Understanding the selective permeability of biological hydrogels

By
Jacob Julian Seid Witten
B.A. Biochemistry & Biophysics and Mathematics
Amherst College, 2014

Submitted to the Program of Computational and Systems Biology in Partial Fulfillment of
the Requirements for the Degree of Doctor of Philosophy in Computational and
Systems Biology

at the
Massachusetts Institute of Technology
June 2019

Signature redacted

Signature of author..

Computational and Systems Biology Graduate Program

March 27, 2019

Signature redacted

Certified by...

Katharina Ribbeck
Hyman Career Development Professor
Thesis Supervisor

Signature redacted

Accepted by...

Christopher Burge
Professor of Biology
Director, Computational and Systems Biology Graduate Program
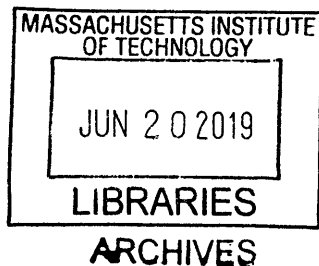
1

Understanding the selective permeability of biological hydrogels

By
Jacob Julian Seid Witten

Submitted to the Program of Computational and Systems Biology in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Computational and Systems Biology

Abstract

Biological hydrogels are fundamental to life, from microbial biofilms to mucus and the nuclear pore in humans. These hydrogels exhibit complex selective permeability behavior, allowing the passage of some particles while blocking the penetration of others. This selective permeability is critical for understanding the biological and medicinal impact of mucus, which coats all non-keratinized epithelia in the body. Mucus controls the penetration of microbes, pollutants, and nanoparticles through a combination of steric and interactive (binding-based) constraints. For small molecules, binding to mucus and in particular mucin, the main gel-forming component of mucus, affects diffusive permeability and may also affect a molecule's biological or therapeutic activity. However, the molecular characteristics leading to mucus binding are not well understood. I therefore developed a mucus binding assay with substantially greater throughput than any existing assay, and combined it with a mucin binding screen to identify a new motif as associated with binding to mucin. I also validate the link between binding to mucin and reduced activity in mucin for the antibiotic colistin. Next, I applied my binding technique to study the binding of a wide range of antibiotics and inhaled drugs to respiratory mucus, and identified previously unknown mucus binding interactions. These binding interactions could impact the activity of the drugs within the mucus or impact their lung residence time in the case of highly muco-obstructive lung diseases. The nuclear pore, which controls the passage of material between the nucleus and the cytoplasm, is similar to mucus in that it too is a selectively permeable network of disordered proteins. Passage through the nuclear pore requires interaction with the network that was initially thought to be purely hydrophobic in character. However, there is evidence that electrostatic interactions also partly govern nuclear pore transport. Here, we apply a peptide-based system to study the interplay of hydrophobic and electrostatic interactions to further dissect the biochemistry underlying nuclear pore function.

Thesis supervisor: Katharina Ribbeck

Title: Hyman Career Development Professor of Biological Engineering

Acknowledgements

I would like to thank my advisor, Prof. Katharina Ribbeck, for helping me become excited and passionate about the unlikely-sounding topic of mucus. Her passion for science and discovery have been an inspiration to me, and her mentorship, support, and fostering of an incredibly positive lab environment have all been instrumental in helping me grow as a scientist.

I would also like to thank my committee, Alan Grodzinsky and Collin Stultz, for their thoughtful suggestions and guidance throughout the last five years. It has been great to get their perspectives on diverse problems and from such different and helpful perspectives. Looking further back, thank you to my undergraduate research advisors, David Hill, Richard Boucher, and Sheila Jaswal, who all supported me greatly and helped me learn what it meant to become a scientist.

Thanks also to the Ribbeck lab, of which literally every member has been outrageously supportive, helpful, interesting, and passionate about science. I have learned a tremendous amount from everyone in the lab and my Ph.D. wouldn't have been close to the incredible experience it was without the Ribbeck lab community. Thanks especially to Tahoura Samad, with whom I have worked on pretty much all of my projects; she has been indispensable as a mentor, a collaborator, and a friend. Particular thanks also to my other collaborators within the Ribbeck lab: Wes Chen (who also helped me get started in the lab), Kate Dupont, Caroline Wagner, Julia Co, Emiko Zumbro, and Gerardo Cárcamo-Oyarce. To the rest of the Ribbeckers old and new, Caroline Werlang, Kelsey Wheeler, Ben Wang, Brad Turner, Miri Krupkin, Chloe Wu, Michaela Gold, Julie Takagi, Erica Shapiro, Nicole Kavanaugh, and Nicole Billings: you have all helped make lab a great place to work every day.

Thank you as well to my parents, Matt Witten and Nancy Seid, who gave me unconditional love and support, and constant support of my curiosity, all of which helped me become who I am today. And to my brother Zack, whose discussions with me over ping pong are probably the defining event for my intellectual development.

To my Amherst friends who have also been in Boston, Chris, Caitlin, Maria, and Tinsae, it's been an amazing 9 years of friendship (It's actually been 9 years! I swear! Our 5th reunion is this June! Wow, we are old) and I've been so grateful to have you all in my life.

Lastly, to a person I suppose I could also technically call a friend from Amherst who went to Boston after graduation, Lisa Witten. Thanks for being a great and supportive friend when I'm feeling down, and a thoughtful scientist who always give me great ideas about everything from experiments to try to how to give a clear presentation. Oh, and also for being the love of my life.

Table of Contents

1 INTRODUCTION

Much of the work presented in this chapter was included in: Witten, J. and Ribbeck, K., 2017. The particle in the spider's web: transport through biological hydrogels. *Nanoscale*, 9(24), pp.8080-8095.[1]

1.1 HYDROGELS ARE UBIQUITOUS SELECTIVE BARRIERS IN BIOLOGY

1.1.1 Hydrogels across organisms

Biological hydrogels, which are composed of hydrated polymer networks, are found throughout every domain of life. For example, both archaea[2] and bacteria can form biofilms, or microbial aggregates surrounded by secreted extracellular polymeric substances (EPS) that act as a protective barrier and create microenvironments within which microbes thrive and adapt to harsh conditions.[3,4] In eukaryotes, hydrogels have evolved to fulfill an astonishingly diverse set of functions. For example, the nuclear pore is a barrier formed from crosslinked intrinsically disordered proteins called nucleoporins, which control the passage of macromolecules between the nucleus and the cytoplasm.[5,6] One of the largest hydrogels in the body is the mucus gel (Fig. 1.1) that lines all wet epithelia and protects the underlying cells against toxins, pollutants, and invading pathogens.[7–11]
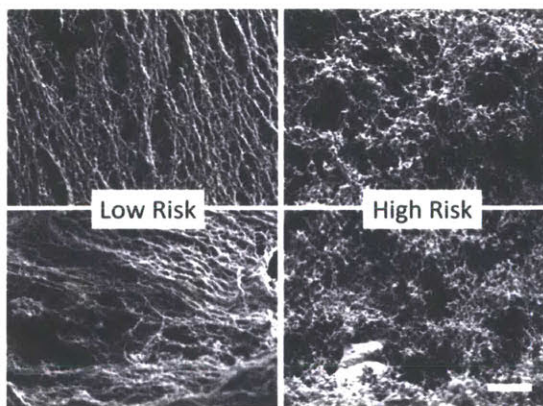

Figure 1.1 Scanning electron microscopy of cervical mucus samples from pregnant patients at low or high risk for preterm birth. Scale bar: 200nm. Reprinted with permission from Critchfield *et al.* (2013).[12] Copyright (2013) PLOS. Published under CC BY license.

1.1.1 Pemeability of hydrogels

The permeability of biological hydrogels poses a challenge for biomedical development, because passive diffusive transport through these gels controls drug delivery and activity in a wide variety of contexts. For example, biofilms are involved in the majority of infections in developed countries; they form on medical implants and wounds, cause middle-ear infections and gingivitis, and more.[13] One driver of antibiotic resistance in biofilms is binding of antibiotics to EPS, which sequesters them and/or reduces their penetration into the biofilm. This sequestration and reduced penetration can lead to sublethal concentrations within the interior of biofilms, preventing the eradication of the biofilm. Mucus is a critical barrier to mucosal administration of drug delivery nanoparticles designed for drug delivery, particularly for inhaled, oral, and vaginal delivery. In these contexts, the interplay of mucus layer turnover and diffusive penetration through the lung, intestinal, and cervical mucus layers respectively determine the extent to which nanoparticles reach the cells they were designed to treat.,[14-19] Likewise, the interaction of mucus with small molecule drugs relying on diffusive transport through mucosal niches can affect their pharmacokinetic and pharmacodynamics properties.[20,21]

More broadly, diffusive transport of nutrients, toxins, inorganic ions, bacterial signaling molecules, oxygen, and more is a critical determinant of the properties of biofilms and mucus.[22-26] Similarly, controlling diffusive transport from the nucleus to the cytoplasm and vice versa is the *raison d'etre* of the nuclear pore and a greater understanding of the mechanisms of transport would be of great value.[6,27-29]

1.1.2 Structure and composition of biological hydrogels

Biological hydrogels are complex molecular assemblies with context-dependent properties. For biofilm EPS, the structure of the matrix varies widely across species, strains, and environments, complicating the characterization of EPS components. However, biofilm EPS commonly contains extracellular DNA and various types of polysaccharides.[3] One well-studied biofilm is that formed by Pseudomonas aeruginosa, an opportunistic pathogen that can cause biofilm-associated infection in wounds[30] or in the lungs of patients with cystic fibrosis (CF).[31] Interestingly, even this single species produces anionic (alginate), cationic (Pel),[32] and neutral (Psl)[33] polysaccharides in differing amounts depending on context and bacterial strain.[34]

In human mucus, the main gel-forming components consist of a family of large, intrinsically disordered secreted glycoproteins called mucins. The main secreted mucins are MUC2, MUC5B, and MUC5AC, although other secreted mucins can be present in smaller amounts as well.[8] Mucins have alternating regions of small, globular hydrophobic domains and highly glycosylated, primarily anionic (due to sialic acid and sulfation) unstructured regions.[8,35,36] Other components of mucus include lipids, soluble proteins and peptides, and nucleic acids.[35] Mucus composition can vary depending on disease state, for example CF, a genetic disease in which improper ion balance results in pathologically thick, dehydrated mucus.[37] This dehydrated mucus is particularly problematic in the lung, where hyperconcentrated airway mucus leads to inflammation and often fatal infection. The presence of necrotic neutrophils in inflamed CF airways results in high levels of free DNA and actin filaments in lung mucus,[38] increasing the

mucus' viscoelasticity[35] and playing important roles in binding to cations. More detailed reviews can be found on composition of biofilms[3,39] and mucus.[7,8,35,40,41]

While biological hydrogels are distinct in terms of their locations and molecular compositions, they share certain common principles that govern selective filtration. Broadly speaking, the transport of a solute through a gel is controlled by the solute's size and its interactions with the components of the gel (Fig. 1.2).[42] Size, or steric, filtering is a universal feature of biological gels, which have a polymeric mesh size that constrains the diffusion of large particles (Fig. 1.2a). Pure steric filtering is an important component of hydrogel selectivity, but it is also crude because it only selects based on one parameter. Filtering based on chemical interactions between solutes and gel components is also a common feature of biological gels. Depending on their chemistry, the gel components interact with solutes across gradients of chemical properties including charge and hydrophobicity, thereby differentially affecting diffusion.[42] A lack of interactions enables unhindered diffusion (Fig. 1.2b), while binding to hydrogels reduces effective solute diffusivity and hinders penetration (Fig. 1.2c). Below, we discuss in more detail the effects on particles of steric hindrance (Section 1.2) and chemical interactions (Section 1.3).

1.2 SIZE-DEPENDENT FILTRATION

Size effects are important for the transport of viruses, bacteria, eukaryotic cells, particulate pollutants, nanoparticles, and any other particle on the same length scale as the gel mesh. In this section, we discuss size filtering while assuming that all particles are inert to (do not interact with) gel components. To a first approximation, steric interactions are quite simple, corresponding in the macroscopic world to the fact that an
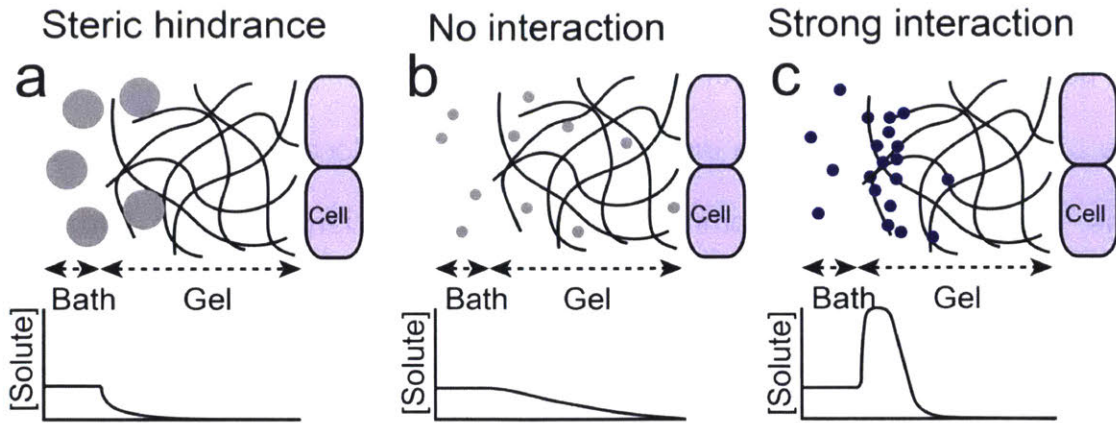
Figure 1.2 Effects of steric hindrance and chemical interactions on gel penetration. a) Particles above the mesh size are unable to penetrate the gel, even if they do not interact with the gel. b) Small inert particles penetrate gels. c) Binding to the gel causes enrichment of solute at the bath-gel interface but slowed gel penetration.

elephant, but not a fly, can be stopped by a chain-link fence, while a tennis ball can pass through with some prodding.

In biological hydrogels, the microscopic mesh is formed by entangled and crosslinked polymers. The distance between adjacent links in the chain-link fence corresponds to the mesh size; objects much smaller than the mesh size diffuse at a rate corresponding to the viscosity of the interstitial fluid (the fluid between fibers; Fig. 1.3a), objects on the order of the mesh size are obstructed but not completely stopped (Fig. 1.3b), and objects much larger than the mesh size are trapped (Fig. 1.3c). This steric barrier is important for mucus because it blocks and/or traps large pollutants and potential pathogens, thus allowing some mechanism, such as mucociliary clearance, to clear the invading particle before it can reach the epithelium.[9] Mesh size is also an important consideration for nanoparticle design, as it sets a maximum possible size for a nanoparticle that must penetrate a hydrogel barrier. The mesh sizes of mucus and
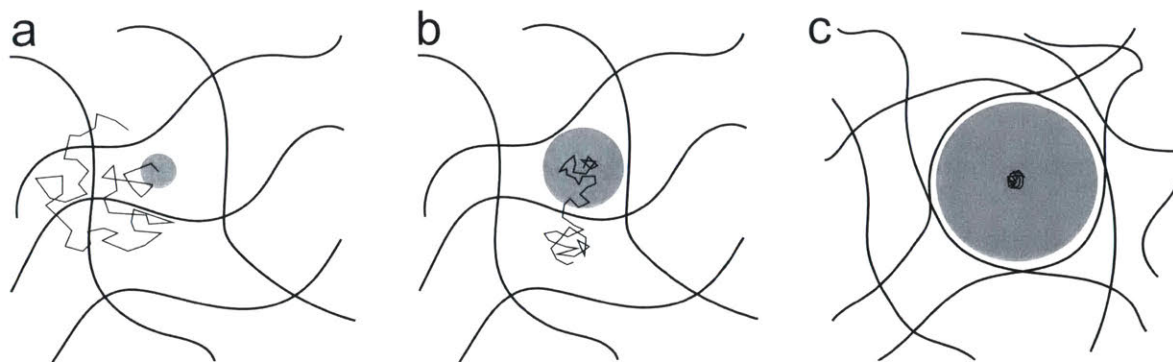
Figure 1.3 Steric effects on diffusion in gels. a) Particles smaller than the mesh size diffuse freely in interstitial fluid. b) Particles on the order of the mesh size have significant steric hindrance but eventually penetrate gels. c) Large particles are trapped.

biofilm EPS vary and are on the order of 10-1000nm.[43–51] Note that measurements of mesh size may be incorrect if the particles used to probe the mesh size interact with the gel, because strong nanoparticle-gel adhesion slows diffusive transport just as surely as steric constraints.[52] As a result, design of non-mucoadhesive "mucus penetrating particles" is an active area of investigation[16,17,53] The nuclear pore also imposes steric constraints on passage: it is filled with intrinsically disordered proteins called nucleoporins that form a mesh with a mesh size ~2.5-5 nm in size that excludes proteins and protein complexes larger than 30-100 kDa, unless they are chaperoned by a nuclear transport receptor (NTR).[54,55]

In addition to the importance of steric constraints for nuclear pore function and for inhibiting mucosally administered nanoparticle delivery, mucus mesh structure may be a useful biomarker for disease states associated with mucosal dysfunction. For example, reduced particle transport in sputum from CF patients is correlated with increased sputum mucin and DNA concentrations, increased disulfide bonding, and reduced lung function.[56] This suggests that particle transport may be a good *in vitro* test for mucus

pathologies. Indeed, recently particle transport was used to examine the effectiveness of disulfide bond reducing mucolytics: in part using passive particle diffusion analysis, Ehre *et al.* showed that a new thiol-based molecule can disrupt CF mucus' gel structure by reducing disulfide bonds.[57] This is a promising approach to normalizing CF mucus or even other lung diseases involving oxidative stress including asthma and chronic obstructive pulmonary disease (COPD), because the oxidative stress drives increased disulfide crosslinking between mucin molecules, which increases lung mucus' viscoelasticity and reduces clearance.[58]

Analysis of mucus gel structure by way of particle transport is not limited to CF, however. For example, an inner colonic mucus layer is believed to prevent contact between the colonic epithelium and bacteria. However, in the inflammatory disease ulcerative colitis, bacteria and bacteria-sized nanoparticles can penetrate this mucus layer to the epithelium. This enhanced penetration is likely due to a compromised mucus mesh.[59] Particle diffusion is also altered in the intestinal mucus of a mouse model of Hirschprung's disease.[60]

Finally, in Appendix B, we present Smith-Dupont *et al.*, where we analyze the permeability of cervical mucus using multiple metrics including particle tracking microrheology. We show that cervical mucus from pregnant patients is associated with reduced particle mobility compared to cervical mucus from ovulating non-pregnant patients, suggesting a tighter gel network in the pregnant case.[61] This is likely because cervical mucus in pregnancy acts to block bacterial penetration and intro-amniotic infection, while ovulatory mucus is structurally reorganized to allow the penetration of sperm.[61]

1.3 INTERACTIVE FILTERS

1.3.1 Overview: effects of chemical interactions on gel penetration and retention

In addition to steric effects, biological hydrogels selectively interact with foreign particles based on the solutes' chemical properties. These interactions often lead to reversible binding, which may be desirable, undesirable, or neutral depending on context. Since reversible binding improves retention within a gel, it may enable longer drug action, as in mucoadhesive drug delivery systems.[10,62] Similarly, binding of native immune factors to cervical mucus may promote their retention at high concentration and prevent microbial penetration through the cervical mucus plug during pregnancy, which could lead to intrauterine infection and preterm labor.[41] On the other hand, reduced diffusivity may slow the penetration through a gel (Fig. 1.2c), impacting the oral bioavailability of drugs or microbial killing by antibiotics. Similarly, binding can sequester a molecule and decrease its effective concentration. Finally, diffusivity could simply be a tunable regulatory parameter, neither intrinsically "good" nor "bad." Whatever the effect of gel binding, unlocking the forces underlying it is crucial to understanding the function of native systems and optimizing drug function. Experimentally, the most common techniques used to measure the diffusion coefficients of small solutes such as drugs and proteins are bulk diffusion assays, fluorescence recovery after photobleaching (FRAP), and fluorescence correlation spectroscopy (FCS).[24,51,63]

To clarify what is meant by slowing solute penetration and improving retention, we briefly present the standard model for diffusion of a solute with first-order reversible binding to a gel. Under the assumptions that 1) binding and unbinding are fast, such

13

that the solute reaches local equilibrium between bound and free states at each point, 2) bound solute is immobile, 3) binding sites within the gel are far from saturated, and 4) binding to the gel does not disrupt gel structure, we can define the *effective* diffusivity D_{eff} as:

$$D_{eff} = D_F (1 + N_T/K_D)^{-1} \tag{1.1}$$

where D_F is the diffusion coefficient of the free solute in the gel (which may be lower than the diffusivity in water, due to steric constrains or increased interstitial fluid viscosity), N_T is the total binding site density in the gel, and K_D is the dissociation constant.

The timescale for a solute to penetrate a gel (or equivalently, the timescale of escape from a gel) τ_{lag} is related to the D_{eff} by:

$$\tau_{lag} \sim L^2/D_{eff} \tag{1.2}$$

where L is the length scale of the gel. These two equations show how increasing solute-gel binding strength or gel binding site density increases τ_{lag}, thus slowing penetration or improving retention.[64] Fig. 1.4 shows this effect in action, as the penetration of tobramycin (an aminoglycoside antibiotic) into a *P. aeruginosa* biofilm is hindered by binding of tobramycin to biofilm components.[65]

Given these results, it is no surprise that specific protein-ligand interactions are commonly exploited to regulate transport and retention within a gel, often for protective purposes or to control signaling.[66,67] For example, sialic acid-terminated O-linked oligosaccharide chains on secreted mucin molecules, particularly Muc5AC, act as decoy receptors for sialic acid targeting bacteria and viruses, including the influenza virus.[68] This interaction slows penetration of sialic acid targeting bacteria and viruses
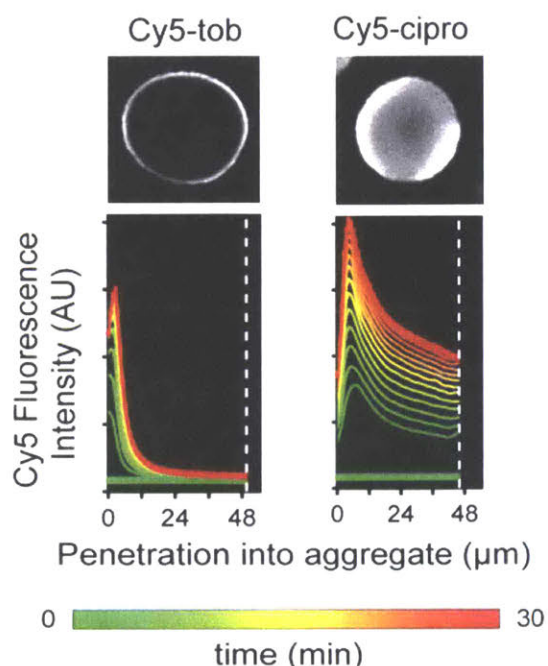
14

Figure 1.4 Penetration of Cy5-labeled (fluorescent) tobramycin and ciprofloxacin into *P. aeruginosa* biofilms. Plots represent quantified timecourses of penetration, with representative images shown above plots. Cy5-tobramycin penetrates the gel less extensively than Cy5-ciprofloxacin. Adapted with permission from Tseng *et al.* (2013).[65] Copyright (2013) Society for Applied Microbiology and John Wiley and Sons Ltd.

through the mucus layer to the vulnerable epithelium[69,70] and competitively inhibits binding to cell-associated sialic acid.[68] A mouse model with lung Muc5AC overexpression has increased resistance to influenza, illustrating this protective function of lung mucus.[71] In addition to non-covalent protein-ligand interactions, covalent linkage to biological hydrogels is also possible; for example, certain mucoadhesive formulations use free thiol groups to form disulfide bonds to mucin.[72]

1.3.2 Biochemistry of reversible binding to the gel matrix

Reversible binding to biological hydrogels is important for molecules throughout the size spectrum;[24] even proton transport can be slowed by gastric mucus, which

protects epithelial tissue from stomach acid.[73] Here, we discuss representative examples that illustrate the importance of various intermolecular forces, as well as several cases showing how modulating gel binding could deliver novel approaches to fighting infections.

The biochemistry underlying reversible binding to gels is difficult to study for several reasons. First, crystal structures of gel components do not usually exist since gel polymers are generally intrinsically disordered. It is also often not obvious to which component(s) of the gel a solute binds, since biological gels contain a complex mix of components.[65] Despite the lack of structural information, existing literature on gel binding and the composition of biological hydrogels yields useful predictions and conclusions about transport in gels; these predictions usually center on how charge and hydrophobicity influence diffusion. Other intermolecular forces such as van der Waals forces and hydrogen bonding are important, but their effect on binding is not easily predictable. In one case, research has shown that hydrogen bonding is important for mucoadhesion of hydrophilic polymers such as poly(acrylic acid).[62] In summary, analysis of a molecule's overall chemical properties may predict that the molecule will bind to *something* within a gel.

An instructive instance in which charge interactions are important is *P. aeruginosa* lung infections in CF patients, which involve two biological gels: mucus and biofilm EPS. One primary class of antibiotics used to treat *P. aeruginosa* infections, aminoglycosides, are highly positively charged and thus bind to polyanions such as mucins, DNA, and actin, which are present at pathologically high levels in CF respiratory mucus. This binding impedes antibiotic penetration and inhibits activity.[74,75]

The *P. aeruginosa* EPS acts as a second protective barrier for biofilm-associated cells. The EPS also contains DNA and, in the case of mucoid strains, high levels of the polyanionic aminoglycoside-binding alginate, although alginate surprisingly seems to not significantly hinder tobramycin penetration.[76–78]

This effect of electrostatic interactions with polyanions on cationic antibiotics is a clinically important topic still under investigation. We address the binding of polymyxins, another class of cationic antibiotics used to treat *P. aeruginosa* lung infections in CF, in chapters 2 and 3, and Appendix A, of this thesis. In addition, in Smith-Dupont *et al.* (Appendix B), we analyze the permeability of pregnant patients' cervical mucus to cationic peptides and anionic peptides. We show that cationic and, surprisingly, sometimes even anionic peptides interact with cervical mucus and that these interactions vary with risk of preterm birth.[61] Thus, gel biochemical properties independent of mesh size effects may also be a source of biomarkers for health and disease.

