased binding to hydrophobic domains with neighboring anionic residues.

The analysis of the Nsp1 repeat consensus sequence (Figure 1A-B) suggests that the conserved location of lysines within 2-3 amino acids from the FG sequences may be relevant for FG-mediated molecular recognition. To determine how the spatial localization of lysines affects FG sequence differentiation between various hydrophobic domains, we tested if placing lysines immediately adjacent to FG sequences (FGK), or moving them 3 amino acids away (FGA_2K), would affect the sequence's selectivity compared to the original arrangement in the FGAK peptide. Using the same conditions as before, both peptides were dissolved at 2% (w/v). While FGA_2K peptides readily formed a stiff material and could be tested for selective uptake, the FGK peptide gel is two orders of magnitude less stiff, flowed when inverted, and dispersed when loading fluorescent reporters (Figure 3A). Therefore FGK could not be analyzed for transport and the rest of the analysis will be a comparison on diffusion and accumulation between only FGAK and FGA_2K. As previously shown in Figure 3C and D, the FGAK gel selectively

enriched for the Hydrophobic (-) reporter but failed to uptake the Hydrophobic (+) reporter to the same degree. When comparing the two anionic Hydrophobic (-) and Hydrophilic (-) reporters in FGA_2K, the diffusivity of Hydrophobic (-) is an order of magnitude lower compared to Hydrophilic (-), showing that the combination of hydrophobic and electrostatic attraction synergizes for strong binding. However, the max accumulation at the interface for the two anionic reporters is similar, indicating FGA_2K gel is unable to differentiate between the two reporters to the same degree as the original FGAK gel (Figure 3E). For cationic reporters, the Hydrophobic (+) peptide accumulates within the gel and has an order of magnitude lower diffusivity compared to the Hydrophilic (+) reporter. Moreover, the accumulation of Hydrophobic (+) reporters for FGA_2K gels is higher compared to FGAK gels (Figure 3F). These data show that placing lysines within 2 amino acids of FG-domains (FGAK), approximately corresponding to the 1 nm Debye length at physiologically relevant salt concentrations, such as in our experimental conditions, enables hydrophobic selectivity of substrates with opposite charge. Placing lysines 3 amino acids away (FGA_2K) reduces the contribution of electrostatic interactions in determining hydrophobic FG-mediated selectivity. As a result, by increasing the distance between charged residues and FG sequences, the FG sequences becomes less dependent on the electrostatic profile surrounding the hydrophobic substrates during selective transport.

Charge reversal flips selectivity of FG-domains

Since lysines in the FGAK gel help FG sequences differentiate between reporters containing anionic or cationic hydrophobic domains, we next asked if FG-mediated recognition could be reversed by replacing lysines with anionic glutamic acids (FGAE). Using the same class of fluorescent reporters, we find that the selectivity of FGAE gels is reversed from FGAK-gels

(Figure 4A-D). Again, the selectivity is a function of both hydrophobic and electrostatic interactions as the Hydrophilic (+) reporters accumulated 2-fold less at the interface compared to Hydrophobic (+) reporters (Figure 4D). Moreover, the diffusion coefficient for Hydrophobic (+) is an order of magnitude lower than for Hydrophilic (+). Conversely, both Hydrophobic (-) and Hydrophilic (-) reporters did not interact significantly with the gel and exhibited similar effective diffusion coefficients, showing that electrostatic repulsion prevents hydrophobic interactions with FG sequences. Together, these data suggest that charge proximal to FG sequences can tune hydrophobic selectivity and that charge is essential in determining which hydrophobic moieties FG sequences can recognize.

Sidechain biochemistry as a regulator of assembly

Lastly, we tested how sidechain biochemistry of charged amino acids affect FG sequence function for both self-assembly and molecular recognition of transport molecules. From the Nsp1 consensus sequence analysis, we observed that the predominant cation is lysine rather than arginine, suggesting that the two cations may not be treated equally. Similarly, glutamic acids are conserved closer to FG sequences whereas aspartic acids appear closer to the central region of the spacer sequences and are not as well conserved. To test the role of sidechain differences, we first compare how FGAR gels compared to FGAK. In the diffusion assay, we find that FGAR gels have similar uptake properties when compared to FGAK gels (Figure 4A, E), where the Hydrophobic (-) reporter accumulates and interacts with the gel interface while the Hydrophobic (+) reporter does not accumulate and instead diffuses near freely into the gel (Figure 4E and 4F). These data suggest that that from an electrostatic standpoint, arginines are just as capable as lysines in helping FG sequences differentiate between various types of hydrophobic substrates.

However, mechanically, FGAR forms gels that exhibited an approximate stiffness of 2000 Pa throughout the frequency sweep (Supplementary Figure 4C), which is 4-fold more compliant than FGAK gels. The changes in the mechanical properties of the FGAK and FGAR gels suggest that the lysines and arginines may predominantly affect the structural self-assembly of FG-domains. Indeed, from TEM analysis, we find that FGAR form different structures compared to FGAK (Supplementary Figure 4D). To compare FGAD and FGAE gels, we find that FGAD does not form a gel at the standard 2% (w/v) and forms no repeating structures (Supplementary Figure 4A-B), suggesting that the small chemical changes between glutamic acid and aspartic acid are able to transition the material from a selective gel to a viscous solution. Together, these data show the detailed amino acid biochemistry, in addition to net charge, is an important regulatory parameter in determining the structure of self-assembled FG-domains with minor effects on selectivity.

Tryptophan interactions are similarly modulated by electrostatics

An essential question in the selectivity of FG-domains is how generalizable the tunability of hydrophobic selectivity is to other natural aromatic amino acids. In previous assays with the designed reporters, we chose phenylalanines as the representative aromatic residue for hydrophobic interactions with FG-domains. In native NTRs such as NTF2 and other transport receptors, tryptophan is commonly used for the transient hydrophobic interactions required for transport. To test if the electrostatic dependence still exists when phenylalanines are converted to tryptophans, we compared two new hydrophobic reporters Hydrophobic (+W) and Hydrophobic (-W) to the original hydrophobic phenylalanine reporters (now termed +F and –F, respectively, for the hydrophobic residue). As seen in Figure 5, the tryptophan reporters are modulated by

342 electrostatic interaction in a similar manner to phenylalanine reporters. We find that the max
343 accumulated fluorescence at the interface and diffusion between Hydrophobic -F and -W
344 reporters are not significantly different (Figure 5A). For cationic reporters, when phenylalanines
345 (+F) are substituted for tryptophans (+W), the reporters exhibit similar max interface properties
346 but tryptophan (+W) reporters show a slight trend toward slower diffusion (Figure 5B),
347 indicating that although both aromatic residues can be modulated by electrostatic interactions,
348 choice of hydrophobic residue will have small but potentially significant differences in transport
349 properties. These results reveal that the phenomenon of electrostatic interactions regulating
350 hydrophobic interactions may be a generalizable phenomenon and is not restricted to
351 phenylalanines. Moreover, the choice of type of non-polar amino acid is now established as
352 another parameter in which selective transport can be tuned.
353
354 **Discussion**
355 Our study shows how the environment surrounding FG sequences tunes hydrophobic interactions
356 necessary for correct molecular recognition of hydrophobic substrates. Here, we explicitly show
357 how the presence, placement, and type of charged amino acid are all essential parameters in
358 controlling FG function for these reduced FG repeats. The method of using peptides based on
359 consensus sequences of Nsp1 reveals how complex function can be encoded in minimal
360 sequence space. The work here supports previous theoretical predictions that electrostatic
361 interactions may be an important regulator of nuclear pore selectivity (22) and sequence
362 observations that the bias in the net charge of nuclear transport receptors may be important in
363 NPC transport (21). Moreover, recent reports on how charge influences interactions with
364 hydrophobic surfaces (8,34) emphasize that the interplay between electrostatics and

hydrophobicity may be a fundamental process for different biological functions such as molecular recognition and protein folding. Here, we provide an example of how these electrostatics and hydrophobicity can tune the molecular recognition properties of biological supramolecular structures.