Although charge-mediated binding to polyanions is important for cationic antibiotics and peptides, reversible gel binding does not seem to be a major factor in many other antibiotic treatments of biofilms.[79] For example, Fig. 1.4 shows that ciprofloxacin easily penetrates a model *P. aeruginosa* biofilm. We likewise show in chapter 3 and Appendix A that ciprofloxacin does not bind to mucus. In addition to slowing antibiotic penetration, biofilms have other resistance mechanisms including slow metabolism and hypoxia in the biofilm interior (which are related to reaction-diffusion dynamics of nutrients and oxygen, emphasizing the broad importance of diffusion in gels), and slow-dividing "persister cells."[13,80,81] The various resistance

mechanisms may be cooperative; slowed penetration could allow more time for bacterial adaptations to antibiotic treatment.[82]

The importance of hydrophobic interactions in small molecule transport has mainly been studied in the context of oral drug bioavailability, which is partially dependent on the ability to diffuse through intestinal mucus to the epithelium. Many hydrophobic drugs such as paclitaxel, testosterone, and cinnarizine bind to exposed hydrophobic sites on mucin, lipids present in mucus, or both; binding slows penetration and renders intestinal mucus a barrier to absorption into the bloodstream.[20]

However, for many molecules it is not immediately clear whether hydrophobic or electrostatic interactions drive interactive properties. For example, cationic antimicrobial peptides (CAMPs), an important class of antibiotics and immune molecules, contain both cationic and hydrophobic residues; the charge, or the hydrophobicity, or both, could lead to gel binding. EPS-CAMP interactions are particularly important, as many antimicrobial peptides are under investigation as possible anti-biofilm treatments, but binding to EPS components often inhibits their effectiveness.[83–85] For example, the Deber group designed a set of synthetic anti-*P. aeruginosa* CAMPs and investigated their interaction with alginate.[86–89] While alginate is anionic, highly soluble, and contains no large hydrophobic domains, alginate blocked the penetration of synthetic anti-*P. aeruginosa* CAMPs when a hydrophobicity threshold was exceeded.[88] Furthermore, they showed that this interaction was mediated by a combination of electrostatic and hydrophobic interactions with alginate.[86] Similarly, both synthetic and natural CAMPs have been shown to bind mucins, DNA, and F-actin, and stronger binding correlates

with greater inhibition of CAMP activity against pathogenic bacteria in saliva and CF sputum.[90–93]

For small molecule drugs, the situation is often even less clear than it is for the peptides discussed above. While the behavior of purely hydrophobic drugs such as testosterone or cationic drugs such as tobramycin are relatively well understood, the vast majority of drugs do not have such extreme physicochemical properties. In Chapters 2 and 3, we consider the question of small druglike molecule binding to mucus.

1.3.3   The effect of nanoscale heterogeneity on solute-gel interactions

While knowing the composition of a gel gives an overall sense of the gel's net charge and helps predict how solutes behave, it is also important to understand how nanoscale molecular heterogeneity, or spatially varying chemical properties, tunes solute-gel interactions. Heterogeneous surfaces are commonly found in gel-related molecules: protein surfaces and antimicrobial peptides contain charged, neutral hydrophilic, and hydrophobic residues, and larger objects such as viruses have complex surfaces (Fig. 1.5a).[94] In addition, the gels themselves often have locally varying properties that render simplistic models of gel charge problematic.

As an example of the effect of nanoscale heterogeneity on the solute side, when Li *et al.* examined the diffusion of two peptides, with the same near-neutral net charge but a different arrangement of these charges, these peptides diffused differently into a reconstituted mucin gel. The block-charge peptide interacted weakly with mucin, while the alternating-charge peptide did not (Fig. 1.5b).[95] This result is indicative of how variations in molecular structure beyond simple net charge affect transport properties in
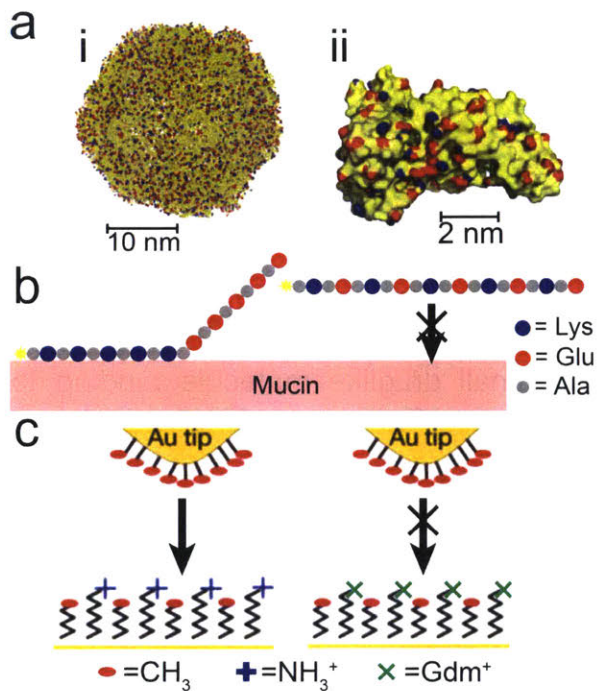
Figure 1.5 Nanoscale heterogeneity of solutes and effect on gel diffusion. a) Surfaces of (i) human rhinovirus (PDB: 2rm2) and (ii) human albumin (PDB: 2bxi), with positive charges in blue and negative charges in red. Reprinted with permission from Cone (2009),[94] (Copyright (2009) Elsevier). Rhinovirus and albumin are densely coated with opposing charges. b) Two fluorescently labeled peptides with the same net charge but different spatial arrangement. The "block" peptide at left interacts weakly with mucin while the "alternating" peptide at right does not;[95] schematic shows potential mechanism for this difference. c) Effect of immobilized charges on nearby hydrophobic interactions, probed by measuring the adhesion of a hydrophobically functionalized gold (Au) atomic force microscopy tip to surface monolayers. Amine (NH_3^+) groups strengthen hydrophobic interactions between these hydrophobic surfaces, while guanidinium (Gdm^+) groups weaken or eliminate them. Adapted by permission from Macmillan Publisher Ltd: Nature (Ma *et al.*, 2015).[96] Copyright (2015).

gels, which could have applications for drugs whose penetration is inhibited by gel binding. Electrostatic and hydrophobic interactions can also alter each other where both are present (Fig. 1.5c), such as in the case of molecularly complex solute diffusion in hydrogels.

The interplay between charge and hydrophobicity is likely also important in the nuclear pore complex, which contains a central channel filled with intrinsically disordered proteins called nucleoporins. These nucleoporins contain many instances of short hydrophobic repeats called FG repeats, which likely act as crosslinkers. The nuclear pore inhibits nuclear translocation of proteins larger than roughly 30-100kDa and even some smaller proteins such as histones.[54,97] However, NTRs efficiently carry large cargo through the nuclear pore.[55] This transport ability of NTRs derives in large part from exposed hydrophobic patches on NTR surfaces, which have hydrophobic and/or π-π interactions with the FG repeats on nucleoporins.[98] The mechanism by which these interactions mediate selective transport is still under debate, and multiple models exist. Of relevance is the fact that although hydrophobic interactions with FG repeats are required for transport through the nuclear pore, charge is important as well. Our group has found that nucleoporins and NTRs tend to have complementary net charges (positive and negative respectively), thus implying that charge assists in the recruitment and translocation of NTRs;[99] this hypothesis is supported by theoretical[100] and structural[101] analysis. In Chapter 4, we discuss the interplay of charge-based and hydrophobic effects using a simplified peptide-based model for FG repeats.

1.4 THESIS OVERVIEW

This thesis details progress on understanding gel permeability and binding. In Chapter 2, I focus on identifying patterns behind the selective interaction properties of mucins. To do this, I integrate equilibrium dialysis (ED) with quantitative liquid chromatography-mass spectrometry (qLCMS) to increase the throughput of small molecule-mucus binding measurements. Combined with a small molecule microarray

screen for mucin-binding molecules, I identify 2,4-diaminopyrimidine as a motif associated with binding to mucins and other mucus-associated polyanions. I also examine how different molecules' charge and hydrophobicity alter binding to mucus, and how ionic strength and pH may affect binding.

In Chapter 3, I extend my ED experiments to other mucus models and different types of functional molecules. In particular, I introduce mucus secreted by human bronchial epithelial cells (as opposed to the purified pig gastric MUC5AC used in Chapter 2) as another lung mucus model. I measure the binding of this mucus to a wide variety of antibiotics, as well as to a variety of inhaled drugs. In appendix A, Tahoura Samad (the first author of the paper presented in appendix A) showed that mucins and mucus confer protection against antibiotics to *P. aeruginosa*. I combined my binding experiments with her microbial work to show that antibiotic-mucin binding explains some, but not all, of this protection.

In Chapter 4, I consider transport in the nuclear pore. As a co-first author with Wesley Chen, we use a hydrogel-forming peptide-based model for FG repeats to examine the diffusion of proteins and charged and/or hydrophobic probes in these hydrogels. We show how varying the spatial location of different charged and hydrophobic groups alters binding of probes to FG repeat peptides and the subsequent effect on diffusion. In Appendix B, Smith-Dupont et al., I also modeled peptide diffusion, the time for the experiments described above where we showed that cervical mucus' diffusive permeability to peptides stratifies patients for risk of preterm birth.

Finally, Chapter 5 contains the conclusions from this work and some discussion of potential future work.

2 MOLECULAR CHARACTERIZATION OF MUCUS BINDING

The work presented in this chapter was published in the journal *Biomacromolecules*. The authors were Jacob Witten, Tahoura Samad, and Katharina Ribbeck.[102] It represents the combined work of all co-authors on the paper.

2.1 ABSTRACT

Binding of small molecules to mucus membranes in the body has an important role in human health, as it can affect the diffusivity and activity of any molecule that acts in a mucosal environment. The binding of drugs, and of toxins and signaling molecules from mucosal pathogens, is of particular clinical interest. Despite the importance of mucus-small molecule binding, there is a lack of data revealing the precise chemical features of small molecules that lead to mucus binding. We developed a novel equilibrium dialysis assay to measure the binding of libraries of small molecules to mucin and other mucus components, substantially increasing the throughput of small molecule binding measurements. We validated the biological relevance of our approach by quantifying binding of the antibiotic colistin to mucin, and showing that this binding was associated with inhibition of colistin's bioactivity. We next used a small molecule microarray to identify 2,4-diaminopyrimidine as a mucin binding motif, and confirmed the importance of this motif for mucin binding using equilibrium dialysis. Furthermore, we showed that for molecules with this motif, binding to mucins and the mucus-associated biopolymers DNA and alginate is modulated by differences in hydrophobicity and charge. Finally, we showed that molecules lacking the motif exhibited different binding trends from those containing the motif. These results open up the prospect of routine testing of small molecule binding to mucus, and optimization of drugs for clinically relevant mucus binding properties.

2.2 INTRODUCTION

Mucus is a viscoelastic hydrogel layer that coats every non-keratinized epithelial surface of the body, including the lungs and intestines.[36] Its activity has important ramifications for biology and medicine:[1,17,103,104] the selective binding of small molecules to mucus may affect the activity and diffusivity of any molecule acting in a mucosal environment.[1] Medically relevant molecules affected by mucus include bacterially produced toxins and signaling molecules, and inhaled therapeutics for diseases including cystic fibrosis (CF), chronic obstructive pulmonary disease (COPD), and asthma.[1,17]

The main gel-forming components of mucus are large secreted glycoproteins called mucins, which contain alternating regions of globular hydrophobic domains and disordered polyanionic oligosaccharide brushes with sialic acid and sulfated groups.[36,105] Since mucins are a major component of mucus, the ability of a molecule to bind to mucin is a key determinant of mucus binding. Mucin overproduction is a hallmark of respiratory diseases including CF, COPD, and asthma,[106] suggesting that mucin binding is correspondingly more important in these pathologies. Synthetically crosslinked, purified mucin gels are a second environment in which mucin binding plays a role; they are a promising biomaterial for sustained drug delivery, and mucin binding helps determine the kinetics of drug release.[107] Mucins can also nonenzymatically catalyze some organic reactions by binding to relevant molecules.[108]

Besides mucins, lung mucus contains lipids, proteins, extracellular DNA (eDNA), and bacterial polysaccharides (e.g. the polyanion alginate, often present in infections by the *Pseudomonas aeruginosa*, an opportunistic pathogen[109]). Both eDNA and bacterial

polysaccharides are commonly found in CF lung mucus and may also bind small molecules.[1]

Despite the clinical importance of small molecule binding to mucus, relatively little is known about the chemical properties that regulate small molecule binding. What is known has largely been inferred from diffusion measurements: slower diffusion in mucus indicates increased mucus binding, since small molecules are smaller than the mesh size of mucus.[17] In these experiments, increased hydrophobicity typically correlates with slower drug diffusion.[20] Lipids, and the hydrophobic domains of mucins, are believed to account for this slowed diffusion by binding to hydrophobic drugs[1,17,104,110] although the glycans grafted to mucins may also play a role.[108] Drug charge only appears to play a major role for highly cationic molecules such as aminoglycoside and polymyxin antibiotics (+5 molecular charge in the case of the anti-*P. aeruginosa* antibiotics tobramycin and colistin).[21,75,111-114] This binding is believed to be mediated by the polyanionic glycosylated domains of mucins and, in CF, eDNA and bacterial polysaccharides.[1,104]

However, beyond an apparent correlation between diffusion and hydrophobicity, and electrostatic binding of highly charged molecules, few details are known about which molecules can bind mucus. One primary reason for this knowledge gap is that studies of mucus-small molecule interactions typically use small datasets, with fewer than twenty molecules per study (with one interesting exception that tested 41[115]).[24,116] This low volume of data is due to the difficulty of acquiring large amounts of mucus or mucus-like models. In addition, diffusion and binding measurements have been based on molecular detection techniques such as radiolabeling, fluorescence, or

27

electrochemical methods,[24] which cannot be easily multiplexed. Thus, a data-driven approach to identifying molecules that bind mucus has so far not been feasible.

Our goal in this work is to overcome this limitation by analyzing the binding of small molecules to mucus components, including MUC5AC, DNA, and alginate, in a relatively high throughput manner. To do this, we developed an equilibrium dialysis (ED) technique to measure binding of small molecules to these biopolymers. This label-free technique can measure up to 20 molecules in a single experiment, minimizing sample volume requirements and enabling the generation of a large dataset (96 molecules in total) which constitutes a substantial increase over previous studies of drug-mucus interactions.

First, we validated the biological relevance of our binding assay by showing that the cationic polymyxin antibiotic colistin binds MUC5AC, a prominent lung mucin,[105] but not the neutral polysaccharide methylcellulose (MC), and that binding was associated with inhibition of antibiotic activity. Next, we used a small molecule microarray (SMM) approach[117] to screen for MUC5AC binding of thousands of molecules, which revealed a previously unknown chemical motif associated with mucin binding. We used our novel ED method to confirm the importance of this motif in mucin binding, and further showed that charge and hydrophobicity play a context-dependent role in predicting binding. Together, these findings demonstrate the strength of our technique for enabling a data-driven approach to understanding mucus binding, and illustrate important insights about the effect of mucus on small molecules.

2.3 EXPERIMENTAL

2.3.1 Buffer preparation and specifications.

10x Phosphate-buffered saline (PBS) was purchased from AccuGene. At 1x concentration, it consisted of 1mM KH_2PO_4, 2mM Na_2HPO_4, and 150mM NaCl at pH 7.4. Phosphate-citrate buffer (PCB) at 1x consisted of 10mM each of phosphate and citrate, and 58mM sodium, at pH 7.0.

2.3.2 Mucin preparation and fluorescent labeling.

MUC5AC was purified from fresh pig stomach scrapings as described in Celli *et al*,[118] with the exception that the cesium chloride density gradient ultracentrifugation step was only performed for mucin used in the SMM, not for mucin tested using ED. Briefly, mucus was scraped from fresh pig stomachs (Research 87, Inc., Boylston, MA, USA) and mucins were purified using size exclusion chromatography.

To fluorescently label the mucin, 0.6mg was solubilized in 300μL of pH 9 0.1M sodium bicarbonate buffer overnight at 4°C. Alexa Fluor 647 succinimidyl ester (ThermoFisher Scientific) dissolved in dimethyl sulfoxide (DMSO) at 10 mg/mL was added to the mucin solution to a concentration of 0.02 mg/mL (1:100 mass ratio with respect to the 2mg/mL mucin) and the mixture was incubated at room temperature for 1 hour with shaking. An equal volume of 50mM tris(hydroxymethyl)aminomethane was added to quench the reaction and the mixture was left at room temperature for 5 minutes with shaking. The volume was diluted to 20mL with PBST (PBS with Tween 20) in a Corning Spin-X 100 kDa molecular weight cutoff ultrafiltration concentrator (Fisher) and centrifuged at 3,000xg for 25min at 4°C; the flow-through was discarded. These steps (dilute to 20mL with PBST, centrifuge, discard flow-through) were repeated twice

29

more; the volume remaining after each centrifugation was approximately 5mL. The labeled mucin was then diluted to the appropriate concentration.

2.3.3 Equilibrium dialysis.

ED was performed with a 12 kDa cutoff Rapid Equilibrium Dialysis device (Thermo Scientific) with 100μL of 5mg/mL biopolymer solution or gel, or buffer control, in the sample chamber and 300μL of matching buffer in the assay chamber. Each experiment began with equal concentrations of the relevant compound cocktail loaded in both chambers ("1x concentration"). 1x concentration was 200nM for every molecule except colistin (10μM). Equilibration took place over 4 hours with shaking at 235 rpm and 37 °C, at which point aliquots were collected from the assay chamber. HPLC separation took place on an Acclaim PolarAdvantage column (3μm pores, 2.1x100mm, VWR). The HPLC method is given in Table S1. The first 5 minutes of eluent were diverted to a waste container to avoid salt contamination of the mass spectrometer (Agilent 6410 triple quadrupole), and the rest of the eluent was analyzed using multiple reaction monitoring (MRM). Source parameters were as follows: temperature 350 °C, gas flow 10 L/min, nebulizer 25 psi, capillary voltage 4000V, with molecular ions and fragments for all experiments given in Tables S2-S4. Multiple transitions were tracked for several of the molecules (including colistin) as a consistency check; only one was used for analysis but all gave consistent results (not shown). Peak areas for each molecule's chromatogram were compared to an external standard curve to measure concentrations of each molecule.

The SMM and NIH Clinical Collection (NCC) compound sets (acquired from ChemBridge [San Diego, CA] and NIH respectively) were tested for biopolymer binding

in 3 and 2 groups, respectively, with one exception. The assigned groupings are given in Tables S2-S3 ("Cocktail group 1" and "Cocktail Group 2" respectively). There were 18 compounds in each group (cocktail) for the SMM runs and 20 or 21 compounds in the NCC drug cocktails, except for the experiments with SMM molecules binding to mucin in PCB, where 96 compounds (see main text) were run in 4 groups of 24 molecules each ("Cocktail group 2" in Table S2), of which 54 molecules were successfully measured. These 54 in the 3 groups of 18 were used for all remaining experiments. The NCC consists of over 700 compounds; 41 molecules were selected randomly with the constraint that molecules with the same mass could not be included in a single cocktail, to ease qLCMS analysis. Several DAP-containing molecules were included in the NCC set (they did not bind mucin significantly more than non-DAP containing NCC molecules; not shown).

The uptake ratio R_U of a compound was calculated by (see Supporting Text for derivation):

$$R_U = 1 + 4\left(\frac{1}{c_E} - \frac{1}{c_C}\right) \tag{3}$$

where c_E is the concentration in the assay chamber of the experimental dialysis system and c_C is the concentration in assay chamber of the buffer-buffer dialysis control. c_E and c_C were both scaled with respect to the 1x concentration. c_E and c_C were each averaged over 3 wells for one full biological replicate of R_U, which was then converted to ΔG, so n=3 represents 9 measured wells.

We used a biopolymer concentration of 0.5% w/v (5 mg/mL). 5 mg/mL is consistent with typical mucin concentrations in lung mucus (0.1-20 mg/mL, depending on disease state[105,119–121]) and DNA concentrations in CF sputum (0.5-5 mg/mL[119,120]). It

is higher than the bulk concentration of alginate in *P. aeruginosa*-infected CF sputum (0.004-0.1 mg/mL[122]), but the local concentration in a *P. aeruginosa* biofilm may approach 5 mg/mL. Alginate (MW 120-190 kDa) and calf thymus genomic DNA were purchased from Sigma.

2.3.4 Small molecule microarray.

SMM slides were prepared as previously described in Bradner *et al.*[123] The experimental protocol was also as described in Bradner *et al.* using procedure steps 21A ("Dish method"), 22A ("Direct detection of fluor-labeled proteins"), and 23, with the following modifications: PBST and PBS (for the final rinse) were used instead of TBST and TBS, slides were incubated at room temperature rather than 4 °C, 5mL of protein and wash solution was used rather than 3mL, and 3 washes were performed instead of 2 in step 22A(i). The slides were scanned for Alexa 647-labeled mucin fluorescence using a GenePix scanner (Molecular Devices) and analyzed by GenePix Pro software (Axon Instruments). Each printed feature's signal-background ratio was quantified (on a per-feature basis per Bradner *et al.*) and normalized to a z-score with mean and standard deviation measured for all molecules under that feature's condition (mucin and detergent concentration).

For the initial SMM screen, each molecule appeared 16 times: twice on each slide where present, and under 8 conditions. 6 conditions were experimental: 0.1, 0.5, and 1 µg/mL mucin with 0.05% and 0.2% Tween 20, and 2 conditions were controls: a PBST-only blank (0.2% Tween 20) and a fluorophore-only control (0.05% Tween 20). Spots with a z-score of 2 or more were determined to indicate binding. A molecule was considered a hit if it appeared as positive on 6 or more of the 12 potential experimental

locations but none of the 4 potential control locations. This combination of z-score cutoff and hit calling was selected to yield a manageable number of hits for ED analysis. For the follow-up screen, there was only one condition (0.5 µg/mL mucin, 0.1% Tween 20) but each molecule was present on two slides. A molecule was considered a hit if it had a z-score of 2 or more in all 4 possible locations. For quality control, this screen also included 144 blanks that were DMSO-only negative controls, each spotted at either 8 or 16 locations. None of these negative controls had a z-score greater than 2 in 3 or more locations, meaning that none of the negative controls were false positives (data not shown).

2.3.5 Computational molecular analysis.

Motif analysis was performed using RDKit (RDKit: Open-source cheminformatics; http://www.rdkit.org) for Python. Molecules were checked for presence of DAP using the substructure query "NC1=NC=CC(N)=N1". Fisher's exact test was performed using the SciPy package.

Molecular charge was calculated as the average charge weighted by abundance over all generated protonation microspecies using the MajorMicrospeciesPlugin Calculator Plugin, Marvin 5.4.1, 2011, ChemAxon (http://www.chemaxon.com). ClogP was calculated using BioByte (BioByte Corp, Claremont, CA).

2.3.6 Generation of plots for local and global correlations with molecular properties.

We generated random "ΔG" values using 40 uniformly distributed random numbers evenly spaced between 0 and 10 (0-10 was the arbitrarily scaled x axis), and made the function continuous with a cubic spline interpolation. We generated random

33

"Property" values using 15 normally distributed random numbers evenly spaced between 0 and 10, also with a cubic spline interpolation. The blue and red "molecules," or locations on the x axis, were selected by hand to illustrate that local and global correlations between ΔG and a molecular property may not be consistent. All figure preparation was done using Matlab (2015a, The MathWorks, Inc., Natick, MA, USA).

2.3.7 Colistin activity in mucin.

Colistin MIC was first determined using a modified broth microdilution assay. *P. aeruginosa* strain PAO1 was grown overnight in BBL Mueller-Hinton II broth (MHB; BD Falcon) at 37°C. Colistin (Sigma) stock solution was prepared at 10 mg/mL in water, then serially diluted 2-fold into MHB. The overnight culture was diluted to a final concentration of approximately 3×10^4 CFU/mL into fresh MHB. 90 µl of this inoculum was added to each well of a 96-well plate with 10 µl of the colistin dilution series, such that the final concentrations evaluated were between 1mg/mL and 0.4 µg/mL. After incubation at 37°C for 24 hours, the MIC of colistin in MHB was determined as 8 µg/mL, by visually inspecting the plate to identify the lowest concentration at which there was no visible cellular growth.

To compare the efficacy of colistin in mucin, methylcellulose and MHB alone, PAO1 was grown overnight in MHB at 37°C. The overnight culture was diluted to a final concentration of approximately 6×10^4 CFU/mL into fresh MHB. 45 µl of this inoculum was added to each well of a 96-well plate. 45 µl of either mucin (1% w/v, dissolved in MHB), methylcellulose (1% w/v, dissolved in MHB) or MHB were added to each well depending on the condition, and then 10 µl of colistin diluted into MHB was added to a final concentration was 8 µg/mL (MIC) or 16 µg/mL (2xMIC). After incubation at 37°C

34

for 24 hours, the number of surviving cells in each condition was evaluated by scraping the wells, homogenizing the contents by pipetting up and down vigorously, serial dilution, plating and counting colony forming units (CFUs).

2.3.8 Statistical analysis and plots.

Except for the Fisher's exact test, statistical analysis and plotting was performed in R (R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.) Tests of significant correlation were performed using F tests. Pairwise comparison of regression coefficients was performed by testing the ability of a property to predict the difference in ΔG between two different gels. For example, a test of the model:

lm((Mucin_dG - DNA_dG) ~ Charge,data=all_data)

would test the significance of a pairwise comparison between the effect of charge on mucin and DNA binding. All tests were two-tailed.

For the ED experiments, the binding of each molecule was measured 3 times, as described in "Equilibrium dialysis" section of Methods.

2.3.9 Data availability.

The datasets generated during the current study are available from the corresponding author on request.

2.3.10 Code availability.

The code written and used for data analysis for the current study is available from the corresponding author on request.

2.4 RESULTS AND DISCUSSION

2.4.1 Development and validation of equilibrium dialysis technique to measure biopolymer binding.

In the ED assay, we measured the partitioning of molecules across a dialysis membrane between an assay chamber containing buffer and a sample chamber containing the biopolymer of interest in the same buffer. We used multiple compounds in the same experiment (Fig. 2.1a), under dilute conditions to minimize competition for binding sites. We then used quantitative liquid chromatography-mass spectrometry (qLCMS) to measure molecule concentrations and quantify the equilibrium partitioning. qLCMS quantitation allowed us measure up to 20 molecules in a single experiment.

We could not measure molecular concentrations in the sample chamber because the biopolymers were not compatible with LC. Instead, we observed binding by comparing the concentration of molecules in the assay chamber to the concentration of molecules in the assay chamber of a buffer-buffer control well (Fig. 2.1a, Methods). From these measurements, we calculated the equilibrium uptake ratio R_U, given by:

$$R_U = [Compound]_{sample}/[Compound]_{assay} \tag{1}$$

Log-transforming R_U gives a value akin to an energy:

$$\Delta G_{uptake}/k_B T = -\ln R_U \tag{2}$$

For simplicity, we consider the unitless quantity ΔG scaled by thermal energy $k_B T$, meaning $\Delta G = \Delta G_{uptake}/k_B T$. This quantity is more likely than R_U to scale linearly with
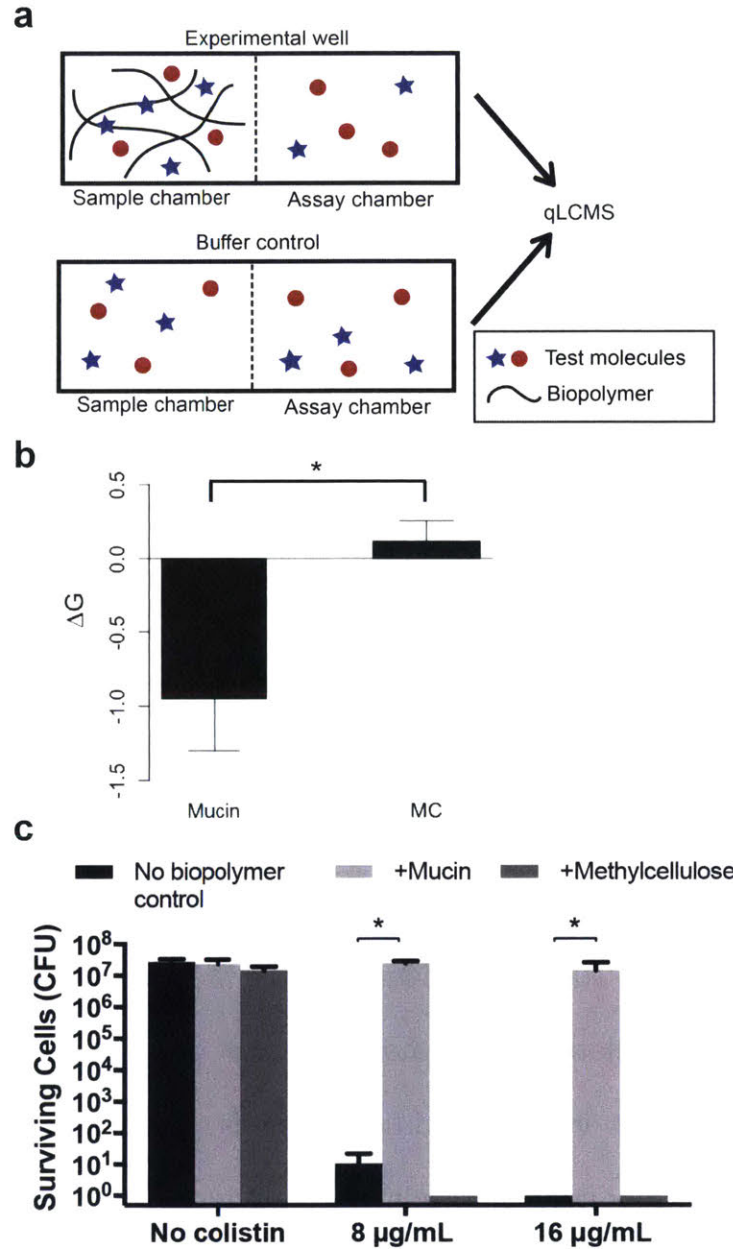
Figure 2.1 (a) Schematic of equilibrium dialysis (ED). Concentration in assay chamber of ED device is compared to that from a buffer control using qLCMS. Multiple compounds may be tested at once. Here, 2 are shown: a compound that binds mucin (blue stars) and a compound that does not (red circles). (b) Binding of colistin to mucin and methylcellulose (MC); negative ΔG implies binding. Colistin binds mucin but not MC. Error bars: SEM. *p=0.48 (t=0.85, df=2). (c) Surviving *P. aeruginosa* strain PAO1 cells (CFU = colony forming units) after 24 hours of growth with colistin in biopolymer-free Mueller-Hinton broth (MHB), in mucin (0.5% w/v dissolved in MHB), or MC (0.5% w/v dissolved in MHB). Colistin activity is inhibited by mucin, but not by MC. No colistin control shows that mucin does not increase growth in the absence of antibiotic. *p<0.0001

molecular properties that affect binding energies, so it is a better representation of binding data for the regression analyses that follow. A negative ΔG implies binding, with more negative ΔG meaning tighter binding and/or an increased binding site density.

We first tested the utility of this assay using colistin, a cationic polymyxin antibiotic that has been tested clinically as an inhaled treatment for *P. aeruginosa* infection in CF.[124] Our binding measurements revealed that colistin bound to mucin but not to MC (a neutral biopolymer) at 0.5% w/v (Fig. 2.1b; we retained this biopolymer concentration throughout this entire work, see Methods). This is consistent with expectations, given colistin's +5 charge.

Next, we showed that these binding measurements correlated with differential activity against *P. aeruginosa*. We incubated *P. aeruginosa* with colistin in biopolymer-free Mueller Hinton Broth (MHB), mucin, and MC for 24 h. As predicted, mucin inhibited the effect of colistin to a substantial degree, while MC had a negligible effect on colistin bioactivity (Fig. 2.1c). These data demonstrate that binding measurements can identify instances where mucus inhibits a drug by binding and sequestering it, which has important implications for mucosally-administered drugs.

Beyond the use of our novel method to predict the effects of mucus on bioactivity for a tested molecule, the ability to predict and guide alterations that would affect mucus binding of untested molecules would be an important advance. While we measured binding of 96 molecules with ED in this work – a substantial increase in throughput over previous techniques – this still constitutes a sparse sampling of chemical space.

2.4.2 Identification of a mucin binding motif by small molecule microarray.
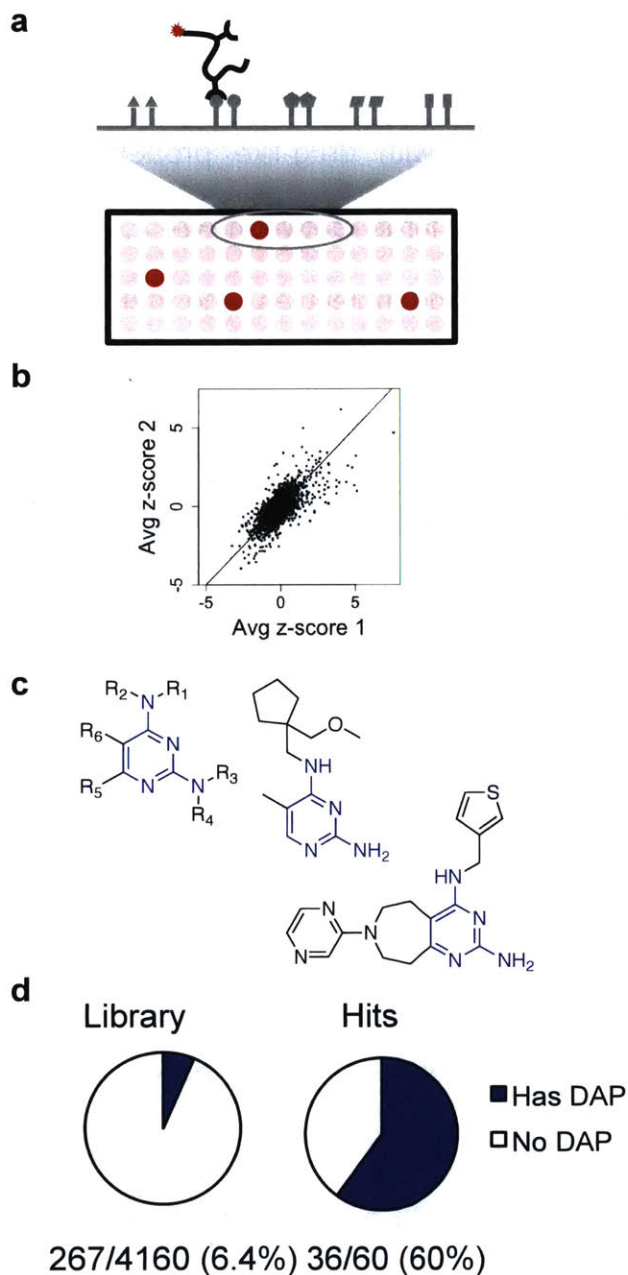
Figure 2.2 (a) Schematic of small molecule microarray (SMM): fluorescently labeled mucin was incubated with slides covalently linked to small molecules at defined locations. After washing and imaging, fluorescent spots on the slide indicated binding of molecule to mucin. (b) Correlation (ρ = 0.65) between average z-scores from two separate SMM experiments indicates high reproducibility. (c) 2,4-diaminopyrimidine (DAP) motif (blue) and two representative SMM hits with DAP. (d) Pie charts show overrepresentation of DAP in hits compared to the entire library.

Towards this end, we used a small molecule microarray to screen for mucin-binding small molecules (Fig. 2.2a, Methods). The glass slides were spotted with an array of covalently attached small molecules with drug-like properties: molecular weights typically between 250 and 400 Da, charge between +/- 2, and moderate hydrophobicity. We first ran an SMM assay on 4160 molecules, and then expanded to a set of 45,656 molecules with similar molecular properties to the initial screen. Both assays were performed in phosphate-buffered saline with Tween 20 added (PBST).

From the initial SMM, we identified 60 putative mucin binding molecules or "hits." Inspection of the structures revealed one motif that was strongly overrepresented in the hits over the non-hits (risk ratio 21.9, p~10^{-28}, Fisher's exact test without multiple hypothesis testing correction). We termed the motif 2,4-diaminopyrimidine (DAP; Fig. 2.2c-d; Table 2.S5 for descriptive statistics). The DAP motif was also highly overrepresented within the 140 hits in the second, expanded SMM screen (risk ratio 9.3, p~10^{-23}).

2.4.3 Equilibrium dialysis supports the presence of mucin binding motif.

We validated the importance of DAP using ED, a more robust method than the SMM. To do this, we selected 96 molecules: the 60 hits from the initial screen, along with 36 assay negatives that include a large number of molecules containing DAP. Of these 96 molecules, 54 were amenable to measurement with qLCMS (Table S6; the others did not ionize sufficiently or were chromatographically non-separable from inorganic salt). We used our ED method to measure binding of these molecules to mucin under two pH and ionic strength conditions: PBS (ionic strength ~160mM, pH 7.4) and a lower salt phosphate-citrate buffer (PCB; 60mM Na^+ was the sole cation) at pH 7.

Exact buffer compositions are given in the Experimental section. These two buffers capture some relevant physiological variation: salt levels in lung mucus are typically isotonic or slightly hypotonic with respect to plasma,[119,125] while pH of lung mucus is neutral to slightly basic for healthy people and slightly acidic in CF patients.[126] For binding to mucin in PCB, the average ΔG was lower for molecules with DAP than for molecules without (p=0.0043, $F_{1,50}$=8.9), meaning that the presence of DAP was positively correlated with increased mucin binding, though this relationship did not hold in PBS (Figs 2.S1, 2.S2). Thus, DAP is indeed associated with mucin binding and this effect is dependent on ionic composition. The higher ionic strength of PBS is the most likely explanation for the difference in binding, although the slight pH difference may have an effect; the relatively low concentration of potassium in our PBS (1mM) is unlikely to have a significant impact.

2.4.4 Molecular properties beyond DAP affect mucin, DNA, and alginate binding.

We next addressed two questions: first, to what extent is binding to mucin distinguished from binding to other mucus-associated biopolymers, and second, are properties beyond the presence of DAP (specifically, hydrophobicity and charge) important determinants of binding to mucus. To answer the first question, we measured binding of the same set of molecules to DNA and alginate, two polymers that are major components of CF lung mucus, and bovine serum albumin (BSA) as a non-polyanionic control. To address the second question, we determined each molecule's predicted log octanol-water partition coefficient (ClogP; the higher the ClogP, the higher the predicted hydrophobicity) and average net charge at pH 7 (for PCB) and 7.4 (for PBS).

41
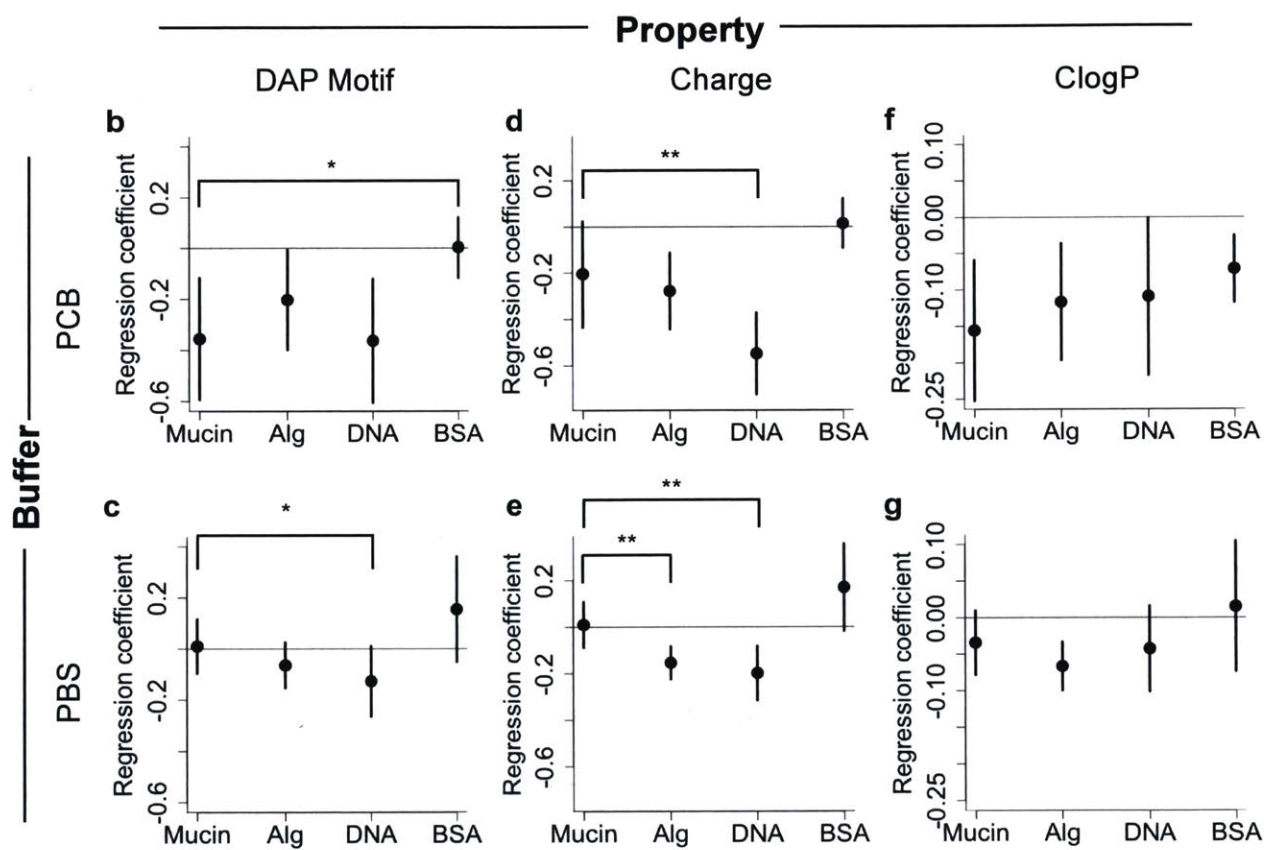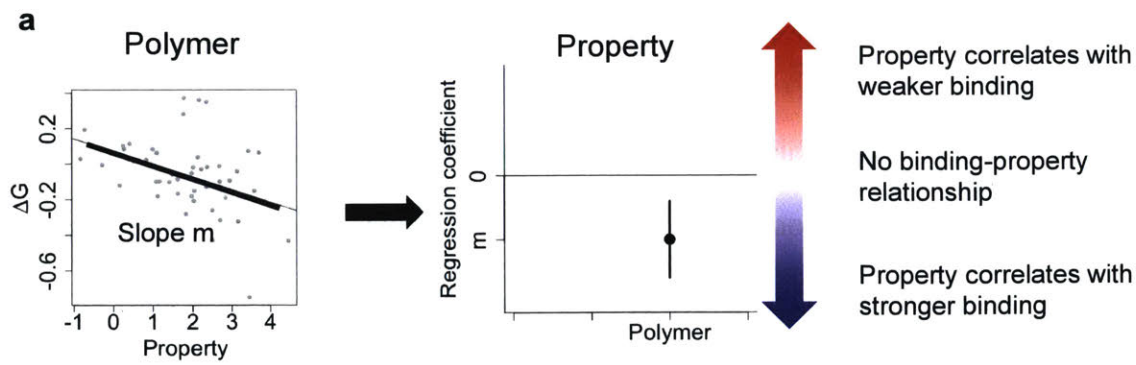
Figure 2.3 Effects of molecular properties on binding of small molecules to mucin, alginate ("Alg"), DNA, and BSA in PBS and PCB. (a) Schematic of analysis technique. Regression coefficient for linear model of ΔG of binding to a biopolymer against a particular molecular property (charge, ClogP, or the presence of DAP) is plotted with 95% confidence intervals. Plot on left: each individual point represents a molecule whose binding was measured along with a molecular property; charge and ClogP can be real numbers as shown in the plot, while the presence of DAP is an indicator variable, either 0 or 1. Negative and positive slopes show that increases in the predictor are associated with stronger or weaker binding respectively. If a confidence interval does not intersect with 0, association between ΔG and the property of interest is statistically significant for the relevant polymer ($p<0.05$). (b,c) Regression coefficient comparing molecules with and without DAP in (b) PCB and (c) PBS. For example, a value of -0.4 means ΔG is 0.4 lower for tested molecules with DAP compared to those without. Pairwise comparisons test whether the effect of DAP varies with different polymers (see Methods for statistical analysis). (d,e) Regression coefficient for ΔG vs. predicted molecular charge in (d) PCB and (e) PBS. Negative values imply that increased charge leads to greater binding (more negative ΔG). (f,g) Regression coefficient for ΔG vs. ClogP in (f) PCB and (g) PBS. Negative values mean that increased ClogP leads to greater binding. Significance for pairwise comparisons only shown for mucin vs. the other polymers. *$p<0.05$, **$p<0.01$ for pairwise comparisons. Exact p values given in Table S7.

The results of these experiments and analyses are presented in Fig. 2.3, which depicts the slopes of best-fit lines for binding to each biopolymer vs. the presence of DAP, net charge, and ClogP (Fig. 2.3; full plots for each individual biopolymer-property pair given in Figs 2.S1-2.S2). In PCB, the presence of DAP was significantly correlated not only with binding to mucin but also with binding to DNA and alginate (Fig. 2.3b), though not to BSA. Thus, DAP is not specific to mucin binding.

Charge and hydrophobicity also affect binding to the biopolymers. In both buffers, increased net charge is significantly correlated with binding to DNA and alginate, though the effect is stronger in PCB, consistent with the fact that increased ionic strength weakens electrostatic interactions. Moreover, there is a trend in the same direction for mucin in PCB ($p=0.080$, $F_{1,50}=3.2$), although the effect of charge is weaker for mucin than for DNA and alginate (Fig. 2.3d,e). These data suggest that (DNA-rich) CF lung mucus, and/or mucus with (alginate-rich) mucoid *P. aeruginosa* biofilms, will likely have

different small molecule binding properties than the mucin-dominated mucus typically observed in COPD and asthma.[106] Increased ClogP was also correlated with stronger binding to each of the biopolymers in PCB (Fig. 2.3f), which is consistent with previous observations that increased hydrophobicity leads to greater mucus binding.[20]

Some caution is required in interpreting these single-variable plots, as there is substantial collinearity between the presence of DAP and the other descriptors (Fig. 2.S3a-c). However, these results all hold true when restricting our dataset to only molecules with DAP (Fig. 2.S4). This means that mucus binding of the tested molecules containing DAP could be modulated by the presence of charged and hydrophobic groups elsewhere on the molecular structure.

2.4.5 ED experiments across a more chemically diverse set reveal multiple previously unknown mucin binding molecules.

Finally, we tested whether the results related to charge and hydrophobicity would apply to molecules that were more structurally diverse, since our ED experiments focused mainly on molecules with DAP. We selected 41 molecules that were amenable to qLCMS analysis from the NIH Clinical Collection (NCC), a library of clinically tested drugs and measured their binding to mucin, DNA, alginate, and BSA in PBS. Interestingly, none of the trends we observed in Fig. 2.3 held for this more diverse set of molecules (Fig. 2.4a,b), suggesting that there are not monotonic relationships between binding, hydrophobicity, and charge across all chemical structures. However, we did identify several molecules able to bind to mucin (Fig. 2.4c), which may reflect mucin binding scaffolds beyond DAP that were unidentified or untested in the SMM. The lack of clear trends observed in the NCC dataset suggests that correlations between

44

a

Charge

b

ClogP

c

i ii

iii iv

d

ΔG

Property

Chemical space

e f

ΔG

Property

ΔG

Property

Figure 2.4 Effect of molecular properties on binding of molecules with diverse structures to biopolymers in PBS. (a) Regression coefficient for ΔG vs. predicted molecular charge. Negative values mean increased charge leads to greater binding (more negative ΔG). (b) Regression coefficient for ΔG vs. ClogP. Negative values indicate increased ClogP leads to greater binding. If a confidence interval does not intersect with 0, association between ΔG and the property of interest is statistically significant for the relevant polymer ($p<0.05$). Significance for pairwise comparisons only tested for mucin vs. the other biopolymers (*$p<0.05$) c) Structures of the four strongest mucin binders ((i) cladribine, (ii) doxorubicin, (iii) diphencyprone, (iv) maprotiline) reveal no similarities: (d-f) Schematic depiction of why correlations between molecular properties (such as charge or logP) and biopolymer binding strength (ΔG) would vary depending on the molecules tested. (d) Depiction of many molecular structures projected onto one axis, which we call "chemical space," plotted against ΔG and a molecular property. Each colored vertical line represents one point in chemical space (i.e. one molecule). Blue lines depict a set of molecules with low chemical diversity (such as molecules with DAP or some other common scaffold), and red lines depict a set of molecules with relatively high chemical diversity (such as the NCC set). (e) Plot of ΔG versus the molecular property based on the blue locations shown in (a), along with best fit line. Increased values of the molecular property are locally associated with stronger binding (lower ΔG), akin to the correlation of ClogP with ΔG for molecules with DAP. (f) Plot of ΔG versus the molecular property based on the red locations shown in (d), along with best fit line. The molecular property is not globally associated with binding (ΔG), akin to the lack of correlation of charge with ΔG for molecules from the NCC set. Exact p values for (a) and (b) given in Table S7.

charge or hydrophobicity and binding to mucus may hold locally, but not globally, in chemical space (Fig. 2.4d-f). This makes sense if one considers that the molecules in the SMM set (which had relatively low chemical diversity) likely bind the same or similar regions on a biopolymer (such as an anionic glycosylated region of mucin) and/or with the same orientation (such as intercalated into DNA). In this case, adding or subtracting hydrophobic groups or charge would be expected to have a systematic effect on binding dictated by the common biopolymer binding site. In contrast, for molecules with vastly different scaffolds such as the ones shown in Fig. 2.4c, binding sites and/or conformations would be more varied, reducing the expectation of correlation between binding and coarse summaries of molecular properties such as charge and ClogP.

2.5 CONCLUSIONS

Binding of small molecules to mucus is an important contributor to their activity and diffusivity within mucus. Here, we described a new technique to measure dozens of binding interactions between drugs and biopolymers commonly found in healthy and diseased mucus, and identified a mucin binding motif, DAP (Table 2.1a).

We also showed that for the molecules containing DAP, increased hydrophobicity and charge were associated with increased binding to various constituent components of mucus (Table 2.1b). Interestingly, for these molecules charge was a better predictor of binding to DNA and alginate than to mucin. This shows that these biopolymers have distinct responses to charge in spite of the fact that they are all polyanions, which is a point of potential relevance for CF.

The dependence of mucus binding on molecular detail was highlighted by the diverse molecules that bound mucus components in the NCC set, where no correlation to simple descriptors of charge and hydrophobicity was evident (Table 2.1b). Fortunately for the future of mucus binding research in drug design, drug development projects typically progress by testing molecules with a common scaffold, meaning that structure-binding analyses take place within a restricted chemical space (similar to the SMM set) so are likely to be enlightening. Furthermore, even in the absence of structure-binding relationships or other molecular understanding, we have established a novel and valuable method to test binding of many molecules to mucus with minimal sample requirements.