Our work shifts the spotlight from analyzing only how hydrophobicity determines nuclear pore selectivity to the underappreciated contribution of electrostatics in molecular recognition (9,11,12,35). The results described here provide insight on how individual hydrophobic FG-sequences may acquire the astounding ability to distinguish a particular subset of hydrophobic substrates using a combination of electrostatic and hydrophobic interactions. Our data also imply that substrate interaction with the FG-containing hydrogel is not a binary effect, but can be tuned by the placement of charge surrounding FG domains to achieve a range of transport behaviors, from complete retention at the interface to free diffusion through the hydrogel, as depicted schematically in Figure 6.

We recognize that although the peptides can answer questions on how the molecular recognition of FG-sequences can be manipulated by the precise placement of charged amino acids, the short peptide gels cannot recapitulate the unique facilitated diffusion properties of the intact nuclear pore. This likely arises from structural differences, as short peptides form fibers and rods, which are not observed and predicted in the native nuclear pore. Future work to address these challenges include synthesizing longer peptides consisting of multiple consensus repeats to determine how repeat number influences structure and selective recognition. Nevertheless, the current peptides provide unique advantages that are absent in more complex and degenerate

388 systems. For instance, previous studies with intact nucleoporins containing multiple repeats of
389 FG-domains suggest that charge primarily plays a structural role in the cohesion of the self-
390 assembled matrix (36), which results in altered selective properties of the gel. In contrast,
391 peptides with single repeats show that electrostatic interactions directly influence how biological
392 substrates interact with hydrophobic FG sequences for molecular recognition. FG-containing
393 peptides can capture at single amino acid resolution on how the interplay between hydrophobic
394 and electrostatic interactions. These results also support recent work where ImpB was recognized
395 to have varying binding capacities with synthetic FG-containing polyacrylamide gels depending
396 on the charged state of the material (37). We believe this is a useful approach that complements
397 other well-established methods used to understand the selective transport mechanism of NPCs
398 such as the *in vivo* minimal nuclear pore complex (18,19), gel and selectivity analysis of
399 individual nucleoporins (9-11,38), and binding interactions with surface-grafted nucleoporin
400 films (36,39-41). Moreover, we expect that this rational approach using conserved repeating
401 sequences can be expanded to further study other biological gel systems such as mucus, byssal
402 threads, and cartilage, where complex disordered proteins based on repeat units and reversible
403 crosslinking contribute a significant proportion of the material (42-44).
404
405 **Materials and Methods:**
406 Sequence analysis:
407 Sequence logo of consensus sequence was generated using Berkeleys' WebLogo software(45) by
408 aligning 15 repeats according to FSFG as references. Conserved fraction was calculated for
409 positive charges (# cationic residues K or R)/(15 repeats) and similarly for negative charges (#
410 anionic residue E or D)/(15 repeats) at each position indicated.

411

412 NTF2 Expression and Labeling

413 pQE30-NTF2-6xHis was a gift from the Gorlich lab and transformed into DH5alpha cells for
414 cloning. A cysteine was inserted for fluorescent labeling after the 6x His-tag using standard site
415 directed mutagenesis techniques using the two primers (5' –
416 ctcagctaattaagcttagcagtgatggtgatggtgatgagatctg - 3' and 5' -
417 cagatctcatcaccatcaccatcactgctaagcttaattagctgag – 3') to form pQE30-NTF2-6xHis-Cys. The
418 W7A-Cys mutant containing the terminal cysteine was created by using pQE30-NTF2-6xHis-
419 Cys and applying standard site direct mutagenesis protocols with the following primers (5' –
420 tgaggagccaatttgttccgcgatcggtttatcacccatg – 3' and 5' – catgggtgataaaccgatcgcggaacaaattggctcctca
421 – 3'). Expression of NTF2-Cys and W7A was completed in OverExpress C41(DE3) cells
422 (Lucigen) and purified using standard nickel column purification and ion exchange columns.
423 Labeling of NTF2-Cys and W7A-Cys was completed using Fluorescein-5-maleimide
424 (Thermofischer Scientific, Catalog #F150) using their recommended protocol and gel purified.
425 Labeling efficiency was approximately 50%. Final concentrations of NTF2 and W7A for
426 transport were adjusted to 10 uM concentrations with 10% of the population labeled.

427

428 Peptide and Gel Preparation:

429 Unless specified otherwise, all chemicals were obtained from Sigma Aldrich. Peptides were
430 prepared by MIT's Koch Institute Biopolymers and Proteomics Facility (Cambridge, MA) and
431 Boston Open Labs LLC (Cambridge, MA). All peptides were HPLC purified unless specified
432 otherwise, desalted using reverse phase HPLC with 0.05% TFA, and lyophilized after synthesis
433 with >95% purity. For fluorescently labeled peptides, a 5-carboxyfluorescein (Anaspec)

fluorophore is added to the N-terminus whereas the C-terminus is modified to be an amide. Fluorescent peptides were diluted into 200 mM NaCl with 20 mM HEPES, pH 7 buffer at 10 uM concentration for transport experiments. The gel peptides were all dissolved in 20 mM NaCl, 20 mM HEPES, pH 7 buffer at 2% (w/v). To facilitate solubilization and gel formation, peptides were vortexed for 30 seconds and briefly sonicated in a bath sonicator (Branson 2510) to reduce aggregation.

Capillary Assay and Analysis:

1.5 inch length borosilicate square capillaries with 9 mm cross sectional width (Vitrocom 8290) are loaded by piercing pre-made hydrogels. 10 uM solutions (200 mM NaCl, 20 mM HEPES pH 7) of fluorescent peptides are injected into the capillary and sealed by a 1:1:1 (by weight) mixture of Vaseline, lanolin, and paraffin. Time lapses of peptide diffusion are taken using at 1 minute intervals for up to five hours on a Nikon Ti Eclipse inverted microscope using a Nikon CFI Plan UW 2X or on an AxioObserver D.1 with a EC Plan-Neofluar 1.25x/0.03 WD=3.9 and Hamamatsu C11440-22CU camera. All fluorescence profiles were obtained by averaging the fluorescence intensities across the width of the capillary. Normalized concentration profiles were obtained by normalizing fluorescence intensities to the bath concentration of the capillary at the initial time point. The fluorescence is signal is linear up to 50 uM (Supplementary Figure 5). For concentration profile plotting, the signal is not adjusted past the saturation point and represents the lower bound of the actual concentration of reporters accumulating in the gel. All data presented represents at least three independent replicates. Students t-test was applied to determine p-values between experimental conditions.