We have focused here on lung mucus and MUC5AC in particular. A similar mucin, MUC5B, has the same domain structure as MUC5AC and is also present in lung

Table 2.1 Summarized results showing associations between molecular properties and binding to mucin, alginate, DNA, and BSA. a) The presence of DAP is associated with binding to mucin, alginate, and DNA but not BSA. b) Associations of charge and hydrophobicity with binding to each biopolymer for molecules with DAP from SMM set and for molecules from the more diverse NIH Clinical Collection set.

A

Biopolymer	Correlation w/DAP		
Mucin	+	+	Positive correlation
Alginate	+	0	No correlation
DNA	+	-	Negative correlation
BSA	0		

B

	Biopolymer	Correlation w/charge	Correlation w/hydrophobicity
Has DAP	Mucin	0	+
	Alginate	+	+
	DNA	+	+
	BSA	0	+
NCC set	Mucin	0	0
	Alginate	0	0
	DNA	0	-
	BSA	-	-

mucus, and likely abundant in the lungs of CF and COPD patients.[105] The shared domain structure of the two proteins suggests they may have similar binding tendencies, but important differences may also exist. It would therefore be an interesting extension of this work to test purified MUC5B or to more thoroughly examine mucus from mucus-secreting cell lines. Furthermore, the methods we have employed are easily transferrable to other mucus layers (such as the MUC2-rich intestine),[7,104,127] mucin-based drug delivery systems,[107] and studies of catalysis by mucin,[108] and

possibly to mucus-free biopolymer systems such as biofilms.[22,23] Our results with alginate may be interesting in this regard, and the application of our techniques to study small molecule binding to other biofilm-associated polysaccharides such as the *P. aeruginosa*-synthesized Pel and Psl[82] is an interesting future direction. Additionally, mucus binding to other molecules such as bacterial toxins and quorum sensing molecules may be important for understanding benign and infectious microbial interactions in mucus.[25,128] Overall we believe that the results presented here, particularly our development of an ED method to measure binding in a high throughput manner, show that measuring and predicting mucus binding are approachable problems and promising avenues for further investigation into small molecule behavior in mucus.

2.6 SUPPLEMENTAL INFORMATION

2.6.1 Derivation of Equation (3)

First, let us define our terms. We consider an ED system with some number of molecules of interest, which can be in one of three states: in the assay chamber (state A), in the sample chamber (state S), and nonspecifically bound to the ED system walls and thus inaccessible to measurement (state B). Furthermore, we have two ED systems (see Fig. 2.1): a buffer-buffer control (BB) and a polymer-buffer experiment (PB). We therefore name the 6 variables of interest:

x: fraction of molecules in the assay chamber in the buffer-buffer control

y: fraction of molecules non-specifically bound to the ED system walls in the buffer-buffer control

z: fraction of molecules in the sample chamber in the buffer-buffer control

a: fraction of molecules in the assay chamber in the polymer-buffer experiment

b: fraction of molecules non-specifically bound to the ED system walls in the polymer-buffer experiment

c: fraction of molecules in the sample chamber in the polymer-buffer experiment

Thus, what we seek, the uptake ratio, is given by:

$R_U = 3c/a$ (S1)

The factor of 3 corrects for the fact that the assay chamber contains 3 times the volume of the sample chamber, and we are interested in the ratio of concentrations, not total number of molecules.

We also assume that we have reached equilibrium, and that the biopolymer does not affect equilibration between states A and B, since the assay chamber always only contains buffer and the ED chamber has a constant architecture. Therefore, we have the relation:

$b/a = y/x$ (S2)

In the buffer-buffer control, clearly the concentrations of molecule in the sample and assay chambers will be equal, therefore:

$x = 3\,z$ (S3)

because the assay chamber has 3 times the volume of the sample chamber. Finally, we have the normalization conditions:

$x + y + z = 1$ (S4)

$a + b + c = 1$ (S5)

Returning to Equation S1, we use these equations to get:

$R_U = 3c/a$

$ = 3(1-a-b)/a$

$ = 3/a - 3 - 3b/a$

$ = 3/a - 3 - 3y/x$

$ = 3/a - 3 - 3(1 - x - z)/x$

$ = 3/a - 3 - 3/x + 3 + 3z/x$

$ = 3/a - 3/x + 1$

$ = 1 + 3(1/a - 1/x)$ (S6)

Now let us consider the relationship between a and c_E and between x and c_C. Since the molecules were all added at 1x working concentration to both the assay and sample

chambers, that means that 1x corresponds to having ¾ of the molecules in the assay chamber. Therefore, we have:

$$c_E = 4/3\ a \qquad (S7)$$

$$c_C = 4/3\ x \qquad (S8)$$

Substituting these relations into equation S6 gives us:

$$R_U = 1 + 3(1/(¾\ c_E) - 1/(¾\ c_C))$$

$$R_U = 1 + 4(1/c_E - 1/c_C) \qquad (S9)$$

As desired.

2.6.2 Supplemental Figures
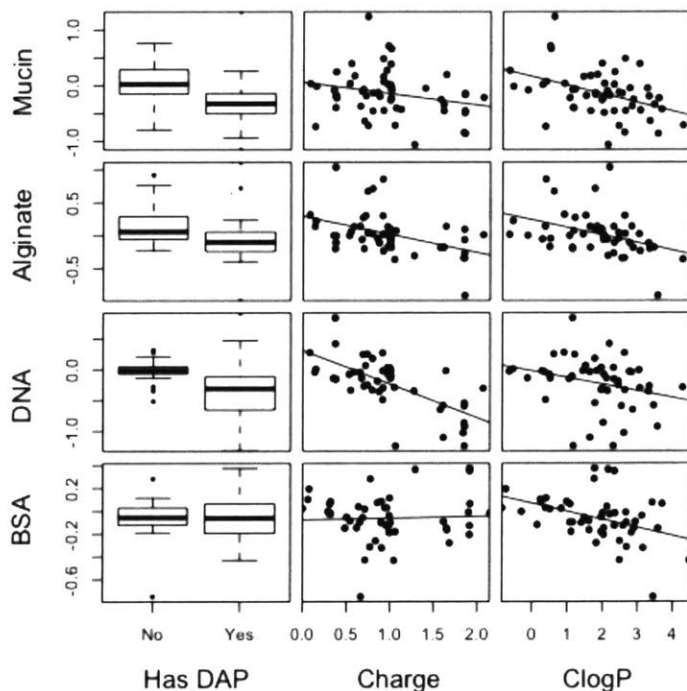
Figure 2.S1 Correlations of presence of DAP, charge, and ClogP with mucin, alginate, DNA, and BSA binding (as measured by ΔG) in phosphate-citrate buffer (PCB). Box plots: center line, median; box limits, upper and lower quartiles; whiskers, 1.5x interquartile range; points, outliers.
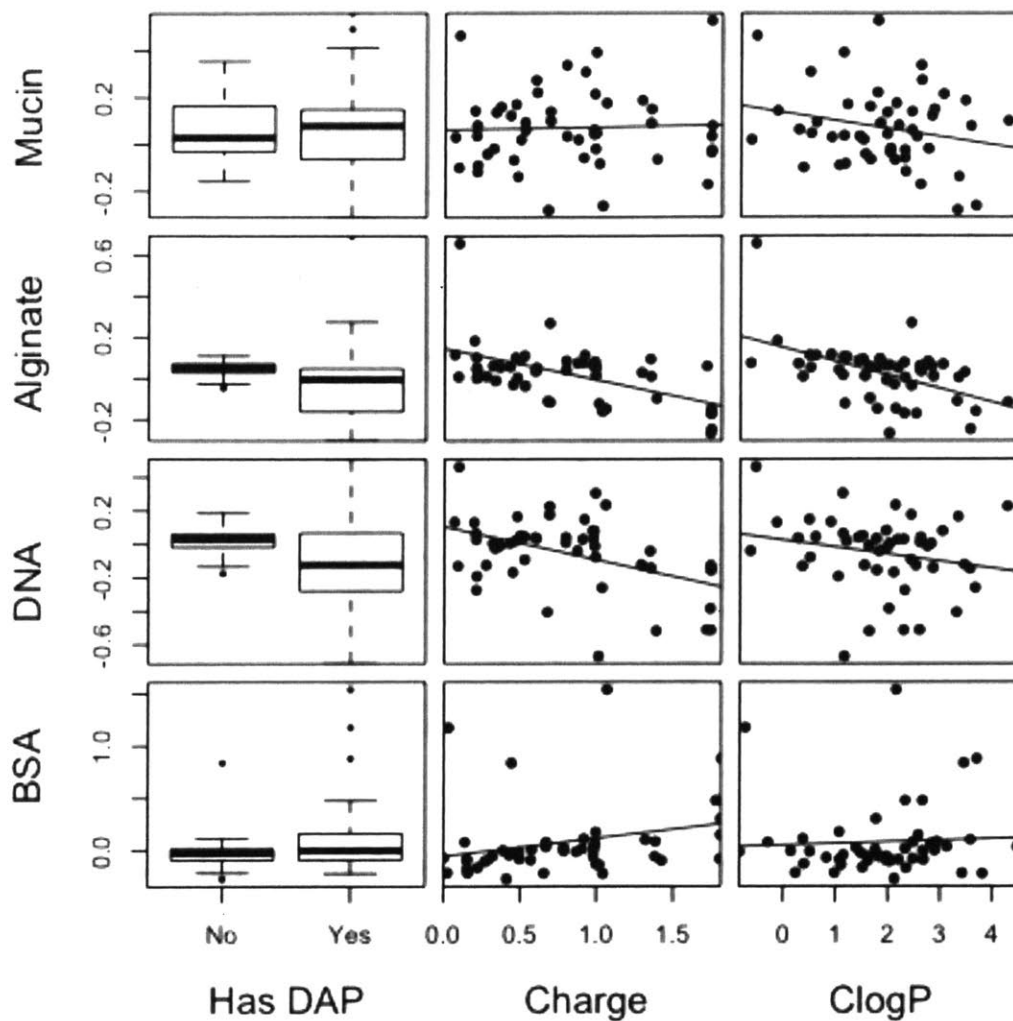
Figure 2.S2 Correlations of presence of DAP, charge, and ClogP with mucin, alginate, DNA, and BSA binding (as measured by ΔG) in PBS. Box plots: center line, median; box limits, upper and lower quartiles; whiskers, 1.5x interquartile range; points, outliers.
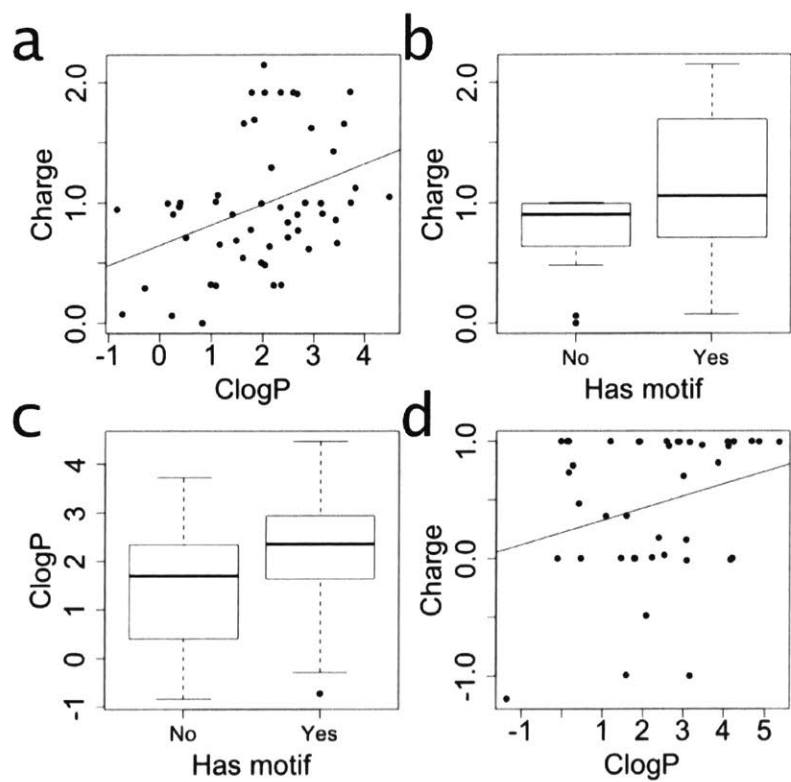
Figure 2.S3. Collinearity of molecular properties. a-c) Correlation of charge, ClogP, and presence of DAP motif for SMM molecule set (charge calculated at pH 7). For c: center line, median; box limits, upper and lower quartiles; whiskers, 1.5x interquartile range; points, outliers. d) Correlation of charge and ClogP for NIH Clinical Collection molecule set.
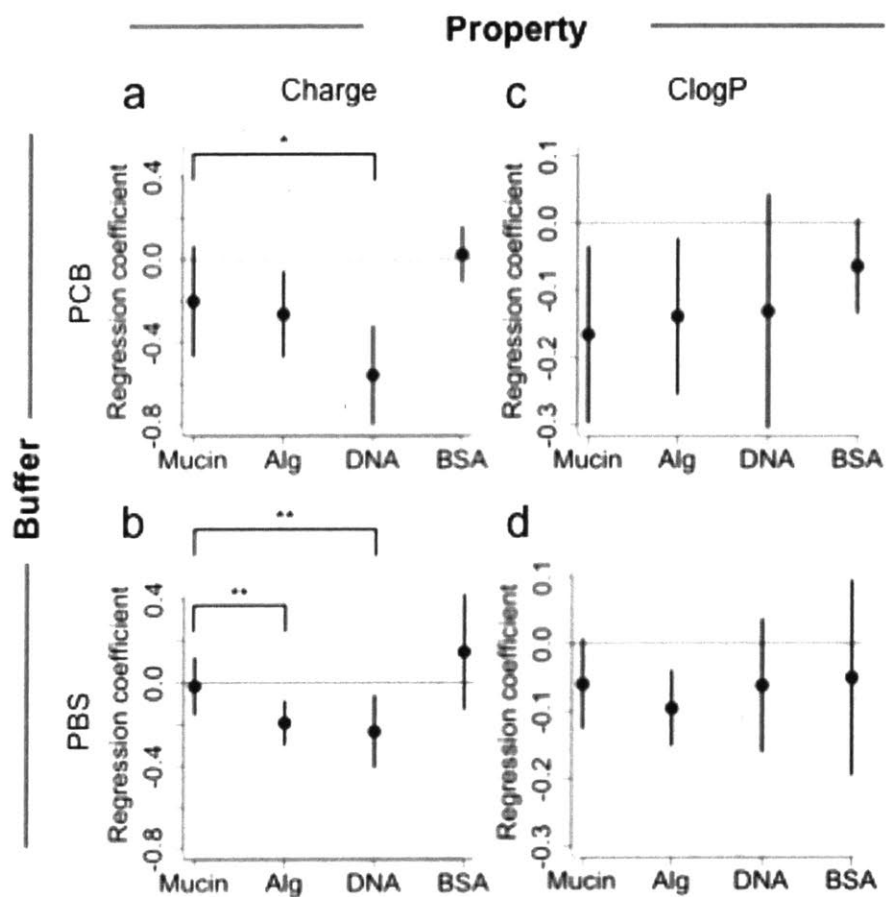
Figure 2.S4. Summarized effect of molecular properties on binding to mucin, alginate, DNA, and BSA in PCB and PBS buffers. As described in Fig. 3, but data is restricted to only molecules with DAP, not all the molecules in the SMM dataset as in Fig. 3. *p<0.05, **p<0.01 for pairwise comparisons

2.6.3 Supplemental Tables

Table 2.S1. HPLC method. Solvent A: ultrapure water with 0.1% v/v glacial acetic acid (VWR), solvent B: HPLC grade acetonitrile (VWR) with 0.1% v/v glacial acetic acid (VWR).

Time (min)	% A	% B	Flow rate (mL/min)
0	100	0	0.3
5	100	0	0.3
20	0	100	0.3
25	0	100	0.3
25.1	100	0	0.3
35.1	100	0	0.3

Table 2.S2. MS settings for SMM molecules

Chembridge ID	Cocktail Group 1	Cocktail Group 2	Precursor Ion
99405121	3	4	300.1
92454529	3	3	381.2
89107391	2	3	299.2
88891724	1	1	354.1
88691716	2	4	355.2
88381944	1	3	335.2
87261711	1	2	296.1
86162542	1	1	360.2
84396124	1	3	365.2
80685707	1	1	348.2

79987012	2	2	329.2
78276598	1	4	381.3
78251495	2	3	340.2
74871325	3	1	310.2
74594727	1	2	339.2
72500924	1	4	354.2
72210516	2	1	370.2
64206473	1	3	306.2
60611543	2	2	359.3
58361985	1	3	290.2
57570910	3	3	488.3
57393919	2	4	290.2
55057377	1	1	251.2
54551735	2	4	351.2
52858814	3	3	345.2
52622972	3	3	355.2
51816129	3	1	342.2

48381932	1	2	379.2
47705523	2	1	359.2
46065915	1	1	357.2
45936238	3	4	292.2
43292549	3	3	324.2
42567250	2	1	382.2
41410851	1	4	342.2
40339734	3	4	348.2
40281582	2	1	378.2
37622775	2	3	287.2
34447012	1	4	325.2
31774206	2	2	349.2
31149690	3	4	353.2
29487940	3	2	354.2
28796746	3	4	359.3
28016876	2	2	335.2
27050181	3	1	289.2

26073561	1	4	316.2
26044756	3	3	339.2
25759080	2	2	344.2
23940862	2	1	309.2
22248652	3	3	361.2
20656559	2	3	354.2
20507984	2	2	322.1
19469566	3	1	334.2
11499081	3	4	378.2
11409740	1	2	348.3
11409740	1	2	348.3

Table 2.S3. MS settings for NCC molecules

Pubchem SID	Cocktail group	Precursor Ion	Product Ion
46386546	1	611.2	303
46386749	1	407.2	222
46386806	1	376.1	165
46386560	1	359.2	165
46386981	1	347.1	112
104170181	1	340.2	239.2
104170085	1	328.2	310.1
104170192	1	318.2	116.1
46386657	1	306.1	236.1
46386844	1	299.1	135
104169981	1	291.1	230.2
104169945	1	280.2	107.1
46386839	1	278.2	250.1
46386575	1	275.1	229.9
104170154	1	266.2	195
46386874	1	256.2	167
46386680	1	256	210.9
46386905	1	247.2	230.1
104169960	1	227.1	106
46386897	1	207.1	178
46386781	2	544.2	361
104170125	2	417.2	234.2
46386593	2	389.2	201
104170027	2	368.2	181
104170038	2	357.1	233
104170016	2	340.2	130.1
46386658	2	337.2	238
104170153	2	330.1	192.1
104170201	2	322.1	212
104170090	2	317.2	261
104170224	2	306.1	158.9
104170137	2	296.1	264.1
46386544	2	286.1	170
104169995	2	278.2	191
46386708	2	275.1	123
104170185	2	260.2	116.2
46386698	2	256	171.9
104170026	2	249.1	232.9
46386846	2	230.1	170.9
46386925	2	189.1	161
46386687	2	177.1	106

Table 2.S4. MS settings for colistin

Molecule	Precursor ion	Product ion
Colistin	385.9	101.1
Colistin	385.9	380.1

Table 2.S5. Descriptive statistics for overrepresentation of DAP among SMM hits (DAP overrepresented with in hits, p~10^{-28})

	SMM Hit	Negative	Total	Hit %
DAP	36	231	267	13.48
No DAP	24	3869	3893	0.62
Total	60	4100	4160	1.44

Table 2.S6. Statistics for SMM molecules measured using ED

	SMM Hit	Negative	Total
DAP	22	8	30
No DAP	15	9	24
Total	37	17	54

Table 2.S7. Exact statistics for tests of association between predictors (the presence of DAP, charge, and ClogP) and binding to a biopolymer (mucin, alginate, DNA, or BSA). Type "Association" means a simple test for association. Type "vs. Mucin" means a comparison of the relative predictive capacity of the predictor for the biopolymer of interest versus mucin (the comparisons shown in Fig. 4).
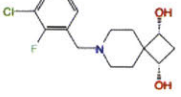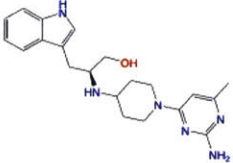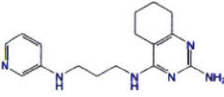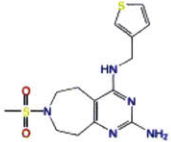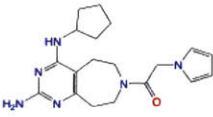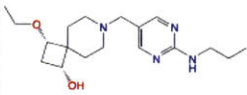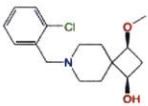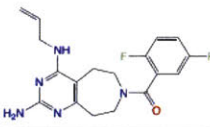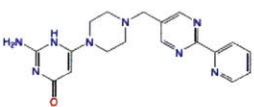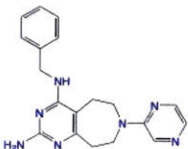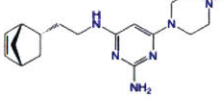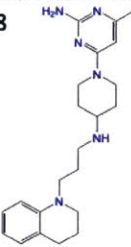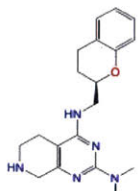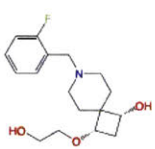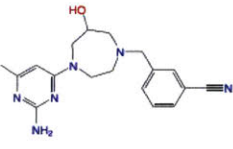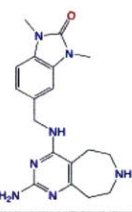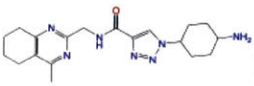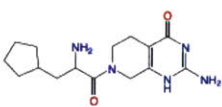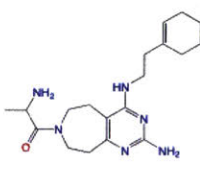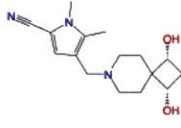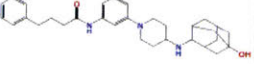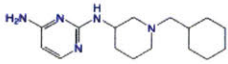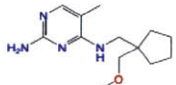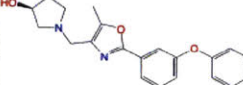
Dataset	Type	Buffer	Predictor	Biopolymer	p value	Statistic
SMM	Association	PCB	DAP	Mucin	0.004	F(1,50)=8.9
SMM	Association	PCB	DAP	Alginate	0.043	F(1,47)=4.3
SMM	Association	PCB	DAP	DNA	0.004	F(1,47)=9.0
SMM	Association	PCB	DAP	BSA	0.96	F(1,47)=0.002
SMM	Association	PCB	Charge	Mucin	0.08	F(1,50)=3.2
SMM	Association	PCB	Charge	Alginate	0.0014	F(1,47)=11.5
SMM	Association	PCB	Charge	DNA	0.0000001	F(1,47)=39.2
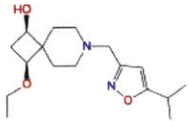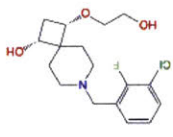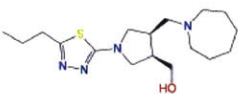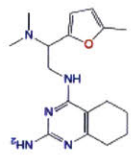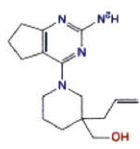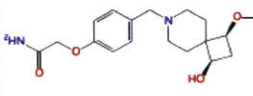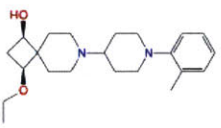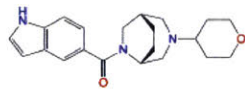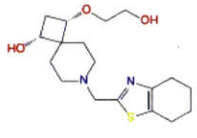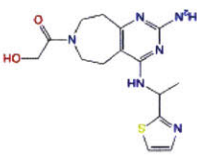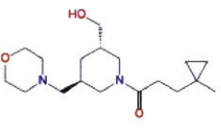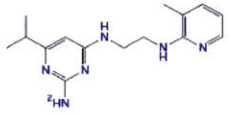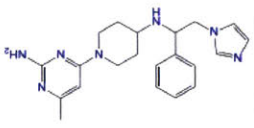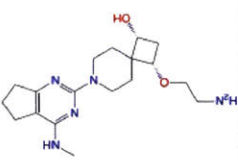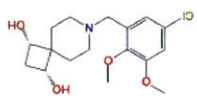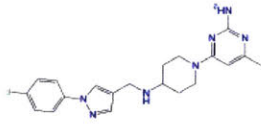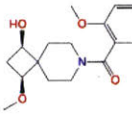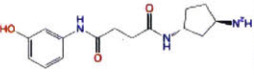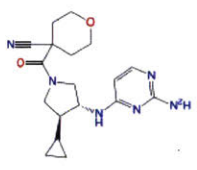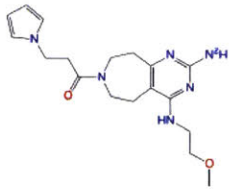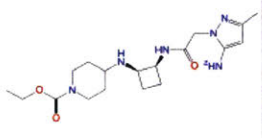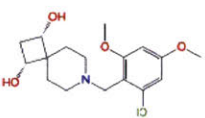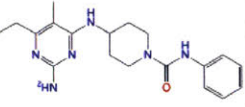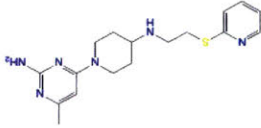
SMM	Association	PCB	Charge	BSA	0.78	F(1,47)=0.074
SMM	Association	PCB	ClogP	Mucin	0.0021	F(1,50)=10.5
SMM	Association	PCB	ClogP	Alginate	0.0052	F(1,47)=8.6
SMM	Association	PCB	ClogP	DNA	0.048	F(1,47)=4.1
SMM	Association	PCB	ClogP	BSA	0.00342	F(1,47)=9.5
SMM	vs. Mucin	PCB	DAP	Alginate	0.19	F(1,47)=1.8
SMM	vs. Mucin	PCB	DAP	DNA	0.94	F(1,47)=0.005
SMM	vs. Mucin	PCB	DAP	BSA	0.0056	F(1,47)=8.4
SMM	vs. Mucin	PCB	Charge	Alginate	0.54	F(1,47)=0.38
SMM	vs. Mucin	PCB	Charge	DNA	0.0058	F(1,47)=8.4
SMM	vs. Mucin	PCB	Charge	BSA	0.075	F(1,47)=3.3
SMM	vs. Mucin	PCB	ClogP	Alginate	0.35	F(1,47)=0.88
SMM	vs. Mucin	PCB	ClogP	DNA	0.33	F(1,47)=0.96
SMM	vs. Mucin	PCB	ClogP	BSA	0.1	F(1,47)=2.8
SMM	Association	PBS	DAP	Mucin	0.84	F(1,46)=0.04
SMM	Association	PBS	DAP	Alginate	0.15	F(1,46)=2.1
SMM	Association	PBS	DAP	DNA	0.07	F(1,46)=3.4
SMM	Association	PBS	DAP	BSA	0.13	F(1,46)=0.13
SMM	Association	PBS	Charge	Mucin	0.83	F(1,46)=0.04
SMM	Association	PBS	Charge	Alginate	0.00007	F(1,46)=19.1
SMM	Association	PBS	Charge	DNA	0.001	F(1,46)=11.6
SMM	Association	PBS	Charge	BSA	0.08	F(1,46)=3.3
SMM	Association	PBS	ClogP	Mucin	0.12	F(1,46)=2.5
SMM	Association	PBS	ClogP	Alginate	0.00022	F(1,46)=16.1
SMM	Association	PBS	ClogP	DNA	0.15	F(1,46)=2.1

SMM	Association	PBS	ClogP	BSA		0.73	$F_{(1,46)}=0.12$
SMM	vs. Mucin	PBS	DAP	Alginate		0.18	$F_{(1,46)}=1.8$
SMM	vs. Mucin	PBS	DAP	DNA		0.017	$F_{(1,46)}=6.1$
SMM	vs. Mucin	PBS	DAP	BSA		0.15	$F_{(1,46)}=2.1$
SMM	vs. Mucin	PBS	Charge	Alginate		0.00076	$F_{(1,46)}=13.9$
SMM	vs. Mucin	PBS	Charge	DNA		0.000026	$F_{(1,46)}=21.9$
SMM	vs. Mucin	PBS	Charge	BSA		0.086	$F_{(1,46)}=3.1$
SMM	vs. Mucin	PBS	ClogP	Alginate		0.17	$F_{(1,46)}=1.9$
SMM	vs. Mucin	PBS	ClogP	DNA		0.73	$F_{(1,46)}=0.11$
SMM	vs. Mucin	PBS	ClogP	BSA		0.24	$F_{(1,46)}=1.4$
NCC	Association	PBS	Charge	Mucin		0.92	$F_{(1,39)}=0.0098$
NCC	Association	PBS	Charge	Alginate		0.11	$F_{(1,39)}=2.7$
NCC	Association	PBS	Charge	DNA		0.77	$F_{(1,39)}=0.089$
NCC	Association	PBS	Charge	BSA		0.007	$F_{(1,39)}=7.9$
NCC	Association	PBS	ClogP	Mucin		0.057	$F_{(1,39)}=3.9$
NCC	Association	PBS	ClogP	Alginate		0.086	$F_{(1,39)}=3.1$
NCC	Association	PBS	ClogP	DNA		0.01	$F_{(1,39)}=7.3$
NCC	Association	PBS	ClogP	BSA		0.0044	$F_{(1,39)}=9.1$
NCC	vs. Mucin	PBS	Charge	Alginate		0.31	$F_{(1,39)}=1.05$
NCC	vs. Mucin	PBS	Charge	DNA		0.72	$F_{(1,39)}=0.13$
NCC	vs. Mucin	PBS	Charge	BSA		0.022	$F_{(1,39)}=5.7$
NCC	vs. Mucin	PBS	ClogP	Alginate		0.3	$F_{(1,39)}=1.1$
NCC	vs. Mucin	PBS	ClogP	DNA		0.79	$F_{(1,39)}=0.079$
NCC	vs. Mucin	PBS	ClogP	BSA		0.21	$F_{(1,39)}=1.6$

2.6.4 Supplemental Charts

Chart 2.S2.1 Structures from SMM tested using ED, listed by Chembridge ID

99405121	92454529	89107391	88891724
88691716	88381834	87261711	86162542
84396124	80685707	79987012	78276598
78251495	74871325	74594727	72500924
72210516	64206473	60611543	58361985
57570910	57393919	55057377	54551735
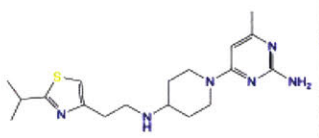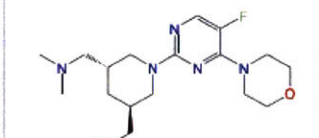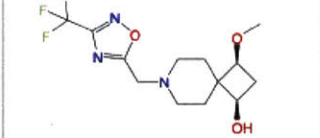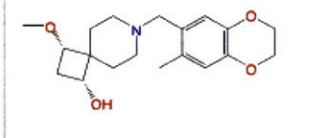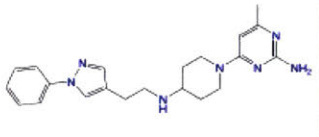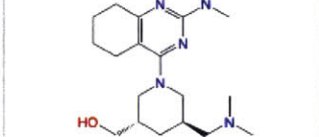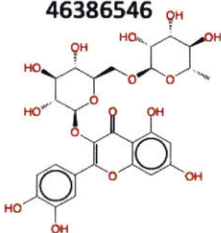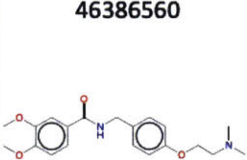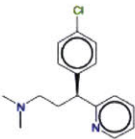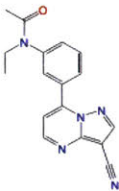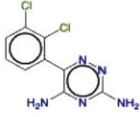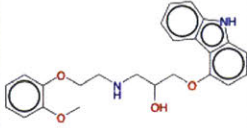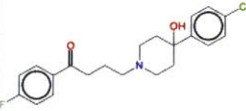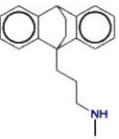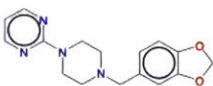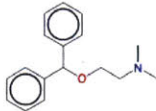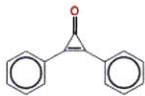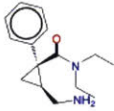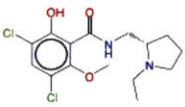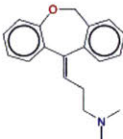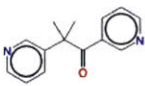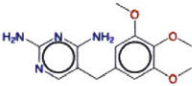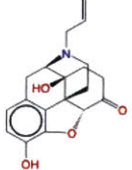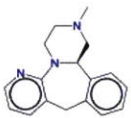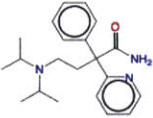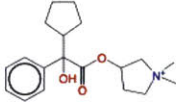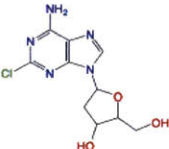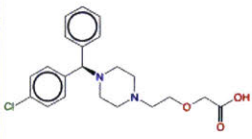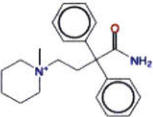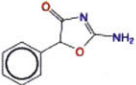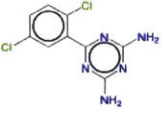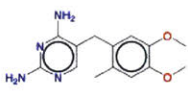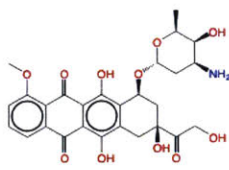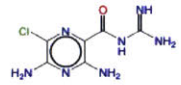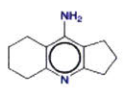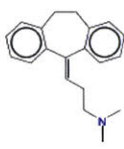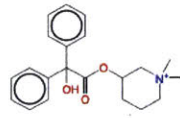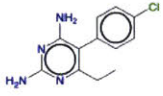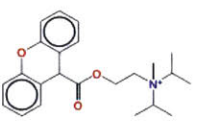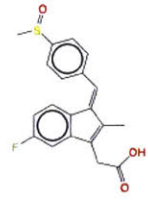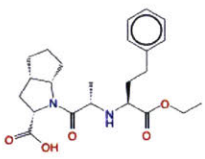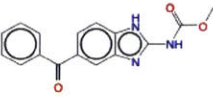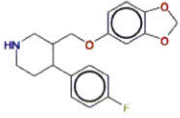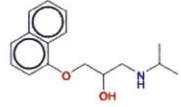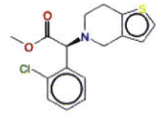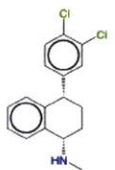
22248652

20656559

20507984

19469566

11499081

11409740

Chart 2.S2 Structures from NCC tested using ED, listed by PubChem SID

46386546	46386560	46386575	46386657
46386680	46386749	46386806	46386839
46386844	46386874	46386897	46386905
46386981	104169945	104169960	104169981
104170085	104170154	104170181	104170192
46386544	46386593	46386658	46386687

46386698	46386708	46386781	46386846

46386925	104169995	104170016	104170026

104170027	104170038	104170090	104170125

104170137	104170153	104170185	104170201

104170224

3 BINDING TO RESPIRATORY MUCUS FROM HUMAN BRONCHIAL EPITHELIAL CELL CULTURE

3.1 INTRODUCTION

In Chapter 2, we discussed the binding of small molecules to MUC5AC and two mucus-associated polymers: DNA, found in the respiratory mucus of people with CF, and alginate, found in mucoid *P. aeruginosa* infections of people with CF. However, we did not examine binding to other mucins or mucus layers in great detail, and we did not focus specifically on molecules whose activity is known to be relevant in a mucosal context (other than colistin).

There is some evidence to suggest that different mucins have at least broadly similar binding properties. In Appendix A, I used my equilibrium dialysis technique to measure binding of polymyxin and fluoroquinolone antibiotics to purified MUC5AC, MUC5B, and MUC2. Colistin (also known as polymyxin E) and polymyxin B bound each of these mucins, consistent with the idea that the major secreted mucins are polyanionic due to their glycosylation. The two fluoroquinolone antibiotics I tested, levofloxacin and ciprofloxacin, are predominantly zwitterionic at neutral pH with a single positive and negative charge each[129] and did not bind any of these mucins. This is roughly expected given a model where most molecules without strong hydrophobicity or positive charge do not strongly bind mucin. On a larger scale, PEGylated particles have shown effectiveness for penetration through cervical,[130] intestinal,[18] and respiratory[131] mucus. Since penetration is only possible in the absence of mucus binding,[1,94] this supports the common theme that a combination of hydrophilicity and charge neutrality leads to lack of interaction with mucus, regardless of which mucin is predominant.

However, it is likely that more detailed binding properties can be mucin- or mucus layer-specific. Thus, purified MUC5AC is unlikely to be a perfect proxy for binding to all

70

mucosa. Therefore, in this chapter we address the specific question of drug binding to respiratory mucus using mucus secreted from cultured primary human bronchial epithelial cells (HBE mucus). This mucus consists mostly of MUC5B, though some MUC5AC is also present. The MUC5B in HBE mucus does not share the exact glycosylation pattern of MUC5B identified from patients, which has two "glycoforms" containing different negative charge densities, while HBE mucus has just one glycoform.[132] Despite this slight difference, HBE mucus is still a commonly used model for respiratory mucus and respiratory MUC5B in particular.[133,134]

We set out to test three groups of molecules, each of particular interest for their interactions with respiratory mucus: pyocyanin, a toxin secreted by *P. aeruginosa* with clinical importance for CF; antibiotics, which are important for treating lung infections; and non-antibiotic inhaled drugs for lung diseases. Antibiotics are important because of the high prevalence of *P. aeruginosa* and *Staphylococcus aureus* lung infections in CF, as well as lung infections by these and other pathogens in ventilator-associated pneumonia (VAP), chronic obstructive pulmonary disease (COPD), and non-CF bronchiectasis. As a result, a number of antibiotics have been used or developed as inhaled therapies, because inhalation allows for a high concentration of antibiotics within lung mucus while limiting side effects on other organs.[135] We consider numerous classes of antibiotics that are in various stages of clinical testing and approval for treatment of respiratory infections. Furthermore, we measured binding of pyocyanin and these antibiotics to eight different biopolymers including three types of mucin: HBE mucus, pig gastric MUC5AC, and bovine submaxillary mucin (BSM). This allowed us to

compare results from the different mucin types to see how consistent the binding results are. Pyocyanin and the antibiotics are discussed in Section 3.3.1.

One issue of note is that we found the HBE mucus to be not entirely sterile, containing a very small density of viable bacteria. While this made experiments of antibiotic activity as performed in Fig. 2.1c infeasible since there are unknown bacteria present, we do not believe that these bacteria are at a high enough concentration to substantially affect binding experiments as these experiments take place over a shorter time period (4 hours) and in the absence of media.

The third group of molecules included investigative or clinically approved inhaled non-antibiotic treatments for asthma, COPD, and CF. These molecules (discussed in Section 3.3.2), as classic druglike molecules with limited charge and no more than moderate hydrophobicity, exhibited only weak to nonexistent binding.

In both Sections 3.3.1 and 3.3.2, we discuss the binding of the drugs in structural and mechanistic classes rather than trying to make overarching statements about binding as a whole. This is because our analysis in Chapter 2 suggested that looking within specific molecule classes is more enlightening than searching for global effects.

Finally, in Section 3.3.3, we present a theoretical model to understand the pharmacokinetic impacts of mucus binding of inhaled drugs. It is a common hypothesis that binding to mucus, or mucoadhesion, will increase residence time of nanoparticles within a mucosal layer[136] and it therefore seems natural to expect that the same may be true of small molecules. We discuss the extent to which mucoadhesion is likely to impact residence time, and put into perspective the binding strengths observed in Section 3.3.2.

I thank Tahoura Samad for some pilot experiments testing the effect of HBE mucus on antibiotics and specifically for identifying the small number of bacteria present in the mucus. I also thank Gerardo Cárcamo-Oyarce for doing some of the initial experimental setup for the antibiotic binding experiments, which allowed me to be blinded while I performed the bulk of the experiments. All other results and analysis (binding experiments, qLCMS, and theory) were mine.

Finally, this chapter is not intended to present a single fully integrated story. This is because we found in Chapter 2 that in the absence of DAP or strong positive charge, global correlations between druglike small molecule chemical properties and binding to mucus are weak to nonexistent under our conditions. Thus, we provide an inventory of results in Sections 3.3.1 and 3.3.2 and encourage the interested reader to focus on whichever molecules are of particular interest. For example, the pyocyanin and/or antibiotic binding results are particularly relevant for microbiology and lung infection research, while the results for bronchodilators are of greater interest for asthma. However, we do consider our conclusion (Section 3.4) to be of general interest. In particular, we present five simple rules for predicting and understanding the binding of novel small molecules to mucus, with the hope that these rules will provide helpful guidelines for future researchers into mucosal biology.

3.2 MATERIALS AND METHODS

3.2.1 Materials

MUC5AC from fresh pig stomach scrapings, DNA, and alginate were from the same sources described in Chapter 2. BSM, BSA, and MC (15cP) were acquired from Sigma-Aldrich. CMC (90 kDa) was acquired from Acros Organics.

HBE mucus, generously provided to us by the Hill lab at the University of North Carolina, was harvested from human bronchial epithelial (HBE) cell cultures as previously described.[133,134,137–139] In brief, over a 6-week interval, confluent cultures developed cilia, generated a periciliary layer (PCL) surrounding the cilia, and formed a mucus layer. Washings from >100 cultures were pooled in 3,500 kDa cut off dialysis tubing (Fischer Scientific) and concentrated against Spectra/Gel (Spectrum Labs) to the designated % solids. Concentrated mucus was dialyzed against PBS overnight to ensure proper salt content.

All of the antibiotics, and pyocyanin, were acquired from Sigma-Aldrich, with the exception of aztreonam (TCI America) and iclaprim (WuXi AppTec).

The inhaled drugs were provided to us by Novartis.

3.2.2 Chemical structures

Chemical structures were drawn using ChemDraw version 18.0 (PerkinElmer).

3.2.3 Equilibrium dialysis

ED was performed as described in Chapter 2, in PBS buffer and 0.5% w/v biopolymer. Compounds were measured in two cocktail groups, given in Table 3.1 along with the concentration used for each molecule. Different molecular concentrations were used for the second group because the ionization efficiencies of the molecules varied, so the minimum concentration required for a linear qLCMS response varied as well. Also, for the first group (antibiotics and pyocyanin), I (the primary experimenter)

was blinded to which ED well contained which biopolymer until the qLCMS analysis was complete.

In the case of the group 4 molecules (the inhaled drugs from Novartis), two concentrations (varied by a factor of five) were used for each molecule. The lower concentration was used for the first four biological replicates, and the higher concentration was used for the second four. The higher concentration was required because the mass spectrometer's sensitivity dropped in the weeks between the first and second set of replicates. The concentrations were therefore increased to capture enough signal. However, even these higher concentrations only went up to a maximum of 5 µM per molecule, still a relatively low concentration unlikely to cause competition for binding sites.

Table 3.3.1 Compound cocktail group of each molecule and concentration(s) of each molecule used in binding assay. KF-80-NW22 and OF-80-NS22: investigational structures without standard names.

Molecule	Compound group	Concentration (µM)
Pyocyanin	1	1
Polymyxin B	1	15
Colistin	1	15
Erythromycin	1	1
Vancomycin	1	5
Levofloxacin	1	1
Ciprofloxacin	1	1
Aztreonam	1	1
Methicillin	1	1
Iclaprim	1	1
Trimethoprim	1	1
Tiotropium	2	0.2, 1
Nedocromil	2	1, 5
Formoterol	2	1, 5
Ipatropium	2	0.2, 1
KF-80-NW22	2	0.2, 1
OF-80-NS22	2	0.1, 1
Salbutamol	2	1, 5
Amiloride	2	0.2, 1

3.3   RESULTS AND DISCUSSION
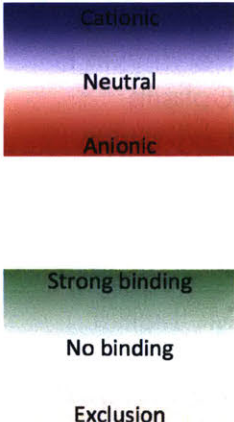
3.3.1 Pyocyanin and antibiotic binding

3.3.1.1 Relationship between pyocyanin and antibiotic binding to different biopolymers

First, we measured binding of pyocyanin and various mucus-relevant antibiotics to eight biopolymers using our ED method developed in Chapter 2. The data from these experiments (Table 3.2) was analyzable in two ways. First, in the same spirit as Fig. 2.3 and Fig. 2.4, we were able to compare the binding behavior to different biopolymers.

Table 3.3.2 ΔG measurements for each tested antibiotic and each biopolymer. HBE binding to ciprofloxacin was not measured consistently so was excluded.

Molecule	Charge	HBE	MUC5AC	DNA	Alginate	CMC	BSM	BSA	MC
Polymyxin B	5	-1.27	-0.92	-0.51	-0.53	-0.66	0.00	0.14	-0.32
Colistin	5	-1.42	-0.95	-0.56	-0.42	-0.37	0.12	0.06	0.12
Erythromycin	1	0.42	0.12	0.11	0.21	0.03	0.07	0.07	0.09
Vancomycin	1	0.40	0.29	0.05	0.02	0.25	0.04	-0.03	-0.25
Iclaprim	1	-0.66	-0.26	-0.03	0.40	0.01	0.08	-0.14	-0.14
Trimethoprim	1	-0.01	0.46	0.11	0.52	0.06	0.11	0.08	0.04
Pyocyanin	0	0.27	0.02	-0.92	0.25	0.05	0.01	-0.06	0.00
Levofloxacin	0	-0.10	0.12	0.02	0.42	0.14	0.06	-0.14	0.19
Ciprofloxacin	0		0.07	-0.10	0.34	0.15	0.06	-0.12	-0.16
Methicillin	-1	0.36	0.10	0.06	0.38	0.15	-0.02	-0.01	-0.01
Aztreonam	-2	-0.40	0.00	0.00	0.53	0.08	0.14	0.01	0.03

Cationic

Neutral

Anionic

Strong binding

No binding

Exclusion

While we had fewer molecules here (11) than in Chapter 2, we had even more biopolymers, eight in total: HBE mucus (which is a mixture of biopolymers including

mostly MUC5B and some MUC5AC[132]), pig gastric MUC5AC, DNA, alginate, carboxymethylcellulose (CMC; a polyanionic methylcellulose analogue), BSM, BSA, and methylcellulose (MC).

In order to summarize the similarities and differences in binding behavior, we considered the pairwise correlations in ΔG measurements (Fig. 3.1). One striking feature is that correlations are high between non-BSM polyanions (HBE mucus, MUC5AC, DNA, alginate, and CMC) and low between all other pairs. This is due primarily to polymyxin B and colistin binding, which is strong for these five polyanions

	HBE	MUC5AC	DNA	Alginate	CMC	BSM	BSA	MC
HBE	1	0.90	0.40	0.64	0.85	-0.28	-0.28	0.17
MUC5AC		1	0.58	0.82	0.90	0.04	-0.34	0.21
DNA			1	0.55	0.54	0.25	-0.14	0.11
Alginate				1	0.82	0.23	-0.54	0.38
CMC					1	0.07	-0.62	0.25
MBS						1	0.03	0.43
BSA							1	-0.10
MC								1
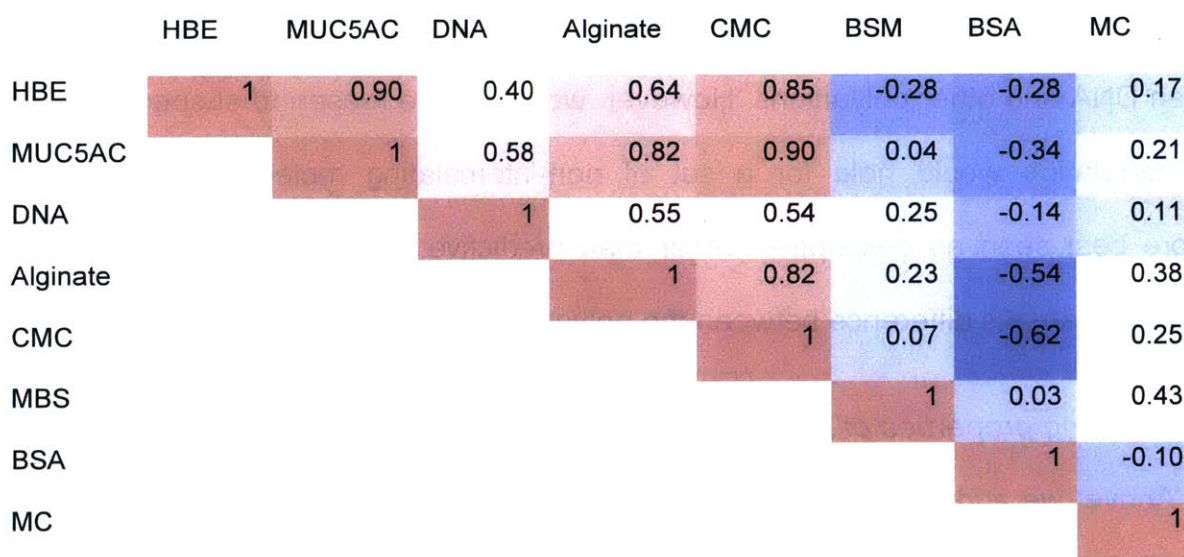
Figure 3.1. Pairwise Pearson correlation matrix for antibiotic and pyocyanin binding to each of the eight biopolymers (or biopolymer mixture, in the case of HBE). Higher correlation indicates similar binding behavior. Red: higher correlation, blue: lower/more negative correlation (note that the coloring is different from that in Table 3.3.2).

but not for BSM, BSA, or MC. Seven out of eight of these polymyxin binding results were as expected based on what we observed in Fig. 2.1b and Appendix A. The one interesting exception is the lack of apparent polymyxin or colistin binding to BSM, which may imply that BSM has lost a substantial portion of its charged moieties during the purification process. This could occur if BSM lost a substantial portion of its charged

moieties during Sigma-Alrich's proprietary purification process. In addition, it is not specified what the primary mucin type(s) are in BSM. This result is a useful reminder that simply using "mucin" as a mucus model is not sufficient to reproduce all of native mucin's properties.

We note that high correlation (such as CMC's 0.85 and 0.90 Pearson correlations with HBE mucus and MUC5AC respectively) does not necessarily mean that that the biopolymers are interchangeable in terms of binding properties because the sample size is not large enough to make such judgments. For example, pyocyanin's uniquely strong binding to DNA (Table 3.2) via intercalation[140] substantially reduces the correlation between DNA and other polyanions. However, we have little reason to suspect that this lower correlation would hold for a set of non-intercalating molecules. Fig. 3.1 is therefore best seen as descriptive rather than predictive, except in the sense that it depicts the dramatic difference between the polyanions and non-polyanions.

3.3.1.2 Binding properties of individual molecular classes

Above, we addressed overall trends of how various molecules bind differently to different biopolymers. Now, we take a second approach to analysis and quickly survey what the results might mean for molecular activity of specific molecules or classes of molecules in mucus (structures given in Fig. 3.2), as well as how well our results agree with previous work where it exists.