Effective diffusion rates were fit by minimizing the squared error of a simulated concentration timecourse in a region of the capillary on the gel side of the interface, over a 100-minute window. To do this, the diffusion equation for the concentration of probe *c*:

$$\frac{\partial c(x,t)}{\partial t} = D\frac{\partial^2 c(x,t)}{\partial x^2}$$

was numerically solved using Matlab's *pdepe* function (MathWorks; Natick, MA). The initial condition $c(x,0)$ was set by the concentration profile at the first timepoint used, and the boundary conditions $c(0,t)$ and $c(L,t)$ (for a fit over length L) were similarly determined by the concentration profiles at the edges of the region of interest. The D that minimized the squared difference between simulated and actual concentration profiles was the value reported; minimization of error took place iteratively using a home-written modified gradient descent algorithm. The time period from 150-250 minutes was generally used for fitting, but for particularly fast-diffusing probes (D>1000 μm^2/min) 50-150 or 30-130 minutes was used. The earlier time period allowed for the fitting to take place before steady-state or pseudo-steady state was reached, which was crucial for fitting to be precise. Code for the diffusion analysis is provided in the supplementary materials.

Rheological Testing:

Experiments were performed on an Anton Paar MCR 302 Rheometer in a cone-plate geometry with a 25mm diameter, 1 degree cone angle, and 51 micron truncation. The temperature was maintained at 25C and evaporation controlled with an H2O-filled solvent trap. To identify the linear regime, amplitude sweeps were conducted at ω = 10 rad/s from γ_0 = 0.01% to 100% strain. In the linear regime, frequency sweeps were conducted using the previously determined strain amplitude from ω = 100 rad/s to 0.1 rad/s.

480

481 Transmission Electron Microscopy:

482 Images were taken using a JEOL-1200 transmission electron microscope. Gels were formed as
483 aforementioned and spotted onto glow discharged carbon-coated copper-Formvar grids (Ted
484 Pella). Excess liquid and gel were removed using Whatman paper or parafilm. Grids were
485 submerged in 1% uranyl acetate solution, blotted, and air dried for 15 minutes prior to imaging.
486 Images of gels are representative of at least four images taken of each gel type.

487

488 Phenyl-sepharose chromatography:

489 Fluorescent reporters were dissolved in 200 mM NaCl, 20 mM HEPES pH 7 and loaded into
490 high-copy FF 1 mL capacity phenyl-sepharose columns equilibrated with three volumes of 200
491 mM NaCl, 20 mM HEPES pH 7. For elution, flow rates were set to 1 ml/min with 0.5 mL
492 fraction volumes. The fraction concentrations are representative traces of elution profiles.

493

494 Nile Red Fluorescence Assay:

495 Using the same transport methods, Nile Red dye (Thermo Fischer Scientific, N1142) was loaded
496 into capillaries at 10 ug/ml concentrations in 200 mM NaCl, 20 mM HEPES buffers, pH 7.

497

498 Acknowledgements:

499 This work was supported by the MRSEC Program of the National Science Foundation under
500 award number DMR – 0819762, Defense Threat Reduction Agency under award number
501 HDTRA1-13-1-0038, NSF RO1 R01-EB017755, and NSF Career PHY-1454673. W.G.C. was
502 supported by the NIGMS/NIH Interdepartmental Biotechnology Training Program under T32

GM008334. J.W. was supported in part by the National Science Foundation Graduate Research Fellowship Program under Grant No. 1122374. We wish to thank Prof. Alan Grodzinsky at MIT for his valuable feedback and helpful discussions, and Felice Frankel for the images of the gels.

Competing interests:

The authors declare no competing interests.

Figure Legends:

Figure 1: Identification of conserved repeat sequences and design of simplified FG self-assembling peptides

A) Conserved sequence identification of the C-terminal end of essential yeast nucleoporin Nsp1. B) Conservation of charge between FSFG domains in Nsp1 sequences. Positive/negative values refer to the presence of cationic or anionic residues. Amino acid numbers correspond to position in the consensus repeats as seen in Figure 1A. C) Simplified peptides consisting of 14 amino acids were designed with the sequence structure FSFGAXAAXAFSFG where X represents the substituted amino acids lysine (K), glutamic acid (E), or serine (S) to determine how presence of charge affects FG-mediated selectivity. C) Designed peptides where the lysines are moved either immediately adjacent (FGK) or placed three away (FGA_2K) from FSFG domains to test how spatial localization of charge influences FG selectivity. D) Class of peptides where lysines (K) are substituted with arginine (R) or aspartic acid (D) to determine how biochemistry affects FG-mediated selectivity and self-assembly.

Figure 2: FGAK and FGAE self-assembly and selectivity of native transport receptor NTF2

A) Frequency sweep of FGAK gel with G' (storage) and G" (loss) moduli reported at 2% (w/v) with TEM image of self-assembled peptides and macroscopic gel (inset). B) Frequency sweep of FGAE gel with G' (storage) and G" (loss) moduli reported at 2% (w/v) with TEM image of self-assembled peptides and macroscopic gel (inset). C) Schematic of slab 1-D transport system within a sealed capillary providing no flux boundary conditions. Example diffusion of fluorescent molecules is depicted at 10 minute, 30 minute, and 60 minute timepoints. The interface of the gel is depicted by the increase in fluorescence signal between the buffer and hydrogel. D-E) Transport of representative fluorescently labeled NTF2 and W7A proteins into FGAK and FGAE gels, respectively. The fluorescent images and corresponding concentration profiles are at the 5-hour time point. F) The max concentration at the interface and effective diffusion coefficient are quantified for both NTF2 and W7A in FGAK and FGAE gels. All error bars are standard deviations of at least three independent replicates. * indicates $p<0.05$ using unpaired Student's t-test.