Pyocyanin, a phenazine. Our results showed that pyocyanin binds strongly to DNA, and does not bind other biopolymers. The DNA result agrees with previous observations that pyocyanin, a planar molecule (Fig. 3.2a), binds DNA via intercalation.[140] Interestingly, the lack of pyocyanin binding to any of the mucin models appears to be in

conflict with Darch *et al.*, who observed that BSM, but not DNA, slowed diffusion of pyocyanin in a cystic fibrosis mucus-mimicking medium.[25,141] However, this work used a concentration of BSM (5 mg/mL) nearly 10x higher than their DNA concentration (600 μg/mL). Despite the fact that pyocyanin is smaller than the mesh size of the medium, the authors attribute its slowed diffusion to increased viscosity driven by the presence of BSM.[25] This is reasonable because high biopolymer concentrations may still affect the interstitial fluid viscosity.[14] Thus, this result does not necessarily contradict ours, as the DNA concentration in their system may be too low to noticeably affect diffusion of pyocyanin by binding.

In a mucus context, binding between pyocyanin and DNA may be important because it may elevate the viscosity of DNA-containing mucus and also assist in *P. aeruginosa* biofilm formation.[140] Thus, pyocyanin may compound the negative impact of high DNA concentrations in CF lung mucus by making lung mucus both more tenacious and more susceptible to chronic infection.[140]

Polymyxins: colistin and polymyxin B (Fig. 3.2b). As discussed extensively elsewhere in this thesis (Chapter 2, Appendix A, and above), these cationic antibiotics bind to all polyanions we have tested, with the exception of BSM (which may or may not be polyanionic under the proprietary purification conditions) and are inhibited by multiple types of mucin. Since colistin is used as an inhaled therapy to treat lung infections, recognizing its binding to mucus is highly important.[135]
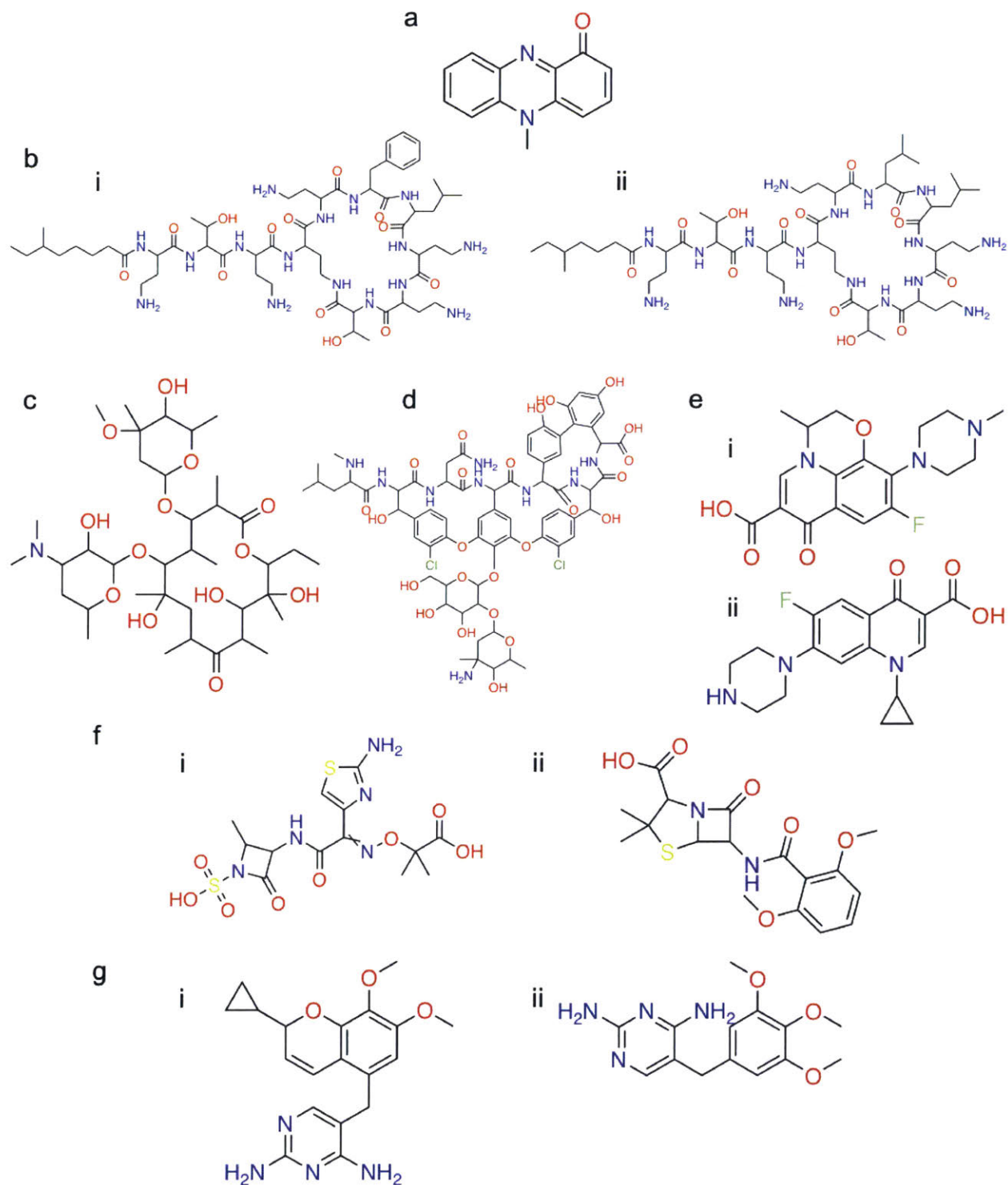
Figure 3.2. Structures of: a) pyocyanin. b) polymyxins, i) polymyxin B, ii) colistin. c) erythromycin. d) vancomycin. e) fluoroquinolones, i) levofloxacin, ii) ciprofloxacin. f) beta-lactams, i) aztreonam, ii) methicillin. g) dihydrofolate reductase inhibitors, i) iclaprim, ii) trimethoprim.

Erythromycin, a macrolide (Fig. 3.2c) While it is not used extensively to treat lung infections, erythromycin's close analogue azithromycin is frequently chronically prescribed to CF patients. Interestingly, it appears that the beneficial effects, which are unclear in any case, may be more due to azithromycin's anti-inflammatory properties than to its antibiotic activity.[142] Regardless, erythromycin does not bind to any of the biopolymers to any relevant degree, so binding to mucus is not likely to be of great relevance.

Vancomycin, a glycopeptide (Fig. 3.2d) Vancomycin used to treat methicillin-resistant *S. aureus* (MRSA) infections, a particular problem for CF patients.[143] It is currently undergoing clinical trials as an inhaled treatment, sponsored by Savara Pharmaceuticals. We do not observe a great degree of binding to any of the biopolymers, suggesting that vancomycin will not be inactivated by binding to CF sputum components.

The positive ΔG particularly in HBE mucus, represents an uptake ratio less than 1 and may be due to steric exclusion from the mucus due to vancomycin's relatively large size (1.5 kDa). The steric exclusion is consistent with the positive ΔG measured for erythromycin, another relatively large antibiotic (734 Da). This explanation is, however, somewhat speculative and the mesh size of 0.5% HBE mucus is undoubtedly greater than the <10nm diameters of these two drugs.

Boegh *et al.* studied the diffusive penetration of vancomycin through mucus, finding that its steady-state permeation was substantially reduced. This reduced permeation was attributed to interactions with mucus.[144] While this is possible, it is also possible that the reduced permeation was due to increased interstitial viscosity in the

thick (pig intestinal mucus or 5% w/v mucin) mucus or even steric exclusion from the mucus gel similar to what we might be observing as positive ΔG values in HBE mucus. Moreover, other differences between Boegh *et al.* and our study include that their 5% mucin concentration used was 10-fold higher than in our study, and that they used industrially purified porcine gastric mucin (PGM) rather than HBE mucus or natively purified gastric mucin (our MUC5AC).[144] As we have seen with industrially purified BSM, industrially purified PGM is likely to have substantially different interactive properties.

Fluoroquinolone antibiotics: levofloxacin and ciprofloxacin (Fig. 3.2e). Inhaled fluorquinolone antibiotics are under various stages clinical development for treatment of lung infections, particularly *P. aeruginosa* infections.[135] We do not observe noteworthy binding of either of these antibiotics to any biopolymer suggesting that binding is not likely to be a major factor in determining activity within the lung. This is in line with further investigations discussed at length in Appendix A.

Beta-lactam antibiotics: aztreonam and methicillin (Fig. 3.2f) While we are not aware of any investigations into inhaled methicillin, inhaled aztreonam is a clinically approved anti-*P. aeruginosa* treatment.[135] The one striking feature about this pair of molecules is that aztreonam, but not methicillin, bound to HBE mucus. The reason for this is unclear, as aztreonam is anionic and not particularly hydrophobic and does not bind to MUC5AC or any of the other biopolymers. We are not aware of any results, positive or negative, in the literature regarding aztreonam-mucus interactions.

We therefore see two possibilities. First, we cannot rule out the possibility that this result is a fluke. While the results for aztreonam were consistent and significantly below zero (one-sample t-test: $t=13.469$ on 2 d.f., $p=0.0055$), we made 87

82

measurements in this dataset so applying a Bonferroni correction gives p=0.476. Alternatively, and more interestingly, there may be something unique about the MUC5B, or another component, in HBE mucus that leads to binding to aztreonam.

Dihydrofolate reductase inhibitors: iclaprim and trimethoprim (Fig. 3.2g). Finally, we come to perhaps the most interesting pair of antibiotics in the dataset, iclaprim and trimethoprim. Iclaprim is currently under investigation by Motif Bio as an inhaled or intravenous therapeutic for the treatment of MRSA infections in CF. Based on the results from Chapter 2, we may suspect that these antibiotics will bind mucin because they both contain the DAP motif, and iclaprim particular would be expected to bind particularly strongly because it is more hydrophobic than trimethoprim. That said the effects of DAP and hydrophobicity of a molecule were weak in mucin suspended in PBS (the conditions here), so these characteristics of the molecules may not drive strong binding.

Iclaprim was in fact observed to bind to MUC5AC and, even more strongly, HBE mucus – the lowest non-polymyxin ΔG value for both MUC5AC and HBE mucus (Table 3.2). Trimethoprim, on the other hand, bound neither (the ΔG value was positive in the MUC5AC case, which is difficult to interpret). This difference is consistent with the DAP plus hydrophobicity trends in Chapter 2. This finding not only identifies a binding interaction of an important antibiotic, but also suggests a novel mechanism of binding– hydrophobicity combined with the DAP motif. These results imply that lung mucus may modulate iclaprim's activity, and further investigation into this possibility appears warranted.

Interestingly, one study showed that intravenously delivered iclaprim appears to accumulate in pulmonary epithelial lining fluid (PELF), a general designation for all fluid coating the lung epithelium, which includes mucus.[145,146] Our results suggest that one explanation for this could be that mucus soaks up iclaprim. However, the data on iclaprim accumulation in PELF is somewhat anomalous and requires further validation.[146]

Aminoglycoside antibiotics. Aminoglycoside antibiotics are a class of antibiotics which includes tobramycin, a frontline inhaled antibiotic for respiratory *P. aeruginosa* infections. Aminoglycosides are heavily cationic and have been observed to interact with polyanionic mucus components,[21,147,148] and we would have liked to include the study of them here. However, aminoglycosides are quite hydrophilic and are therefore not chromatographically separable from inorganic salt, rendering them not measurable using our qLCMS assay because inorganic salt heavily depresses mass spectrometry signal. We attempted to separate multiple aminoglycosides from the salt-heavy aqueous phase using an ethyl acetate extraction under basic conditions (to deprotonate the amine groups in aminoglycosides), but even under these conditions we were unable to extract sufficient aminoglycosides to make reliable measurements at reasonable (sub-millimolar) concentrations.

3.3.2 Non-antibiotic inhaled drugs

We next turned to the question of non-antibiotic inhaled drugs for lung diseases. For this set, we tested 17 molecules and were able to extract measurements from the nine of them described in Table 3.1. As we only tested the detailed binding of these molecules in HBE mucus, we proceed directly to the characterization of each of their

84

binding. The full dataset is given in Table 3.3 and structures in Fig. 3.3. We begin by considering amiloride and two of its analogues, as they are a particularly instructive case for considering the complex effects of mucus binding.

Table 3.3 ΔG measurements for each tested antibiotic and each biopolymer, with standard error of the mean in parentheses after each value (n≥3). KF-80-NW22 and OF-80-NS22 are amiloride analogues (Figure 3.3).

Molecule	HBE
Amiloride	-0.18 (0.20)
KF-80-NW22	0.05 (0.04)
OF-80-NS22	-0.06 (0.08)
Ipratropium	0.42 (0.16)
Tiotropium	-0.30 (0.24)
Salbutamol	0.03 (0.03)
Formoterol	-0.51 (0.15)
Nedocromil	-0.07 (0.23)
Fluticasone	-0.17 (0.17)
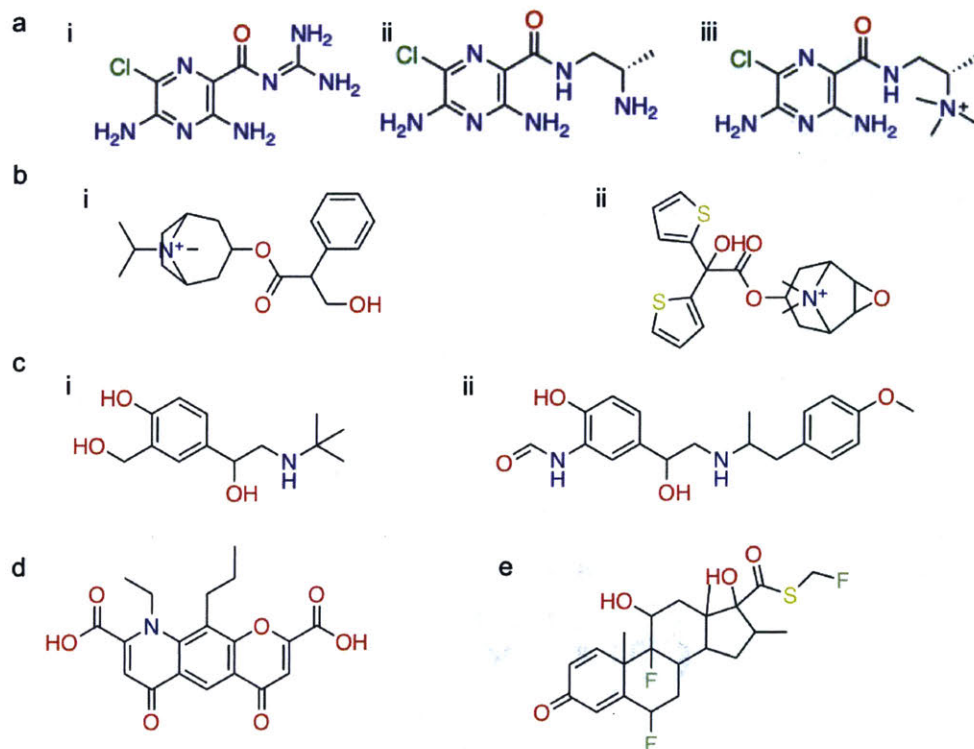
Figure 3.3 Structures of: a) Epithelial sodium channel inhibitors, i) amiloride, ii) OF-80-NS22, iii) KF-80-NW22. b) muscarinic antagonists, i) ipratropium, ii) tiotropium. c) beta-2 agonists, i) salbutamol, ii) formoterol. d) nedocromil. e) fluticasone

85

Epithelial sodium channel inhibitors: amiloride and analogues (Fig. 3.3a).

The epithelial sodium channel (ENaC) is an important regulator of respiratory mucus hydration and is therefore an important therapeutic target for CF and potentially other muco-obstructive diseases such as COPD.[149] Amiloride is an ENaC inhibitor which initially showed some promise as an inhaled therapy, but it was absorbed quickly from the lung into the bloodstream, thus severely limiting the duration of action.[150] As a result, aerosolized amiloride treatment yielded no significant clinical benefit to CF patients.[150] Thus, further optimization of ENaC inhibitors, in particular to increase time spent in the lung, would be an important advance. In principle, binding to mucus could accomplish this because while the activity of a mucus-binding drug may be initially reduced, the slowed diffusivity due to binding would slow clearance from the lung.

However, neither amiloride nor either of its two analogues showed any binding. While the amiloride ΔG is negative, it is within error of 0 and is still relatively low – we discuss the potential significance of this binding magnitude in section 3.3.3.

This weak to nonexistent binding is not surprising, in that these molecules are neither especially hydrophobic or cationic. While the shared chlorine-substituted aromatic ring is close to the DAP motif, it is not identical and in particular is not likely to be ionized at pH 7.4.

Bronchodilator muscarinic antagonists: ipratropium (short acting) and tiotropium (long acting) (Fig. 3.3b). Ipratropium and tiotropium are clinically approved inhaled bronchodilators with different durations of action but the same mechanism of action and similar quaternary amine-containing structures. Interestingly, tiotropium had a lower ΔG than ipratropium, though the potential mechanism for this is not clear

because both are hydrophilic and they have an identical +1 charge. Given these results, it is tempting to attribute at least some of its high duration of action to mucus binding. However, this is unlikely given the relatively low magnitude of effect; again, see Section 3.3.3. Furthermore, tiotropium's long acting behavior is believed to be due to tight binding to muscarinic receptors and resulting slow dissociation.[151]

Bronchodilator beta-2 agonists salbutamol (short acting) and formoterol (long acting) (Fig. 3.3c). Though salbutamol (also known as albuterol) and formoterol are both beta-2 agonists, their structures are quite different. Salbutamol does not bind HBE mucus, as expected given its single positive charge and only moderate hydrophobicity. Formoterol does appear to bind HBE mucus to some degree, although as with the muscarinic antagonists, the magnitude of the effect is unlikely to be tightly related to the difference in duration of action; again, see Section 3.3.3. It would, however, be interesting to test whether HBE mucus inhibits formoterol's activity.

Antiinflammatory mast cell stabilizer: nedocromil (Fig. 3.3d). Nedocromil, an anionic small molecule, is not observed to bind HBE mucus (Table 3.3). This result is expected because the molecule is anionic and not particularly hydrophobic.

Antiinflammatory glucocorticoid agonist: fluticasone (Fig. 3.3e). While fluticasone is highly hydrophobic, we do not observe strong binding (Table 3.3), in contradiction to the idea that strongly hydrophobic molecules bind mucus which inhibits mucus penetration.[20] However, it has been hypothesized that lipids rather than mucins may be the primary mediator of this binding.[110,152] Dialyzed HBE mucus undoubtedly contains far fewer lipids than intestinal mucus, the source where most studies on hydrophobic small molecule binding to mucus have been performed.[20] Thus, this result

supports the hypothesis that lipids drive the reduced diffusive permeability of purely hydrophobic molecules in mucus.

3.3.3 Theoretical considerations for the effect of mucus binding on drug pharmacokinetics

As we mentioned above, mucoadhesive materials are commonly developed to extend residence time at mucosal surfaces.[136] We therefore consider the theoretical question of whether the same could hold true for small molecule drugs.

We model the pharmacokinetics of inhaled drugs as a one-dimensional diffusion problem, in which drugs begin at the air-liquid interface at time $t = 0$ and must diffuse a distance L to the epithelium, which we model as an absorbing boundary because the bloodstream will dilute the drug to a negligible concentration (Fig. 3.4). We assume that the drug is perfectly membrane permeable and ignore drug binding to its target receptor, to uniquely isolate the effect of diffusion through the mucus layer. Membrane impermeability would increase lung residence time by preventing absorption into the bloodstream, and drug-receptor binding with high enough affinity (in particular, low off rates) would of course also increase duration of action. We consider the case where neither of these processes are limiting, which essentially represents an upper bound on the importance of diffusion in mucus for altering residence time. Throughout this section, we continue that theme – as the model contains many variables and we cannot consider every possible situation, we always err on the side of emphasizing the greatest possible effect of mucus binding.

It is important to note here that this model assumes a homogeneous delivery of the drug onto the mucus layer. If there is inhomogeneity in a molecule's distribution

throughout the lung, which could be due to inconsistent drug deposition, the length scale in question is the length scale of the heterogeneous distribution of the drug. Interestingly, others have tackled this particular question in the context of *P. aeruginosa* signaling in a CF infection model.[25] Examining heterogeneity in drug distribution is outside the scope of the model we present here.

Our model relies on several key equations. Firstly, the characteristic diffusion time τ that a drug takes to travel distance L, i.e. to be absorbed into the blood stream, is related to L and to the "effective diffusion coefficient" D_{eff} by:[153]

$$\tau \sim L^2 / \left(\pi^2 D_{eff} \right) \tag{3.1}$$
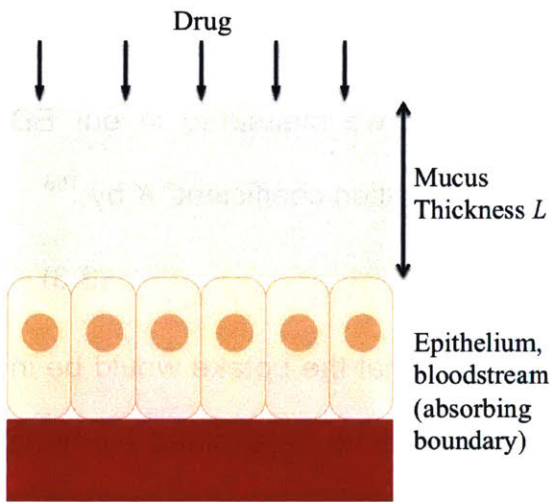
Figure 3.4 Diffusion model for inhaled drugs. Molecules are deposited uniformly on top of the mucus layer and diffuse through mucus of thickness L to the epithelium and bloodstream, which act as absorbing boundaries.

The effective diffusion coefficient of a drug is the diffusion coefficient when taking binding interactions into account. For simplicity, we assume a single, first-order,

reversible binding reaction with dissociation constant K_D and concentration of binding sites within mucus N_T. In this case, we have:[153]

$$D_{eff} = \frac{D_{free}}{1+N_T/K_D}$$ (3.2)

where D_{free} is the diffusion coefficient of the free, unbound drug. In the absence of steric effects, D_{free} will be the same as D_{water}, the diffusion coefficient in water, which is approximately predictable using the Stokes-Einstein equation and the known viscosity of water.[154] In practice, the interstitial fluid viscosity of respiratory mucus is generally higher than the viscosity of water, though likely not by more than an order of magnitude.[14,155]

Assuming N_T is high enough that binding sites do not become saturated (saturation of binding sites would reduce the potential impact of mucus binding by reducing the binding), the uptake ratio R_U (which is what we measured in our ED experiments) is related to the other quantities, and to the "partition coefficient" K by:[153]

$$R_U = K\left(1 + \frac{N_T}{K_D}\right)$$ (3.3)

The partition coefficient K is defined as, essentially, what the uptake would be in the absence of the binding reaction. Specifically, it is likely to be determined by steric hindrance and by Donnan partitioning, a non-specific electrostatic interaction that does not affect diffusivity.[153] As we discussed above without referring to K, steric hindrance based partitioning may be responsible for positive ΔG ($R_U<1$) measurements, particularly for vancomycin and erythromycin.

For a given measured $R_U>1$ (a situation where binding may be causing increased lung residence time due to binding), we consider an approximate upper bound on the effect of binding on diffusion, which occurs when $R_U=1$, so the sole contributor to R_U is

binding. This is an upper bound because while Donnan partitioning does not reduce effective diffusivity, binding does; we say an "approximate" upper bound because R_U may in fact be less than 1 in the case of steric partitioning, but we do not expect this effect to be large in the case of drugs in mucus so we neglect this possibility. We consider this upper bound case for two reasons: first, distinguishing binding and partitioning is not possible without explicit diffusion experiments or saturating out binding sites, neither of which we did. Second, we are interested in whether increasing lung residence time by slowing diffusion in mucus is feasible, so the appropriate way to address the binding-partitioning uncertainty is to assume diffusion is as slowed as it can possibly be. Considering this upper bound, we write:

$$R_U \geq 1 + \frac{N_T}{K_D} \tag{3.4}$$

We can now combine equations 3.2 and 3.4 to get:

$$D_{eff} \geq \frac{D_{free}}{R_U} \tag{3.5}$$

And from there we can relate the uptake ratio to the lung residence time:

$$\tau \leq R_U L^2 / \left(\pi^2 D_{free} \right) \tag{3.6}$$

The result of this theory is an intuitive result. Suppose R_U=3 for a drug. This implies that up to two-thirds of the drug is bound, so each drug molecule may be diffusing only one third of the time. Therefore, it should take up to three times as long to get anywhere, so if diffusion is limiting the lung residence time, the residence time should increase by a factor of up to three because it takes up to three times as long to make it to the epithelium and bloodstream.

We now pick a set of realistic values for the different quantities in equation 3.6, and evaluate in each case whether increased mucus binding is likely to have an impact on pharmacokinetics.

D_{free}: Matsui et al. observed that the diffusion coefficient of N-(3-oxododecanoyl)-L-homoserine lactone (3-oxo C_{12} HSL: a signaling molecule produced by P. aeruginosa, with MW = 297 Da) was:

- 49 $\mu m^2/s$ in 8% HBE mucus, a concentration characteristic of CF.

- 480 $\mu m^2/s$ in 2.5% HBE mucus, a concentration characteristic of a healthy mucus layer[133]

The diffusion coefficient of 3-oxo C_{12} HSL in 2.5% HBE mucus is also likely be the diffusion coefficient of 3-oxo C_{12} HSL in water. This is because the diffusion coefficient of sucrose (MW = 342 Da) in water is 523 $\mu m^2/s$, very close to 480 $\mu m^2/s$.[156] Because 3-oxo C_{12} HSL is a similar size to sucrose, their diffusion coefficients in water are also likely to be similar. In other words, for typical small molecule drugs, binding-independent diffusion is typically unhindered in healthy mucus layers with a D_{free} of 500 $\mu m^2/s$. A D_{free} 10x less than that, 50 $\mu m^2/s$, is probably the lower end; note that the diffusion coefficients measured in Matsui et al. were effective diffusion coefficients that took binding into account and thus are lower bounds on D_{free}.[133]

R_U: we will use a range representative of the range of R_U values that we measured above along with several values higher than any we have observed. We offer a reminder here that R_U is related to ΔG by $R_U = e^{-\Delta G}$.