Figure 3: Effects of spatial localization of charge on FG-mediated self-assembly and selectivity A) Frequency sweep of FGA_2K and FGK gels with G' (storage) and G" (loss) moduli reported at 2% (w/v). B) Schematic of the four fluorescent reporters that contain a charged domain and a hydrophilic (N) or hydrophobic (F) tail. Fluorescent reporters are loaded into capillaries containing the gel of choice where diffusion is monitored for up to 5 hours. C) Representative fluorescence images of Hydrophobic (-) and Hydrophilic (-) reporters into FGAK gel with the corresponding concentration profile, max concentration at gel interface, and estimated effective diffusion coefficient. D) Representative fluorescence images of Hydrophobic (+) and Hydrophilic (+) reporters into FGAK gel with the corresponding concentration profile, max

549 concentration at gel interface, and estimated effective diffusion coefficient. An additional low
550 salt condition was added (20 mM NaCl) for Hydrophobic (+) reporters as a comparison to
551 standardized (200 mM NaCl) conditions for max concentration and effective diffusion
552 estimation. E) Representative fluorescence images of Hydrophobic (-) and Hydrophilic (-)
553 reporters into FGA_2K gel with the corresponding concentration profile, max concentration at gel
554 interface, and estimated effective diffusion coefficient. F) Representative fluorescence images of
555 Hydrophobic (+) and Hydrophilic (+) reporters into FGA_2K gel with the corresponding
556 concentration profile, max concentration at gel interface, and estimated effective diffusion
557 coefficient. All error bars are standard deviations of at least three independent replicates. *
558 indicates $p<0.05$. ** indicates $p<0.01$ between hydrophobic and hydrophilic reporters using
559 unpaired Student's t-test.
560
561 Figure 4: Glutamic acid reverses selectivity of FG-based gels and arginine displays similar
562 selectivity to lysine
563 A)-B) Representative Hydrophobic (-) and Hydrophobic (+) fluorescence images in FGAK gels
564 for comparison to FGAE and FGAR gel selectivity. C) Representative fluorescence images of
565 Hydrophobic (-) and Hydrophilic (-) reporters into FGAE gel with the corresponding
566 concentration profile, max concentration at gel interface, and estimated effective diffusion
567 coefficient. D) Representative fluorescence images of Hydrophobic (+) and Hydrophilic (+)
568 reporters into FGAE gel with the corresponding concentration profile, max concentration at gel
569 interface, and estimated effective diffusion coefficient. E) Representative fluorescence images of
570 Hydrophobic (-) and Hydrophilic (-) reporters into FGAR gel with the corresponding
571 concentration profile, max concentration at gel interface, and estimated effective diffusion

572 coefficient. F) Representative fluorescence images of Hydrophobic (+) and Hydrophilic (+)
573 reporters into FGAR gel with the corresponding concentration profile, max concentration at gel
574 interface, and estimated effective diffusion coefficient. All error bars are standard deviations of
575 at least three independent replicates. * indicates $p<0.05$. ** indicates $p<0.01$ between
576 hydrophobic and hydrophilic reporters using unpaired Student's t-test.
577
578 Figure 5: Comparison of selective transport of tryptophan and phenylalanine in FGAK gels
579 A) Comparison of Hydrophobic (-F) and Hydrophobic (-W) reporters into FGAK gels with
580 representative fluorescent images and corresponding concentration profiles. Estimated max
581 interface concentrations and effective diffusion coefficients are presented. B) Comparison of
582 Hydrophobic (+F) and Hydrophobic (+W) reporters into FGAK gels with representative
583 fluorescent images and corresponding concentration profiles. Estimated max interface
584 concentrations and effective diffusion coefficients are presented. All error bars are standard
585 deviations of at least three independent replicates. No significant differences are reported using
586 Student's t-test.
587
588 Figure 6: Role of electrostatics in modulating hydrophobic FG function
589 Schematic of the influence of electrostatic on hydrophobic interactions mediated by FG-domains.
590 At physiologically relevant salt concentrations, electrostatic interactions typically have a range of
591 ~1 nm (Debye length), which determines how much a charged residue can influence what
592 hydrophobic substrates FG-domains can recognize and bind to. By moving charged residues
593 further away, the electrostatic effects will be less significant on FG-mediated self-assembly and
594 selective transport.

595 Supplementary Figure 1: Frequency sweep of F→S substitution (FGAK → SGAS) to determine
596 effect of phenylalanines on self-assembly of peptides. The elastic modulus (G') and loss modulus
597 (G") are reported, however, the measured values are below the sensitivity of the rheometer using
598 the specified cone-plate geometry.
599
600 Supplementary Figure 2: A) Transport of Nile Red dye into FGAK and FGAE gels at the 0 hour
601 and 3 hour timepoint. Fluorescence indicates the dye is able to access hydrophobic environments
602 created by FG-domains within the gels. Images are of representative gels from three independent
603 replicates.
604
605 Supplementary Figure 3:
606 Fractionation of hydrophilic reporters compared to their hydrophobic counterparts in
607 phenylsepharose columns. The fluorescence signals of each fraction is collected and normalized
608 to the signal with the highest intensity of emission. For both cationic and anionic reporters, the
609 hydrophobic reporters are eluted later, indicating that they show an increased retention time, and
610 thus stronger binding to phenyl sepharose beads.
611
612 Supplementary Figure 4:
613 A) Frequency sweep of FGAD peptide solution with G' (storage) and G" (loss) moduli reported
614 at 2% (w/v) and corresponding TEM image (B). C) Frequency sweep of FGAR gel with G'
615 (storage) and G" (loss) moduli reported at 2% (w/v) and corresponding TEM image (D).
616
617 Supplementary Figure 5:

618 Quantification of fluorescence signal of fluorophores as a function of concentration.

619 Fluorescence signal is approximately linear up to 50 uM concentrations and saturates by 100 uM.

620 All concentrations are reported according to the experimental curve developed and represent

621 lower values of the actual concentrations if values exceed 100 uM.

622

623 **References:**

624 1. Stewart M. Molecular mechanism of the nuclear protein import cycle. Nature Reviews
625 Molecular Cell Biology. 2007 Mar 1;8(3):195–208.

626 2. Stewart M, Baker RP, Bayliss R, Clayton L, Grant RP, Littlewood T, et al. Molecular
627 mechanism of translocation through nuclear pore complexes during nuclear protein
628 import. FEBS Lett. 2001 Jun 8;498(2-3):145–9.

629 3. Görlich D, Kutay U. Transport between the cell nucleus and the cytoplasm. Annu Rev
630 Cell Dev Biol. Annual Reviews 4139 El Camino Way, P.O. Box 10139, Palo Alto, CA
631 94303-0139, USA; 1999 Nov;15(1):607–60.

632 4. Ando D, Zandi R, Kim YW, Colvin M, Rexach M, Gopinathan A. Nuclear pore complex
633 protein sequences determine overall copolymer brush structure and function. Biophys J.
634 2014 May 6;106(9):1997–2007.

635 5. Ribbeck K, Görlich D. Kinetic analysis of translocation through nuclear pore complexes.
636 EMBO J. 2001 Mar 15;20(6):1320–30.

637 6. Ghavami A, Veenhoff LM, van der Giessen E, Onck PR. Probing the Disordered Domain
638 of the Nuclear Pore Complex through Coarse-Grained Molecular Dynamics Simulations.
639 Biophysj. Biophysical Society; 2014 Sep 16;107(6):1393–402.

640 7. Gamini R, Han W, Stone JE, Schulten K. Assembly of Nsp1 nucleoporins provides insight
641 into nuclear pore complex gating. PLoS Comput Biol. 2014 Mar;10(3):e1003488.