L: the thickness of the lung mucus layer is both difficult and especially important to estimate precisely, because the lung residence time depends on the square of L. It

has been estimated at 55 μm[157] or less than 20 μm[158] and may depend on the airway diameter.[157] In the case of muco-obstructive lung disease, mucus plugs may form that substantially increase the thickness of the mucus layer at least where the plugs are present. We therefore include many possible values of L.

Table 3.4 presents some characteristic diffusion times given different choices for the parameters and equation 3.6.

Table 3.4 Lung residence times as a function of mucus thickness, D_{free}, and R_U

L (µm)	D_{free}	ΔG	R_U	τ (Penetration/residence time, min)
20	50	0.00	1	0.01
20	50	-0.69	2	0.03
20	50	-1.39	4	0.05
20	50	-2.08	8	0.11
20	50	-2.77	16	0.22
100	50	0.00	1	0.34
100	50	-0.69	2	0.68
100	50	-1.39	4	1.35
100	50	-2.08	8	2.70
100	50	-2.77	16	5.40
300	50	0.00	1	3.04
300	50	-0.69	2	6.08
300	50	-1.39	4	12.16
300	50	-2.08	8	24.32
300	50	-2.77	16	48.63
1000	50	0.00	1	33.77
1000	50	-0.69	2	67.55
1000	50	-1.39	4	135.09
1000	50	-2.08	8	270.19
1000	50	-2.77	16	540.38
20	500	0.00	1	0.00
20	500	-0.69	2	0.00
20	500	-1.39	4	0.01
20	500	-2.08	8	0.01
20	500	-2.77	16	0.02
100	500	0.00	1	0.03
100	500	-0.69	2	0.07
100	500	-1.39	4	0.14
100	500	-2.08	8	0.27
100	500	-2.77	16	0.54
300	500	0.00	1	0.30
300	500	-0.69	2	0.61
300	500	-1.39	4	1.22
300	500	-2.08	8	2.43
300	500	-2.77	16	4.86
1000	500	0.00	1	3.38
1000	500	-0.69	2	6.75
1000	500	-1.39	4	13.51
1000	500	-2.08	8	27.02
1000	500	-2.77	16	54.04

Table 3.4 suggests that only in extreme conditions is it possible in principle to use slowed diffusion to increase lung residence time past an hour. Certainly, where no mucus plug is present ($L \leq 1000$ μm), there appears to be little to no chance that binding alone will increase residence time for significant amount of time. Even for R_U values of 16, corresponding to ΔG = -2.77, a higher value than we have ever observed though it may be reachable in hyperconcentrated mucus layers with a greatly increased binding site density N_T, penetration time is be on the order of several minutes for a 100 μm-thick mucus layer. It is true that the "characteristic penetration time" is only a timescale estimate, and that after time τ there will still be some drug remaining in the lung. However, after several times τ the remaining concentration of drug is likely to be very small (independent of other residence time-increasing mechanisms such as plasma membrane impermeability, as discussed above).

Where macroscopic, millimeter-or-thicker plugs are present, diffusion though these plugs may possibly be limiting, and therefore changes in mucus binding strength could make a substantial impact on drug residence time and thus duration of action. However, recall the numerous assumptions we made to maximize the possible importance of diffusion. We assumed that off rates of the drug-receptor complex were high and that membrane permeability was high so the lung epithelial barrier was negligible. We also neglected the possible effects of cough, exercise, and mucociliary clearance that may serve to nondiffusively transport drugs. It is therefore unclear whether diffusion through mucus plugs could systematically increase residence time

throughout the lung, and this effect would likely be highly heterogeneous from patient to patient.

This is especially true for the relatively weak binding we observed for the inhaled drugs, with ΔG values all greater than -0.5, corresponding to $R_U < 1.6$. In order to determine the largest possible effect on residence time, let us suppose that the mucus layer is hyperconcentrated, 5% w/v as opposed to the 0.5% w/v we used in our experiments. With a tenfold increase in mucus content, the density of binding sites N_T would then increase by a factor of at most 10, less if concentration-driven aggregation leaves some binding sites inaccessible. Letting $R_U = 1.6$, equation 3.4 gives $N_T/K_D = 0.6$, so a tenfold increase in N_T would give $N_T/K_D = 0.6$ and $R_U = 7$. Even with binding at this strength, millimeter-thick mucus plugs would be required for diffusion-based lung residence time to be above an hour.

To summarize, the binding strengths that we observed for the inhaled drugs suggest that diffusion in mucus is not likely to effect an hours-long change to lung residence time following inhalation unless large mucus plugs are present. In such a case, diffusion may or may not be important, and by extension the effects of mucus binding on diffusion could also be important. Finally, drugs' fast mucus penetration may appear curious in light of the fact that mucus penetration is such an important and difficult problem for inhaled nanoparticle delivery.[1,131,159] The explanation for this apparent contradiction is threefold. First, nanoparticles are much larger than small molecules so they diffuse substantially more slowly, even in water. Second, nanoparticles experience steric constraints if they are large enough to be on the order of the mesh size. Third, binding interactions between nanoparticles and mucus can slow

the particles by a factor of tens of thousands because nanoparticles are capable of polyvalent interactions with the mucus mesh.[1,17,94,160] In contrast, small molecules are only capable of monovalent binding interactions so, as we have seen, mucus binding can generally only slow diffusion of small molecules by an order of magnitude at most.

3.4 Conclusions

We have examined the binding of numerous molecules in which interactions with lung may play an important role in their activity. Notably, we showed polymyxin antibiotics bind to all polyanions, that aztreonam may have some previously unsuspected binding to HBE mucus, and that the DAP-containing iclaprim may bind to MUC5AC and HBE mucus. Tiotropium and formoterol each appear to bind HBE mucus more than their analogues ipratropium and salbutamol.

With respect to the impact of binding, it may be something of a double-edged sword. In some cases, stronger binding is very likely to be antagonistic to therapeutic effectiveness. A classic case in point is polymyxins, which are a) inhibited by binding to polyanions, and b) membrane impermeable, so duration of action is not likely to be limited by absorption into the bloodstream. Indeed, colistin remains detectable in CF sputum as many as 12 hours after inhalation.[161]

On the other hand, some drugs such as amiloride can be delivered at high enough concentrations to produce a significant response immediately upon inhalation, but the effect is short-lifed.[162] In healthy patients, amiloride has an elimination half life of 23 minutes with evidence pointing to epithelial absorption as the primary factor.[163] In principle, a high enough dose of amiloride could increase the duration of action simply because an initially high concentration can survive more half-lives before the dose

drops below the therapeutic threshold. However, drastically increasing the dose would undoubtedly lead to large side effects.

Many strategies may be possible to increase the duration of action. The first is to simply design a drug with tighter binding to its target; with a low enough off rate, clearance would not be limiting. Unfortunately, tightening binding to a receptor is (to understate the situation) not always easy. The second would be to slow absorption by reducing membrane permeability and thus reducing transcellular absorption. This strategy is feasible for inhaled channel blockers such as amiloride, because they do not need to traverse any cell membranes to be active. Other strategies include slow drug dissolution[151] and encapsulating the drug in nanoparticles.[164]

Finally, mucus binding may slow diffusion across the mucus layer, thus slowing absorption. While mucus binding is likely to reduce activity as it does for colistin, the loss in activity during drug concentration peaks may be more than made up for by a longer period of therapeutic activity.

In Section 3.3.3, we therefore developed a theoretical model examining the feasibility of using mucus diffusivity to control drug pharmacokinetics. We found that in the absence of at least millimeter-thickness mucus plugs, this strategy does not appear feasible even given substantially higher binding strengths than we observe measure in Chapter 3. However, it is theoretically possible that mucus plugs could serve as controlled release reservoirs for inhaled drugs. Further pharmacokinetic studies would be required to address this possibility, though they have to be designed carefully because after several hours a drug would not be homogeneously distributed in airway

surface liquid but rather specifically localized at mucus plugs. Thus, inaccurate results may be possible depending on the sampling technique.

An appropriate context for investigating rational drug design to increase lung residence time by increasing mucus binding would satisfy several constraints. First, the drug would have to have a problematically short duration of action that would require an infeasible treatment regime to achieve a therapeutic benefit, but not such a short duration of action (<30 minutes where a chronically high drug concentration is desired, for example) that no amount of mucus binding could make a difference. Second, the drug should act on an intracellular target, so that simply reducing membrane permeability would be inadvisable. Third, the drug should treat a disease involving mucus obstruction and hyperconcentration so that there are large, diffusion-slowing mucus plugs present to retain the drug. Under these conditions, since penetration time of mucus plugs may be on the order of hours, altering the diffusivity of drugs even by less than order of magnitude could have an important impact on duration of action and dosing regimes. A fourth constraint is the timing of mucociliary clearance. If the drug is co-administered with therapies meant to promote mucus clearance via ciliary motion or cough, or is itself intended to promote mucus clearance, a tightly mucus-binding drug may simply be cleared with the rest of the mucus layer. This is an especially large potential constraint because mucolytics to reduce mucus plugging are an important class of therapies for muco-obstructive diseases.[57]

Having discussed a wide variety of chemistries, goals, and mucus types, we conclude with five simple rules for the simple question: given a new small molecule and a mucus context, how will the mucus affect the molecule?

99

1. If the molecule is +5 charged or higher, it will bind mucus. If the charge is +3 or +4 we do not have enough data to say for sure, but binding is likely.

2. If the molecule contains DAP, it is reasonably likely to bind mucus. More hydrophobic DAP-containing molecules are more likely to strongly bind.

3. If the molecule is a classic druglike small molecule without DAP, mucus binding is unlikely to be a significant consideration. The only way to know for certain is to test it, as global rules for binding prediction do not exist.

4. Binding to mucus may inhibit a drug's activity, with the effect approximately proportional to the strength of binding.

5. Binding to mucus is unlikely to affect lung residence time of a drug. However, under a restricted set of circumstances (thick, hyperconcentrated, and static mucus layer; anomalously strong mucus binding), lung residence time could theoretically be significantly extended via mucus binding. We are not aware of any documented examples of this occurring in humans or in animal models.

4 CHARGE INFLUENCES SUBSTRATE RECOGNITION AND SELF-ASSEMBLY OF HYDROPHOBIC FG SEQUENCES

4.1 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, many of which are interspersed with highly charged amino acid sequences. Despite the high density of charge in certain FG domains, if and how charge influences FG-domain self-assembly and selective binding of nuclear transport receptors is largely unexplored. Using rationally designed short peptide sequences, we determined that the charge type and identity of amino acids surrounding FG sequences impact the structure and selectivity of FG-based gels. Moreover, we showed that spatial localization of the charged amino acids with respect to the FG sequence determines the degree to which charge influences hydrophobic interactions. Taken together, our study highlights that charge type and placement of amino acids regulate FG-sequence function and are important considerations when studying the mechanism of nuclear pore complex transport in vivo.

4.2 INTRODUCTION

The nuclear pore complex (NPC) is a megadalton structure that controls the exchange of material between the nucleus and cytoplasm [166–168] through a combination of passive and facilitated diffusion. Above an ~30-kDa cutoff, proteins require complexation with nuclear transport receptors (NTRs) to efficiently translocate at rates of nearly 1000 molecules/s [169–172]. This fast translocation rate relies on transient interactions between hydrophobic phenylalanine-glycine (FG) domains on intrinsically disordered FG-containing nucleoporins (FG nucleoporins) and hydrophobic patches on NTRs [96,98,166,170,173–183]. Without this assistance from NTRs, proteins largely 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 for specific hydrophobic domains on NTRs and what parameters tune the molecular recognition necessary for facilitated transport remain open questions. Sequence analysis and MD 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 NPC [99,100,169,171,184,185]. However, due to the complexity and redundancy of FG nucleoporins within NPCs, systematic biochemical dissection of the amino acid space surrounding individual FG sequences has remained an experimental challenge.

Here, we investigated whether 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. In peptide research, rationally designed peptides and polypeptides have previously been helpful to identify individual amino acids or short peptide sequences important to the function of much larger polymer-based biomaterials such as extracellular matrices [186], silk proteins [187,188], and elastinlike polypeptides [189–191]. Here, we applied this approach and created systematically varied peptide sequences to identify how amino acids surrounding FG domains in native nucleoporins may contribute to selectivity. Although peptides may not recapitulate all properties of the original protein, they are a suitable model system for this study because they yield insight into the contributions of individual amino acids and their positioning to overall protein function, a task not feasible with intact proteins. Our data demonstrate that the identity and charge of amino acids surrounding FG sequences can impact the structure and stiffness of FG-based gels. Moreover, we determined that charged amino acids enable FG-containing peptide gels to discriminate substrates of varying charge and hydrophobicity. Last, we found that the distance at which a charged amino acid is localized with respect to the FG sequence determines the degree to which charge influences hydrophobic interactions. Together, our data suggest that FG domain function may be determined by the placement and type of amino acids surrounding FG sequences.

4.3 MATERIALS AND METHODS

4.3.1 Sequence analysis

The logo of the consensus sequence was generated using Berkeley's WebLogo software [192] by aligning 15 repeats according to FSFG as reference. Conserved fraction was calculated for positive charges (number of cationic residues K or R)/(15 repeats) and similarly for negative charges (number of anionic residue E or D)/(15 repeats) at each position indicated.

4.3.2 NTF2 expression and labeling

pQE30-NTF2-6xHis was a gift from the Görlich lab and transformed into DH5α cells for cloning. A cysteine (C) was inserted for fluorescent labeling after the 6× His-tag using standard site-directed mutagenesis with the two primers (5′ - CTCAGCTAATTAAGCTTAGCAGTGATGGTGATGGTGATGAGATCTG-3′ and 5′ - CAGATCTCATCACCATCACCATCACTGCTAAGCTTAATTAGCTGAG-3′) to form pQE30-NTF2-6xHis-C. The W7A-C mutant containing the terminal cysteine was created by using pQE30-NTF2-6xHis-C and applying standard site-directed mutagenesis with primers 5′ -TGAGGAGCCAATTTGTTCCGCGATCGGTTTATCACCCATG-3′ and 5′ -CATGGGTGATAAACCGATCGCGGAACAAATTGGCTCCTCA-3′ . Expression of NTF2-C and W7A was completed in OverExpress C41(DE3) cells (Lucigen, Middleton, WI) and purified using standard nickel column purification and ion-exchange columns. Labeling of NTF2-C and W7A-C was completed using Fluorescein-5-maleimide (catalog No. F150; Thermo Fisher Scientific, Waltham, MA) in accordance with the

manufacturer's protocol; labeled product was gel purified. Labeling efficiency was ~50% (data not shown). The final concentrations of NTF2 and W7A for transport were adjusted to 10 μM with 10% of the population labeled.

4.3.3 Peptide and gel preparation

Unless specified otherwise, all chemicals were obtained from Sigma-Aldrich (St. Louis, MO). Peptides were prepared by the Koch Institute Biopolymers and Proteomics Facility (Massachusetts Institute of Technology, Cambridge, MA) and Boston Open Labs (Cambridge, MA). All peptides were HPLC-purified unless specified otherwise, desalted using reverse phase HPLC with 0.05% trifluoroacetic acid (TFA), and lyophilized after synthesis with >95% purity. For fluorescently labeled peptides, a 5-carboxyfluorescein (Anaspec, Fremont, CA) fluorophore was added to the N-terminus, and the C terminus was modified to be an amide. The dye labeling protocol was as follows: peptide-resin that contained an N-terminal free amine but was otherwise fully protected was washed six times with dimethylformamide (DMF) followed by six washes with dichloromethane (DCM) and dried. The material was subsequently reconstituted with dry DMF to reswell the resin. The dye-labeling cocktail consisted of a 4× molar excess over peptide resin of 5-carboxyfluorescein, N,N ′ -diisopropylcarbodiimide (DCM), and hydroxybenzotriazole reconstituted in a minimal volume of dry DMF and stirred overnight at room temperature in the dark. The cocktail was allowed to preactivate for 30 min before being added to the peptide-resin. Dye-labeled peptide resin was washed six times with DMF, then six times with DCM, and dried in preparation for standard luorenylmethyloxycarbonyl chloride cleavage and deprotection. All quality-control analyses for purity were provided by the Massachusetts Institute of

Technology's Proteomics facility or by Boston Open Labs (Data S1 in the Supporting Material). Neutral hydrophilic (n) and hydrophobic (n) fluorescent reporters could not be HPLC-purified due to aggregation and were used as crude samples.

Fluorescent peptides were diluted into 200 mM NaCl with 20 mM HEPES [pH 7] at 10 μM final concentration for diffusion experiments. Gel peptides were all dissolved in 20 mM NaCl, 20 mM HEPES [pH 7] at 2% (w/v). To facilitate solubilization and gel formation, peptides were vortexed for 30 s and briefly sonicated in a model No. 2510 Bath Sonicator (Branson Ultrasonics, Danbury, CT) to reduce aggregation.

4.3.4 Capillary diffusion assay and analysis

Borosilicate square capillaries (1.5 inch) with 9-mm cross-sectional width (Catalog No. 8290; Vitrocom, Mountain Lakes, NJ) were loaded by piercing premade hydrogels. Ten-micromolar solutions (200 mM NaCl, 20 mM HEPES [pH 7]) of fluorescent peptides were injected into the capillary and sealed with a 1:1:1 (by weight) mixture of petroleum jelly, lanolin, and paraffin. Time lapses of peptide diffusion were taken at 1-min intervals for up to 5 h on a Ti Eclipse inverted microscope using either a CFI Plan UW 2× or an AxioObserver D.1 with an EC Plan-Neofluar 1.25×/0.03 WD = 3.9 (Nikon, Melville, NY) and a model No. C11440-22CU camera (Hamamatsu, Hamamatsu City, Japan). All fluorescence profiles were obtained by averaging the fluorescence intensities across the width of the capillary in the software MATLAB (The MathWorks, Natick, MA). Normalized concentration profiles were obtained by normalizing fluorescence intensities to the bath concentration of the capillary at the initial time point. The fluorescence signal was linear up to 50 μM (Fig. S2) when testing FAM alone in FGAK gels. To plot concentration profiles, the signal was not adjusted past the

saturation point; data therefore represent the lower bound of the actual concentration of reporters accumulating in the gel. All data represent at least three independent replicates. Student's 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 time course in a region of the capillary on the gel side of the interface over a 100-min window. To achieve this fit, we numerically solved the diffusion equation for the concentration of probe c:

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

using MATLAB's pdepe function (The MathWorks). The initial condition $c(x,0)$ was set by the concentration profile at the first timepoint, 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 the simulated and actual concentration profiles was the value reported; minimization of error took place iteratively using a modified gradient descent algorithm. The window from 150 to 250 min was generally used for fitting, but for particularly fast-diffusing probes (D > 1000 μm^2/min), windows of 50–150 or 30–130 min were 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 precise fitting. Fitting was designed to take place away from the interface to avoid interfacial effects, and the interface was determined in one of two ways. If there was an interfacial peak, fitting began at the x value to the right of the peak where the concentration reached 0.5 times the maximum concentration. If there was no interfacial peak (generally with a hydrophilic probe and a

108

charge-matched gel), the interface was identified as the main inflection point of the initial concentration profile, because the buffer was initially loaded with peptide and the gel was not. The inflection point was identified as the minimum (most negative value) of the numerical first derivative. Fitting was started 20 pixels to the right of the interface. Each fit region was also manually inspected to confirm that it was located at the approximately correct location. Examples of the diffusion-coefficient analysis across the time series as well as the general fit appear in Fig. 4.S3.

4.3.5   Nile Red fluorescence assay

Using the diffusion methods mentioned above, Nile Red dye (N1142; Thermo Fisher Scientific) was loaded into capillaries at 10 μg/mL in 200 mM NaCl, 20 mM HEPES [pH 7].

4.3.6   Rheological testing

Rheological tests were performed on a model No. MCR 302 Rheometer (Anton Paar, Graz, Austria) in a cone-plate geometry with a 25-mm diameter, 1° cone angle, and 51-μm truncation. The temperature was maintained at 25°C and evaporation was controlled with an $H2O$-filled solvent trap. To identify the linear regime, amplitude sweeps were conducted at ω = 10 rad/s from $\gamma 0$ = 0.01–100% 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.

4.3.7   Transmission electron microscopy

Images were taken with a JEOL-1200 transmission electron microscope (JEOL, Akishima, Tokyo, Japan). Gels were formed as described above and spotted onto glow-

discharged carbon-coated copper Formvar grids (Ted Pella, Redding, CA). Excess liquid and gel were removed with paper or parafilm (Whatman, Maidstone, United Kingdom). Grids were submerged in 1% uranyl acetate, then blotted and air dried for 15 min before imaging. Images of gels are representative of at least four images of each gel type.

4.3.8 Phenyl-sepharose chromatography

Fluorescent reporters were dissolved in 200 mM NaCl, 20 mM HEPES [pH 7] and loaded into high-prep phenyl FF 1 mL-capacity phenyl-sepharose columns (GE Healthcare Lifesciences, Pittsburgh, PA) equilibrated with three volumes of 200 mM NaCl, 20 mM HEPES [pH 7]. For elution, flow rates were set to 1 mL/min with 0.5 mL fraction volumes. Fraction concentrations are representative traces of elution profiles.

4.4 RESULTS

To determine the relevance of charged amino acids in FG domains, we chose the yeast nucleoporin Nsp1 as a model system because it is essential in *Saccharomyces cerevisiae* [193] and contains a repeating subsequence (284–553) with a high density of charge (Fig. 4.1a). Sequence analysis of 15 repeats of FG domains in Nsp1[284–553] revealed a nonuniform distribution of charge in the sequence space separating FG sequences (Fig. 4.1b). Cationic residues appear near the center and edges of the repeat, whereas anionic amino acids reside at least three positions away from FSFG sequences. Moreover, several highly conserved lysine (K) residues are situated 2–3 amino acids away from the FG sequence, at positions 6 and 17, and a conserved

glutamic acid (E) appears at position 9 (Fig. 4.1a). The nonuniformity in conserved

charge distributions and amino acid identity suggests that the charge and biochemical

A

Nsp1 Consensus Repeat

B

C

Charge Relevance

FGAK
FSFGAKAAKAFSFG

FGAE
FSFGAEAAEAFSFG

FGAS
FSFGASAASAFSFG

D

Spatial Localization

FGK
FSFGKAAAAKFSFG

FGA₂K
FSFGAAKKAAFSFG

E

Sidechain Biochemistry

FGAR
FSFGARAARAFSFG

FGAD
FSFGADAADAFSFG
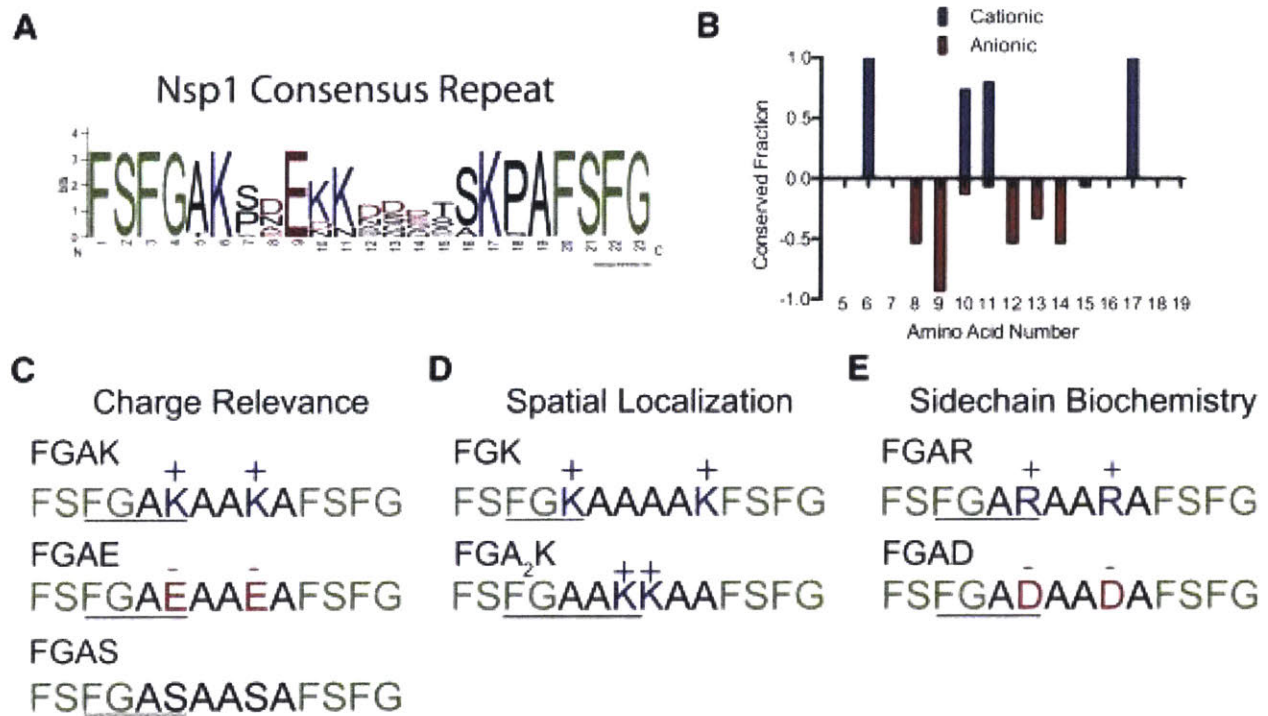
Figure 4.1 Identification of conserved repeat sequences and design of simplified FG self-assembling peptides. (A) Shown here is a conserved sequence identification of the C-terminal end of the essential yeast nucleoporin Nsp1. The letters in the top position indicate the most conserved residue at each location and represent the consensus sequence. (B) Shown here is conservation of charge between FG sequences in Nsp1 repeats. Positive or negative values refer to the presence of cationic or anionic residues, respectively. Amino acid numbers correspond to position in the consensus repeats, as in Fig. 4.1A. (C) Depicted here are designed peptides consisting of 14 amino acids with the sequence structure FSFGAXAAXAFSFG, where X represents the substituted amino acids K, E, or S. These peptides were used to determine how the presence of charge affects FG-mediated selectivity. (D) Depicted here are designed peptides in which K is moved immediately adjacent to (FGK) or placed three amino acids away (FGA2K) from FSFG domains to test how the spatial localization of charge influences FG selectivity. (E) Depicted here is a class of peptides in which K is substituted with R and E is substituted with D to determine how biochemistry affects FG-mediated selectivity and self-assembly.

properties of amino acid side chains may play a role in governing how neighboring FG

sequences respond to environmental substrates.