642 8. Ma CD, Wang C, Acevedo-Vélez C, Gellman SH, Abbott NL. Modulation of hydrophobic
643 interactions by proximally immobilized ions. Nature. 2015 Jan 15;517(7534):347–50.

644 9. Frey S, Görlich D. A saturated FG-repeat hydrogel can reproduce the permeability
645 properties of nuclear pore complexes. CELL. 2007 Aug 10;130(3):512–23.

646 10. Frey S, Richter RP, Görlich D. FG-rich repeats of nuclear pore proteins form a three-
647 dimensional meshwork with hydrogel-like properties. Science. 2006 Nov
648 3;314(5800):815–7.

649 11. Frey S, Görlich D. FG/FxFG as well as GLFG repeats form a selective permeability

barrier with self-healing properties. EMBO J. 2009 Aug 13;28(17):2554–67.

12. Ribbeck K, Görlich D. The permeability barrier of nuclear pore complexes appears to operate via hydrophobic exclusion. EMBO J. 2002 Jun 3;21(11):2664–71.

13. Bayliss R, Littlewood T, Strawn LA, Wente SR, Stewart M. GLFG and FxFG nucleoporins bind to overlapping sites on importin-beta. J Biol Chem. 2002 Dec 27;277(52):50597–606.

14. Bayliss R, Ribbeck K, Akin D, Kent HM, Feldherr CM, Görlich D, et al. Interaction between NTF2 and xFxFG-containing nucleoporins is required to mediate nuclear import of RanGDP. J Mol Biol. 1999 Oct 29;293(3):579–93.

15. Strawn LA, Shen T, Wente SR. The GLFG regions of Nup116p and Nup100p serve as binding sites for both Kap95p and Mex67p at the nuclear pore complex. J Biol Chem. 2001 Mar 2;276(9):6445–52.

16. Grant RP, Neuhaus D, Stewart M. Structural basis for the interaction between the Tap/NXF1 UBA domain and FG nucleoporins at 1A resolution. J Mol Biol. 2003 Feb 21;326(3):849–58.

17. Bednenko J, Cingolani G, Gerace L. Importin beta contains a COOH-terminal nucleoporin binding region important for nuclear transport. J Cell Biol. 2003 Aug 4;162(3):391–401.

18. Strawn LA, Shen T, Shulga N, Goldfarb DS, Wente SR. Minimal nuclear pore complexes define FG repeat domains essential for transport. Nat Cell Biol. 2004 Feb 22;6(3):10–206.

19. Patel SS, Belmont BJ, Sante JM, Rexach MF. Natively unfolded nucleoporins gate protein diffusion across the nuclear pore complex. CELL. 2007 Apr 6;129(1):83–96.

20. Ribbeck K, Kutay U, Paraskeva E, Görlich D. The translocation of transportin–cargo complexes through nuclear pores is independent of both Ran and energy. Current biology. 1999.

21. Colwell LJ, Brenner MP, Ribbeck K. Charge as a selection criterion for translocation through the nuclear pore complex. PLoS Comput Biol. 2010 Apr;6(4):e1000747.

22. Tagliazucchi M, Peleg O, Kröger M, Rabin Y, Szleifer I. Effect of charge, hydrophobicity, and sequence of nucleoporins on the translocation of model particles through the nuclear pore complex. Proceedings of the National Academy of Sciences. 2013 Feb 26;110(9):3363–8.

23. Ando D, Colvin M, Rexach M, Gopinathan A. Physical Motif Clustering within Intrinsically Disordered Nucleoporin Sequences Reveals Universal Functional Features. PLoS ONE. 2013;8(9):e73831.

24. Yamada J, Phillips JL, Patel S, Goldfien G, Calestagne-Morelli A, Huang H, et al. A bimodal distribution of two distinct categories of intrinsically disordered structures with

685 separate functions in FG nucleoporins. Mol Cell Proteomics. 2010 Oct;9(10):2205–24.

686 25. Zhang S. Fabrication of novel biomaterials through molecular self-assembly. Nature
687 Biotechnology. 2003 Oct;21(10):1171–8.

688 26. Hinman MB, Jones JA, Lewis RV. Synthetic spider silk: a modular fiber. Trends
689 Biotechnol. 2000 Sep;18(9):374–9.

690 27. Xu M, Lewis RV. Structure of a protein superfiber: spider dragline silk. Proceedings of
691 the National Academy of Sciences. 1990 Sep 15;87(18):7120–4.

692 28. Meyer DE, Chilkoti A. Quantification of the effects of chain length and concentration on
693 the thermal behavior of elastin-like polypeptides. Biomacromolecules. 2004
694 May;5(3):846–51.

695 29. Nettles DL, Chilkoti A, Setton LA. Applications of elastin-like polypeptides in tissue
696 engineering. Adv Drug Deliv Rev. 2010 Dec;62(15):1479–85.

697 30. Wright ER, Conticello VP. Self-assembly of block copolymers derived from elastin-
698 mimetic polypeptide sequences. Adv Drug Deliv Rev. 2002 Oct 18;54(8):1057–73.

699 31. Hurt EC. A novel nucleoskeletal-like protein located at the nuclear periphery is required
700 for the life cycle of Saccharomyces cerevisiae. EMBO J. 1988 Dec 20;7(13):4323–34.

701 32. Denning DP. Disorder in the nuclear pore complex: The FG repeat regions of
702 nucleoporins are natively unfolded. Proceedings of the National Academy of Sciences.
703 2003 Feb 25;100(5):2450–5.

704 33. Denning DP, Uversky V, Patel SS, Fink AL. The Saccharomyces cerevisiae nucleoporin
705 Nup2p is a natively unfolded protein. … of Biological Chemistry. 2002.

706 34. Chen S, Itoh Y, Masuda T, Shimizu S, Zhao J, Ma J, et al. Ionic interactions.
707 Subnanoscale hydrophobic modulation of salt bridges in aqueous media. Science. 2015
708 May 1;348(6234):555–9.

709 35. Hülsmann BB, Labokha AA, Görlich D. The permeability of reconstituted nuclear pores
710 provides direct evidence for the selective phase model. CELL. 2012 Aug 17;150(4):738–
711 51.

712 36. Ader C, Frey S, Maas W, Schmidt HB, Görlich D, Baldus M. Amyloid-like interactions
713 within nucleoporin FG hydrogels. Proceedings of the National Academy of Sciences.
714 2010 Apr 6;107(14):6281–5.

715 37. Friedman AK, Baker LA. Synthetic hydrogel mimics of the nuclear pore complex display
716 selectivity dependent on FG-repeat concentration and electrostatics. - PubMed - NCBI.
717 Soft Matter. 2016.

718 38. Labokha AA, Gradmann S, Frey S, Hülsmann BB, Urlaub H, Baldus M, et al. Systematic

719 analysis of barrier-forming FG hydrogels from Xenopus nuclear pore complexes. EMBO
720 J. 2013 Jan 23;32(2):204–18.

721 39. Eisele NB, Labokha AA, Frey S, Görlich D, Richter RP. Cohesiveness tunes assembly and
722 morphology of FG nucleoporin domain meshworks - Implications for nuclear pore
723 permeability. Biophys J. 2013 Oct 15;105(8):1860–70.

724 40. Eisele NB, Frey S, Piehler J, Görlich D, Richter RP. Ultrathin nucleoporin phenylalanine-
725 glycine repeat films and their interaction with nuclear transport receptors. EMBO Rep.
726 2010 May;11(5):366–72.

727 41. Zahn R, Osmanović D, Ehret S, Callis CA, Frey S, Stewart M, et al. A physical model
728 describing the interaction of nuclear transport receptors with FG nucleoporin domain
729 assemblies. Elife. eLife Sciences Publications Limited; 2016 Apr 8;5:e14119.

730 42. Holten-Andersen N, Zhao H, Waite JH. Stiff coatings on compliant biofibers: the cuticle
731 of Mytilus californianus byssal threads. Biochemistry. 2009 Mar 31;48(12):2752–9.

732 43. Waller KA, Zhang LX, Elsaid KA, Fleming BC, Warman ML, Jay GD. Role of lubricin
733 and boundary lubrication in the prevention of chondrocyte apoptosis. Proc Natl Acad Sci
734 USA. 2013 Mar 25;:–.

735 44. Rose MC, Voynow JA. Respiratory tract mucin genes and mucin glycoproteins in health
736 and disease. Physiol Rev. 2006 Jan;86(1):245–78.

737 45. Crooks GE, Hon G, Chandonia J-M, Brenner SE. WebLogo: a sequence logo generator.
738 Genome Res. 2004 Jun;14(6):1188–90.

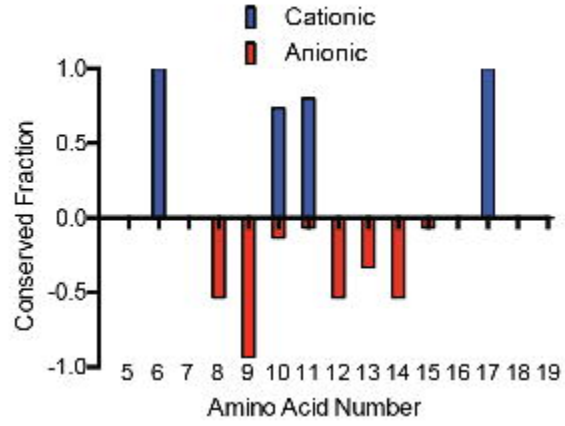
739

Fig 1

A

Nsp1 Consensus Repeat

B



C

Charge Relevance

FGAK

FSFGAKAAKAFSFG

FGAE

FSFGAEAAEAFSFG

FGAS

FSFGASAASAFSFG

D

Spatial Localization

FGK

FSFGKAAAAKFSFG

FGA_2K

FSFGAAKKAAFSFG

E

Sidechain Biochemistry

FGAR

FSFGARAARAFSFG

FGAD

FSFGADAADAFSFG

Fig 2

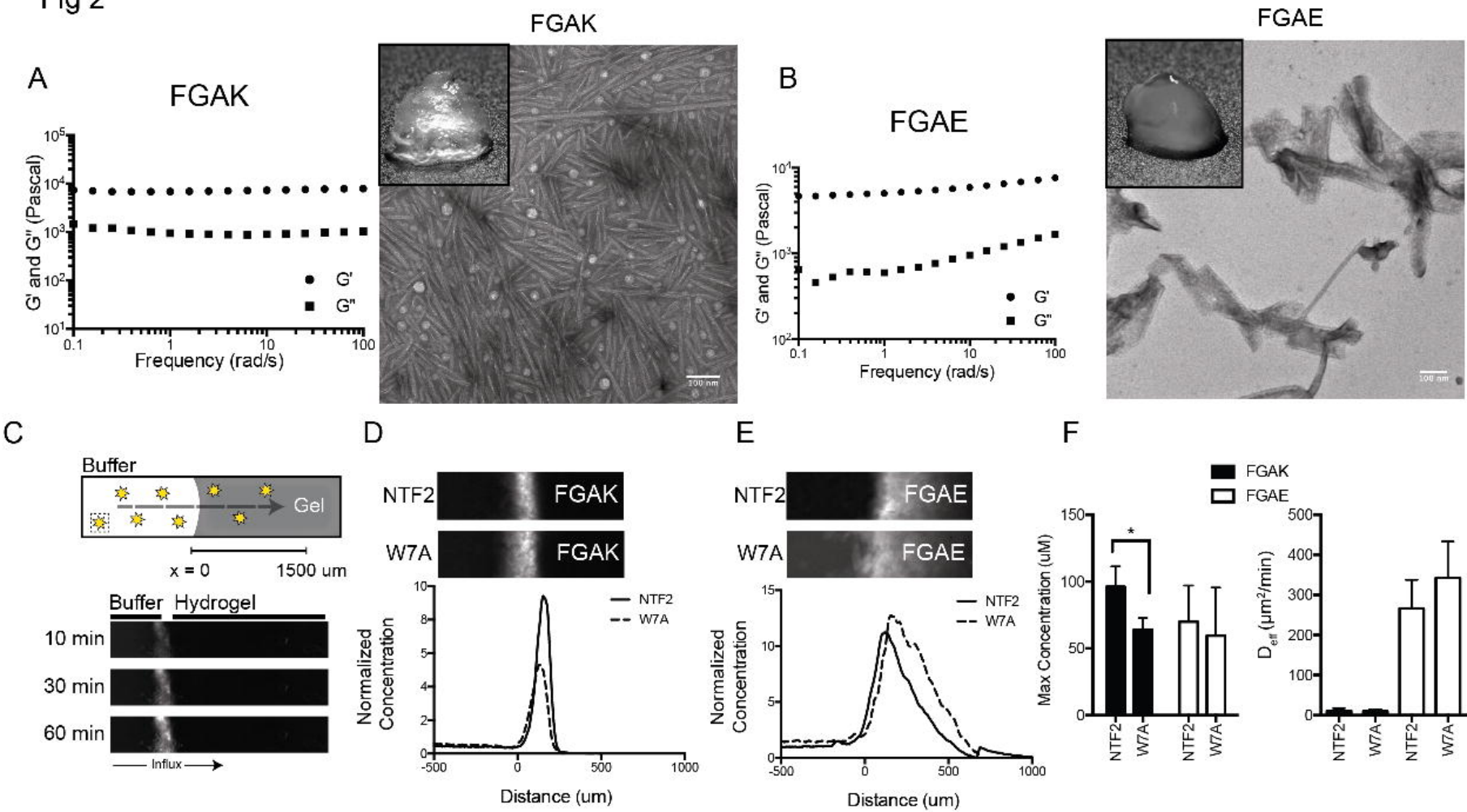

A
FGAK
G' and G'' (Pascal)
Frequency (rad/s)
G'
G''
FGAK
B
FGAE
G' and G'' (Pascal)
Frequency (rad/s)
G'
G''
FGAE
C
Buffer
Gel
x = 0
1500 um
Buffer
Hydrogel
10 min
30 min
60 min
Influx
D
NTF2
FGAK
W7A
FGAK
Normalized Concentration
NTF2
W7A
Distance (um)
E
NTF2
FGAE
W7A
FGAE
Normalized Concentration
NTF2
W7A
Distance (um)
F
FGAK
FGAE
Max Concentration (uM)
NTF2
W7A
NTF2
W7A
D_{eff} (μm^2/min)
NTF2
W7A
NTF2
W7A

Fig 3

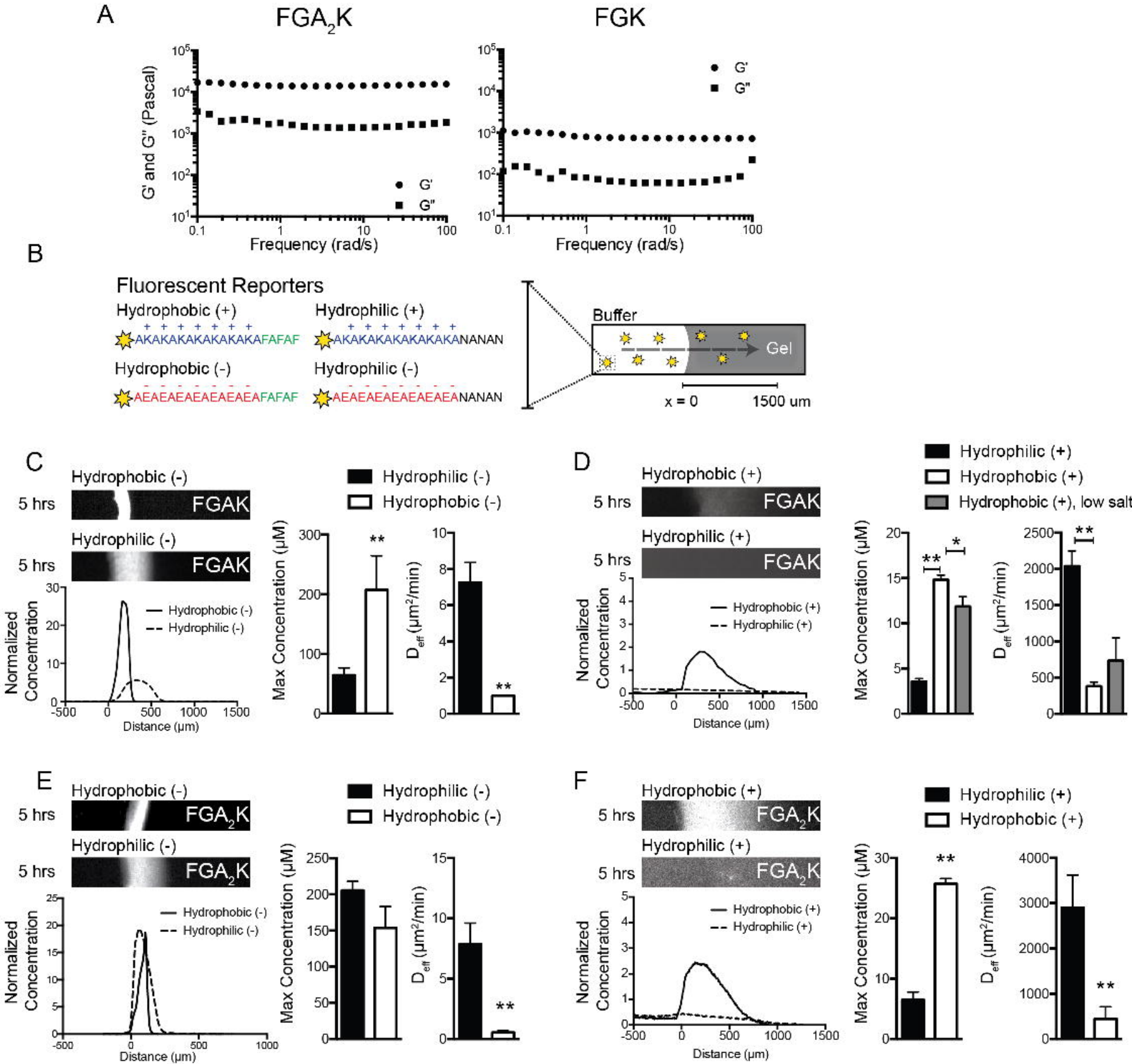

A
FGA2K
FGK
G' and G'' (Pascal)
Frequency (rad/s)
G'
G''
B
Fluorescent Reporters
Hydrophobic (+)
AKAKAKAKAKAKAKAFAFAF
Hydrophilic (+)
AKAKAKAKAKAKAKANANAN
Hydrophobic (-)
AEAEAEAEAEAEAEAFAFAF
Hydrophilic (-)
AEAEAEAEAEAEAEANANAN
Buffer
Gel
x = 0
1500 um
C
Hydrophobic (-)
5 hrs
FGAK
Hydrophilic (-)
5 hrs
FGAK
Normalized Concentration
Hydrophobic (-)
Hydrophilic (-)
Distance (μm)
Hydrophilic (-)
Hydrophobic (-)
Max Concentration (μM)
D_eff (μm²/min)
D
Hydrophobic (+)
5 hrs
FGAK
Hydrophilic (+)
5 hrs
FGAK
Hydrophobic (+)
Hydrophilic (+)
Hydrophilic (+)
Hydrophobic (+)
Hydrophobic (+), low salt
E
Hydrophobic (-)
5 hrs
FGA2K
Hydrophilic (-)
5 hrs
FGA2K
F
Hydrophobic (+)
5 hrs
FGA2K
Hydrophilic (+)
5 hrs
FGA2K

Fig 4

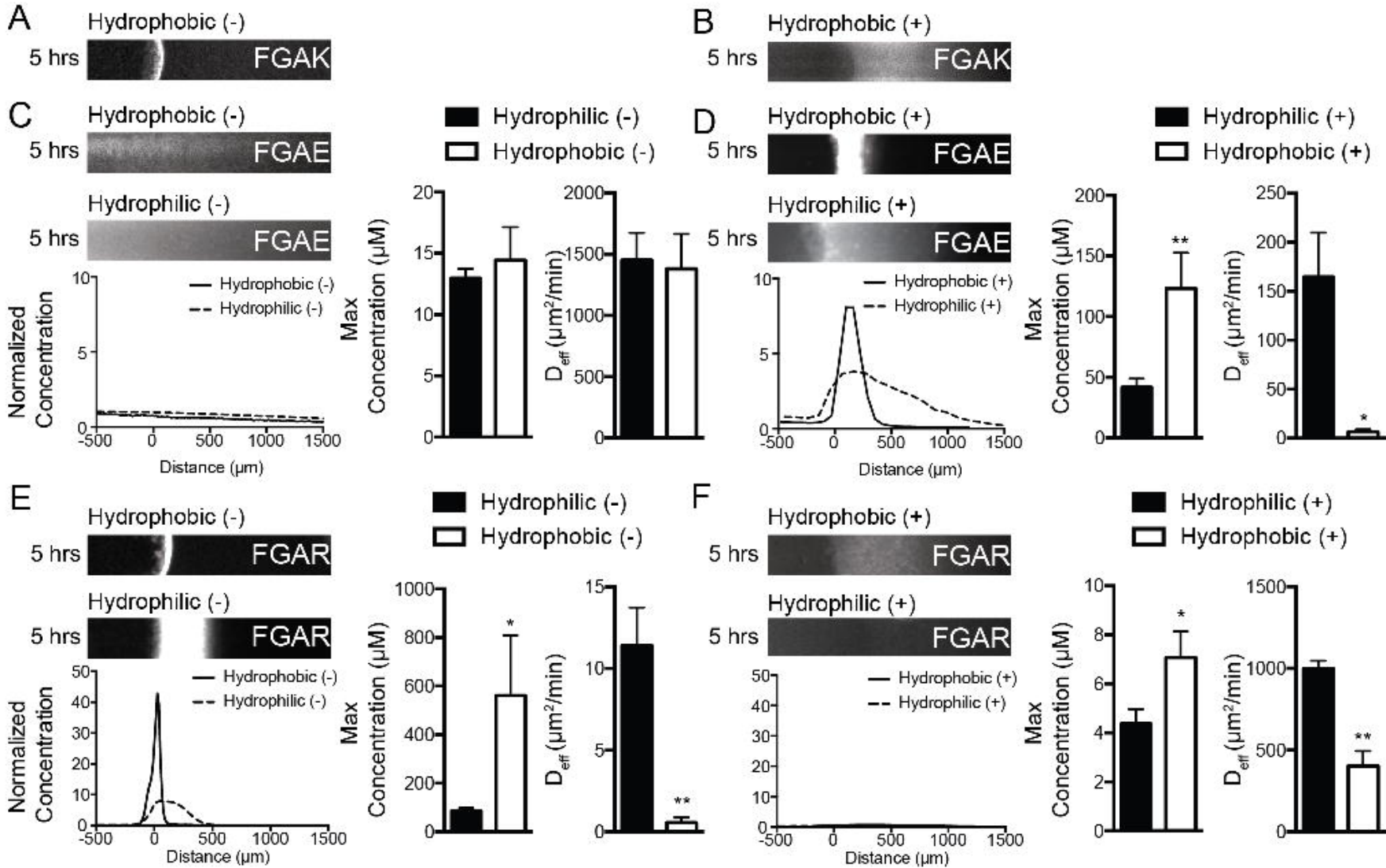

A
Hydrophobic (-)
5 hrs
FGAK
B
Hydrophobic (+)
5 hrs
FGAK
C
Hydrophobic (-)
5 hrs
FGAE
Hydrophilic (-)
5 hrs
FGAE
Normalized Concentration
Hydrophobic (-)
Hydrophilic (-)
Distance (μm)
Hydrophilic (-)
Hydrophobic (-)
Max Concentration (μM)
D_eff (μm²/min)
D
Hydrophobic (+)
5 hrs
FGAE
Hydrophilic (+)
5 hrs
FGAE
Hydrophobic (+)
Hydrophilic (+)
Distance (μm)
Hydrophilic (+)
Hydrophobic (+)
Max Concentration (μM)
D_eff (μm²/min)
E
Hydrophobic (-)
5 hrs
FGAR
Hydrophilic (-)
5 hrs
FGAR
Normalized Concentration
Hydrophobic (-)
Hydrophilic (-)
Distance (μm)
Hydrophilic (-)
Hydrophobic (-)
Max Concentration (μM)
D_eff (μm²/min)
F
Hydrophobic (+)
5 hrs
FGAR
Hydrophilic (+)
5 hrs
FGAR
Hydrophobic (+)
Hydrophilic (+)
Distance (μm)
Hydrophilic (+)
Hydrophobic (+)
Max Concentration (μM)
D_eff (μm²/min)

Fig 5

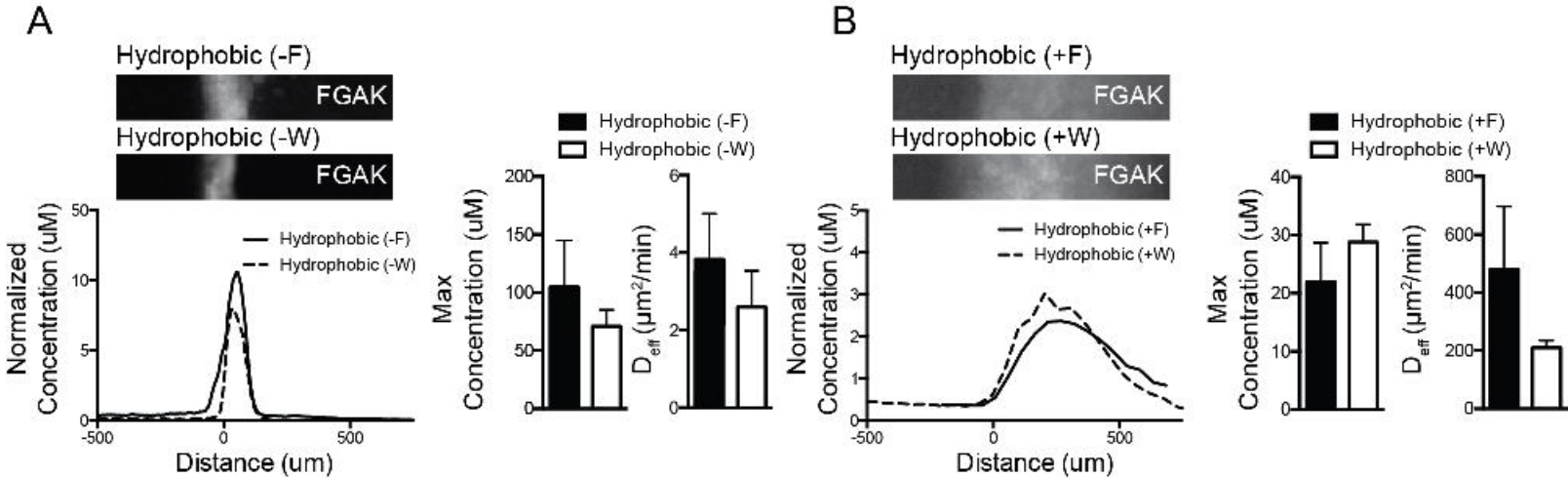

A
Hydrophobic (-F)
FGAK
Hydrophobic (-W)
FGAK
Normalized Concentration (uM)
Hydrophobic (-F)
Hydrophobic (-W)
Distance (um)
Max Concentration (uM)
Hydrophobic (-F)
Hydrophobic (-W)
D_eff (μm²/min)
B
Hydrophobic (+F)
FGAK
Hydrophobic (+W)
FGAK
Normalized Concentration (uM)
Hydrophobic (+F)
Hydrophobic (+W)
Distance (um)
Max Concentration (uM)
Hydrophobic (+F)
Hydrophobic (+W)
D_eff (μm²/min)

Fig 6