To test how the type and placement of charge as well as side-chain chemistry of neighboring amino acids affect FG function, we synthesized the consensus Nsp1 peptide sequence and rationally designed 14-amino acid variations of it. All engineered peptides consisted 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 K, E, or a neutral serine (S) (Fig. 4.1c). Because K is the most conserved residue in this wild-type Nsp1 subsequence, we use the K-containing peptide (the FGAK peptide) as the reference for all other experimental comparisons. To determine whether the substrate-binding properties of FG domains depend upon the type of neighboring charge, we synthesized the anionic peptide FGAE, and as a control the neutral FGAS peptide in which each K was converted to S. To test how the positioning of K relative to the FG sequence affects FG function, we designed two variants of FGAK in which the K was placed directly adjacent to the FG domain or separated by two alanine (A) residues (peptides FGK and FGA$_2$K, respectively; Fig 4.1d). Last, because only certain charged amino acids such as K and E exhibit 100% conservation at positions 6, 9, and 17 (Fig 4.1a), we hypothesized that amino acid biochemistries may also regulate FG-based molecular recognition. To test how the chemical structures of amino acid side chains affect the structure and function of FG domains, we designed a third class of peptides in which K was replaced with cationic arginine (R; FGAR) and E was replaced with anionic aspartic acid (D; FGAD) (Fig 4.1e). With these 14-amino acid nucleoporin-based peptides, we evaluated how single amino acid substitutions alter the selective binding and self-assembling capabilities of individual FG domains.

4.4.1 Characterizing the effect of charge on FG-mediated self-assembly

Before studying selective recognition by FG-based peptides, we first tested whether they form gels. We used a concentration of 2% (w/v) for each of the peptides, which corresponds to 28 mM FG sequence, a value that is well within the range estimated for densely packed NPCs [181,194,195]. To quantify gelation, we measured the stiffness of the resulting material using small amplitude oscillatory frequency sweeps. We report the storage (G') and loss (G'') moduli of the peptide solution (Fig. 4.2). A gel forms when $G' > G''$, which indicates successful self-assembly of a stable network of peptides. The consensus peptide sequence of Nsp1 (Fig. 4.1a) was unable to form a gel and flowed when inverted (data not shown), so it was not suitable for further analysis.

For the engineered FG sequences, we established that the reference peptide FGAK forms a hydrogel with a stiffness of 10^4 Pa across the frequencies tested (Fig. 4.2a). To ensure that the FG sequences were responsible for the self-assembly process, rather than the high density of A residues, which can promote stable self-assembly [196], we converted F to S (SGAK). With this substitution, the peptide remained in the aqueous phase and no longer exhibited a dominant storage modulus (Fig. 4.S4a), suggesting that the F within the designed peptides provides the necessary hydrophobic interactions for gelation, as for intact FG nucleoporins [98,173]. Reversing the charge via an E (FGAE peptides) revealed that the gel-forming properties are maintained with stiffness 5–9 KPa at the frequencies tested (Fig. 4.2b). To determine whether the charge is responsible for maintaining the hydration of the gel, we compared the material to the solution of neutral FGAS peptides. Without the charge, FGAS precipitated out of

solution (Fig 4.S4b), indicating that the presence of charge is essential in maintaining a hydrated network of FG sequences. However, it appeared that too much charge could prevent gelation, presumably by increasing the solubility, as with the consensus sequence peptide. Without electrostatic repulsion in the gels, the network collapses and
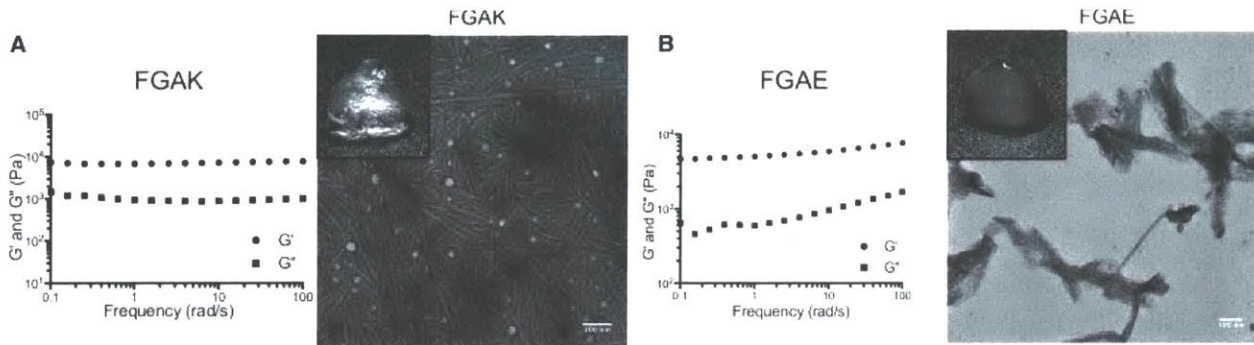
Figure 4.2 FGAK and FGAE self-assembly. Frequency sweeps of (*A*) FGAK gel and (*B*) FGAE gel, with *G'* (storage) and *G''* (loss) moduli reported at 2% (w/v). (*Insets*) Shown is transmission electron microscopy of self-assembled peptides and macroscopic gel. For both FGAK and FGAE peptides, *G' > G''*, indicating gel formation. The replacement of K with E alters the self-assembly properties of FG sequences to form different structures and gels with varying stiffness.

forms a precipitate. Taken together, our results suggest that the sequence adjacent to the FG sequence encodes information that impacts self-assembly and substrate binding by FG domains.

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

To determine whether FG-based peptide gels reconstitute the selective binding properties of native FG nucleoporins, we selected NTF2 as a model receptor because it contains an essential tryptophan (W7) that is required for binding FG domains. We replaced W with an A (NTF2^{W7A}) to ablate FG-binding capabilities [177], enabling us to test whether the FG domains are available in the gel for NTR binding. As cationic native and

114

engineered nucleoporin gels are predicted to bind and facilitate the selective transport of native NTRs such as NTF2 [99,100], we first determined whether the positively charged FGAK gel preferentially selected for NTF2 compared to the NTF2[W7A] mutant. To prepare gels for selective transport assays, we dissolved FGAK peptides at 2% (w/v) as before
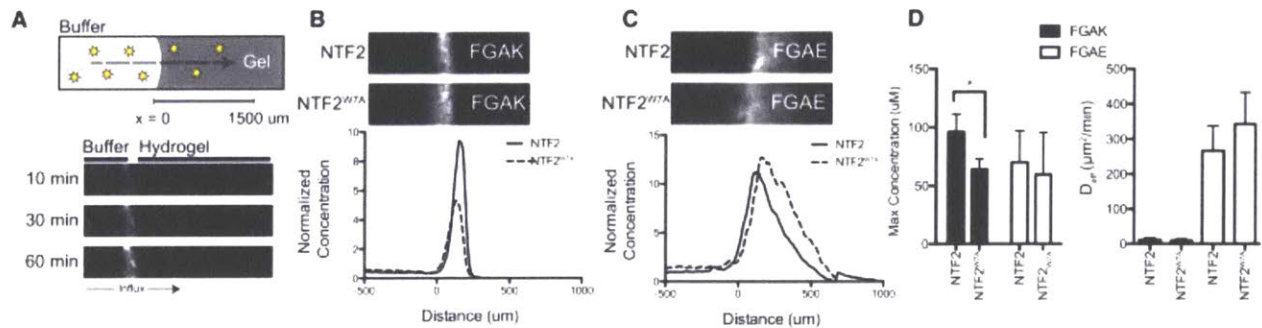
Figure 4.3 Molecular selectivity of the native transport receptor NTF2 in FGAK and FGAE gels. (A) Top, a schematic of a slab 1D transport system within a sealed capillary, which provides no flux boundary conditions during the diffusion of fluorescent molecules (stars). Bottom, the interface of the gel is reflected by the increase in fluorescence signal between the buffer and hydrogel. (Band C) Depicted here is transport of representative fluorescently labeled NTF2 and NTF2W7A proteins into FGAK and FGAE gels at 5 h. FGAK is able to preferentially recognize NTF2 over NTF2W7A mutant whereas FGAE gels are unable to. (D) Left, shown here is maximum concentration of labeled proteins at the interface. Right, shown here is effective diffusion coefficients of labeled proteins. Error bars are SDs of at least three independent replicates. *p < 0.05, unpaired Student's t-test.

and loaded the material into capillaries. Fluorescently labeled NTF2 or NTF2[W7A] was then injected into the capillary and sealed to create a 1D diffusion chamber; the diffusion profile was monitored for up to 5 h at 1-min intervals (Fig. 4.3a).

To quantify the extent to which the gels differentiated between NTF2 and NTF2W7A, we determined the maximum accumulation of the transport receptors at the gel interface and their effective diffusion coefficients inside the gel using time-evolving fluorescence profiles. Lower effective diffusivities and higher accumulation indicate stronger interactions with the gel. For a detailed description of diffusion coefficient

115

analysis for these time series, refer to the Materials and Methods and Fig. 4.S3. NTF2 accumulated at higher concentrations (mean ± SD: 96.2 ± 15.2 µM) at the gel interface than did the NTF2W7A mutant (63.9 ± 8.8 µM) (Fig. 4.3, b and d). Thus, the FGAK peptide has a binding preference for NTF2 over the NTF2W7A mutant, and appears capable of recognizing a single hydrophobic amino acid difference. Although NTF2 displayed increased enrichment at the FGAK gel interface (Fig. 4.3d), its effective diffusivity did not significantly differ from that of the NTF2W7A mutant (NTF2, 10.6 ± 6.5 µm^2/min; NTF2W7A 9.9 ± 4.9 µm^2/min). This experiment illustrates that the minimal in vitro system does not fully reconstitute the selective transport observed in the NPC in vivo, but it enables the characterization of the first step of selective transport in native NPCs, namely, the requirements for the initial selective binding to FG-based gels, which is the focus of this work.

To test whether NTF2 binding is sensitive to charged residues neighboring the FG domain, we determined whether NTF2's interaction with FGAE recapitulated the selective properties of FGAK, despite the reversal of charge. Fig. 4.3c and d, shows that FGAE selects for NTF2 (69.9 ± 27.1 µM) and NTF2W7A (59.5 ± 36.0 µM) equally at the interface, with no significant difference in diffusion coefficients (NTF2 266.5 ± 71.1 µm^2/min and NTF2W7A 342.7 ± 91.0 µm^2/min). These results indicate that the FGAE gel is less effective than the FGAK gel at differentiating between the native and mutant forms of NTF2. The FGAK and FGAE gels exhibit structural differences at the microscopic level (sheets versus fibers; Fig. 4.2, a and a) and the macroscopic level (insets to Fig. 4.2, a and b). Therefore, it is possible that due to structural differences, the FG sequences are not exposed and are unavailable for binding by NTF2 in the

FGAE gel. To test this hypothesis, we determined that Nile Red dye, which fluoresces within hydrophobic environments, was detected in both FGAK and FGAE gels (Fig. 4.S5): the uncharged hydrophobic dye interacted with the FG domains within the peptide-based gels despite the structural variation, showing that hydrophobic domains are available for binding by small molecules such as Nile Red. These results suggest that K and E may regulate the selective properties of the FG sequences from a charge interaction perspective for specific classes of molecules, as opposed to simply altering the microstructure of the gel.

4.4.3 Presence of charge regulates selective recognition by FG domains

Although NTRs such as importin β (Impβ) and NTF2 have well-characterized specificity for particular FG-nup sequences, intact receptors contain multiple binding pockets with varying affinities and charge distribution, or require dimerization for function. These factors complicate the systematic analysis of how charge contributes to FG function. Hence, as a replacement for complex NTRs in our minimal in vitro model, we designed fluorescent peptide reporters with defined spatial arrangements of charged and nonpolar amino acids to systematically test the contribution of charge in hydrophobic selectivity (Fig. 4.4a). The first two reporter peptides harbored three F residues, creating a hydrophobic tail for binding to FG sequences, but one contained an adjacent anionic sequence composed of E residues (termed hydrophobic (−)), whereas the other contained cationic K residues (hydrophobic (+)). As controls, we engineered two more reporters in which F residues were converted to hydrophilic asparagine (N) residues, termed hydrophilic (− or +); these control peptides should not interact with FG domains. To confirm that the synthetic hydrophobic reporters interacted with the

aromatic phenyl group in FG domains in the gels independently of their charged

domains, we used phenyl-sepharose columns to assay hydrophobic interactions. This

assay was previously employed to isolate intact NTRs from cell lysates [175]. Both

hydrophobic reporters displayed longer retention times than their hydrophilic
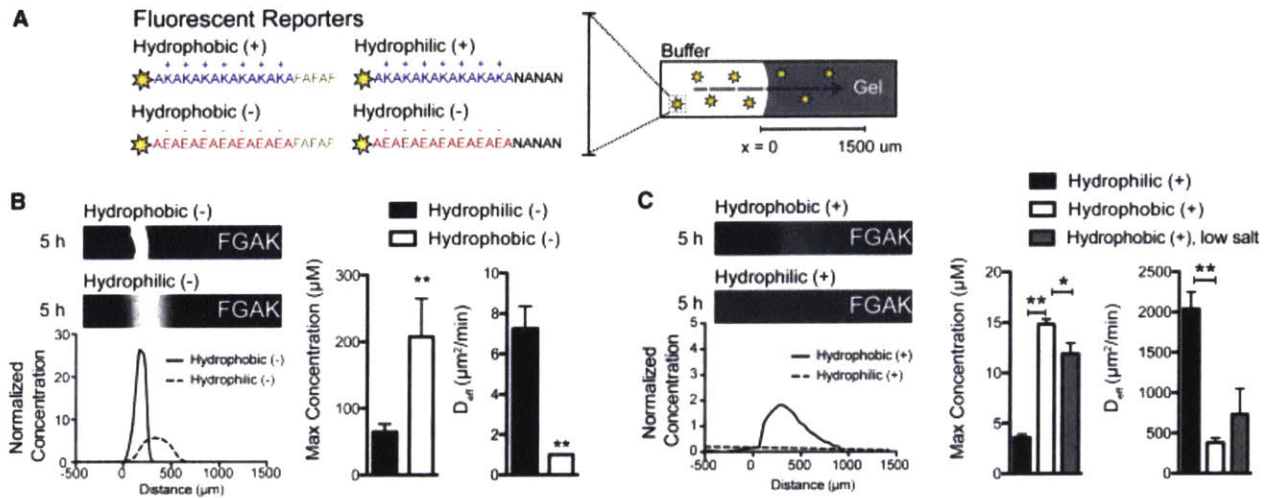
Figure 4.4 Charge affects hydrophobic molecular recognition in cationic FGAK gels. (*A*) Shown here is a schematic of the four fluorescent reporters with a charged domain and a hydrophilic (N) or hydrophobic (F) tail. (*B*) Given here are representative fluorescence images of hydrophobic (−) and hydrophilic (−) reporters in the FGAK gel, showing the increased selectivity for hydrophobic (−) reporters over hydrophilic (−) reporters. (*C*) Depicted here are representative fluorescence images of hydrophobic (+) and hydrophilic (+) reporters, which contain a cationic tail, into FGAK gel showing the increased selectivity for hydrophobic (+) reporters over hydrophilic (−) reporters. A low-salt condition (20 mM NaCl) was added for the hydrophobic (+) reporters to show how electrostatic screening modulates hydrophobic interactions for hydrophobic (+) reporters. For (*B* and *C*), error bars are SDs of at least three independent replicates. $^*p < 0.05$ and $^{**}p < 0.01$, unpaired Student's *t*-test.

counterparts (Fig. 4.S6), indicating that the hydrophobic reporters interact with the

phenyl groups on the sepharose independent of the displayed charge.

We used the FGAK gels to investigate whether the two hydrophobic fluorescent

reporters underwent differential uptake into the gel. The cationic FGAK gel interacted

with the hydrophobic (−) reporters, and accumulated inside the gel. The maximum

118

concentration of peptide at the interface was approximately threefold higher for the hydrophobic (−) reporter (64.6 ± 12.1 µM) than the hydrophilic (−) reporter (207.6 ± 57.2 µM) (Fig. 4.4b), which does not contain a hydrophobic tail, suggesting that FG sequences in FGAK recognize reporter F residues when the reporter is negatively charged. The contribution of hydrophobic interactions is further corroborated by the strong reduction in the effective diffusion coefficient from 7 μm^2/min for hydrophilic (−) to <1 μm^2/min for hydrophobic (−) (Fig. 4.4b), indicating that the hydrophobic interactions induce tighter binding in the context of electrostatic attraction. We note that the interactions are so strong that the hydrophobic (−) reporters essentially do not diffuse over the 5-h period analyzed, suggesting that the binding is irreversible. In contrast, when the reporters contained a cationic tail (Fig. 4.4c), the hydrophobic (+) and hydrophilic (+) reporters diffused into the gel with coefficients of 381.0 ± 57 μm^2/min and 2038.0 ± 211.8 μm^2/min, respectively. The hydrophobic (+) reporters bound 1.5 times above the original bath concentration at the interface (14.81 ± 0.5 µM), whereas the hydrophilic (+) reporters equilibrated (3.6 ± 0.3 µM) with no discernible partitioning (Fig. 4.4c). Thus, hydrophobic interactions can occur in electrostatically repelling environments, but the interactions are much weaker than those detected with the hydrophobic (−) reporters. Moreover, lowering the salt concentrations from 200 mM NaCl to 20 mM (to decrease electrostatic screening and to increase electrostatic repulsion) decreased the accumulation of hydrophilic (+) reporter at the interface (11.9 ± 1.1 µM), whereas the diffusion coefficient (732 ± 317.7 μm^2/min) trended upward (Fig. 4 C), indicating that repulsion weakens the overall strength of the hydrophobic interactions with the FG domains. These results show that the K residues in FGAK gels

distinguish between two substrates that contain the same hydrophobic domain but with different surrounding charge types. In particular, K allows for increased binding to hydrophobic domains with neighboring anionic residues.

4.4.4 Spatial localization of lysine affects FG-mediated self-assembly and binding selectivity

Our analysis of the Nsp1 repeat consensus sequence (Fig. 4.1, a and b) suggests that the conserved location of K within 2–3 amino acids of the FG sequence may be relevant for FG-mediated molecular recognition. To characterize this relationship, we tested whether placing a K immediately adjacent to the FG sequence (FGK) or moving it three amino acids away (FGA2K) affected the sequence's selectivity compared to the original arrangement in the FGAK peptide. Whereas FGA2K peptides readily formed a stiff material and could be tested for selective uptake, the FGK peptide gel was two orders-of-magnitude less stiff (Fig. 4.5, a and b), flowed when inverted, and dispersed when fluorescent reporters were loaded (data not shown). We therefore focused our comparison on diffusion and accumulation in FGAK and FGA2K gels. The FGAK gel selectively enriched for the hydrophobic (−) reporter but failed to uptake the hydrophobic (+) reporter to the same degree (Fig. 4.4, b and c). The effective diffusivity of the hydrophobic (−) (0.5 ± 0.14 μm^2/min) reporter in the FGA2K gel was an order-of-magnitude lower than that of the hydrophilic (−) reporter (7.9 ± 1.7 μm^2/min) (Fig. 4.5 c), showing that the combination of hydrophobic and electrostatic attraction synergizes for strong binding. However, the maximum accumulation at the interface of the two anionic reporters was similar (153. 6 ± 29.6 μM for hydrophobic (−) versus 205.7 ± 12.2 μM for hydrophilic (−)) (Fig. 4.5 c), indicating that the FGA2K gel is unable to differentiate

between the two reporters to the same degree as the original FGAK gel. For cationic

reporters, the hydrophobic (+) peptide accumulated within the gel (25.8 ± 0.8 μM) and

displayed an order-of-magnitude lower effective diffusivity (444.3 ± 267.7 μm^2/min) than
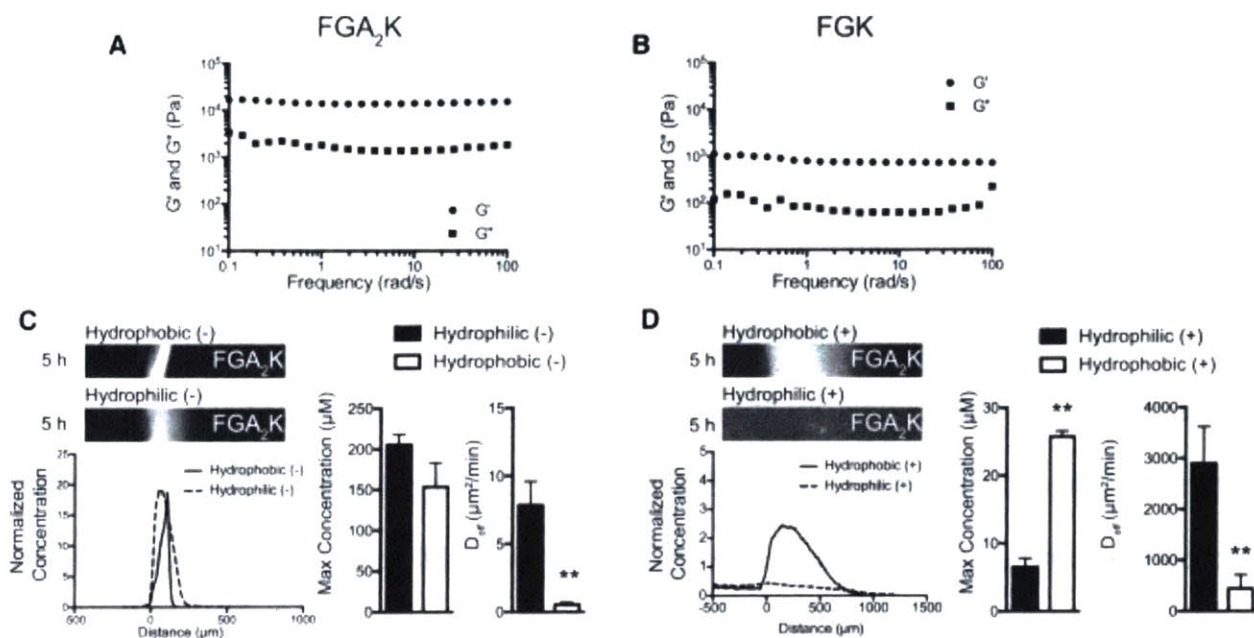
Figure 4.5 Effects of spatial localization of charge on FG-mediated self-assembly and selectivity. Frequency sweeps of (A) FGA2K and (B) FGK gels with G' (storage) and G" (loss) moduli are reported at 2% (w/v). (C) Shown here are representative fluorescence images of hydrophobic (−) and hydrophilic (−) reporters in a FGA2K gel, showing the similarities between the selectivity of the two reporters. (D) Depicted here are representative fluorescence images of hydrophobic (+) and hydrophilic (+) reporters, which contain a cationic tail, in a FGA2K gel showing the ability of FGA2K gels to select for hydrophobic (+) reporters independent of electrostatic repulsion. Error bars are SDs of at least three independent replicates. ∗∗p < 0.01, unpaired Student's t-test.

the hydrophilic (+) reporter (6.5 ± 1.3 μM and 2902 ± 719.8 μm^2/min, respectively) (Fig.

4.5 d). In addition, more hydrophobic (+) reporter accumulates in FGA2K gels than in

FGAK gels (Fig. 4.5 d). These data show that placing K within two amino acids of an FG

domain, approximately corresponding to the 1 nm Debye length at physiologically

relevant salt concentrations, enables hydrophobic selection of substrates with opposite

charge, whereas placing K three amino acids away reduces the contribution of

electrostatic interactions to hydrophobic FG-mediated selectivity. As a result, by increasing the distance between charged residues and FG sequences, the FG sequences become less dependent on the electrostatic profile surrounding the hydrophobic substrate during selective transport.

4.4.5 Glutamic acid reverses the selectivity of FG domains

Because the K residues in the FGAK gel help FG sequences differentiate between reporters containing anionic or cationic hydrophobic domains, we next asked whether FG-mediated recognition could be reversed by substituting anionic E residues. The selectivity of FGAE gels for the same class of fluorescent reporters was indeed reversed from that of FGAK gels (Fig. 4.6a–d). Selectivity was a function of both hydrophobic and electrostatic interactions, as the hydrophilic (+) reporters accumulated threefold less at the interface (41.8 ± 7.1 µM) than did the hydrophobic (+) reporters (123.2 ± 29.32 µM) (Fig. 4.6d). Moreover, the diffusion coefficient of the hydrophobic (+) reporter (6.6 ± 2.6 µm^2/min) was an order-of-magnitude lower than that of the hydrophilic (+) reporter (160.8 ± 32.7 µm^2/min) (Fig. 4.6d). Conversely, neither the hydrophobic (−) nor the hydrophilic (−) reporter interacted significantly with the gel; their similar effective diffusion coefficients (1381.0 ± 282.0 and 1451.0 ± 223.0 µm^2/min, respectively) (Fig. 4.6c) show that electrostatic repulsion minimizes hydrophobic interactions with FG sequences.

One general concern in synthesizing anionic peptides is residual TFA protonating anionic groups such as E. To ensure that the reporters were not neutralized in our buffering conditions, we synthesized neutral reporter peptides hydrophilic (n) and hydrophobic (n) (Fig. 4.S7). In diffusion experiments with FGAK and FGAE gels, the

122

neutrally charged reporters aggregated in solution and interacted minimally with the gels

(Fig. 4.S7), indicating that the anionic peptides are not neutralized by the residual TFA
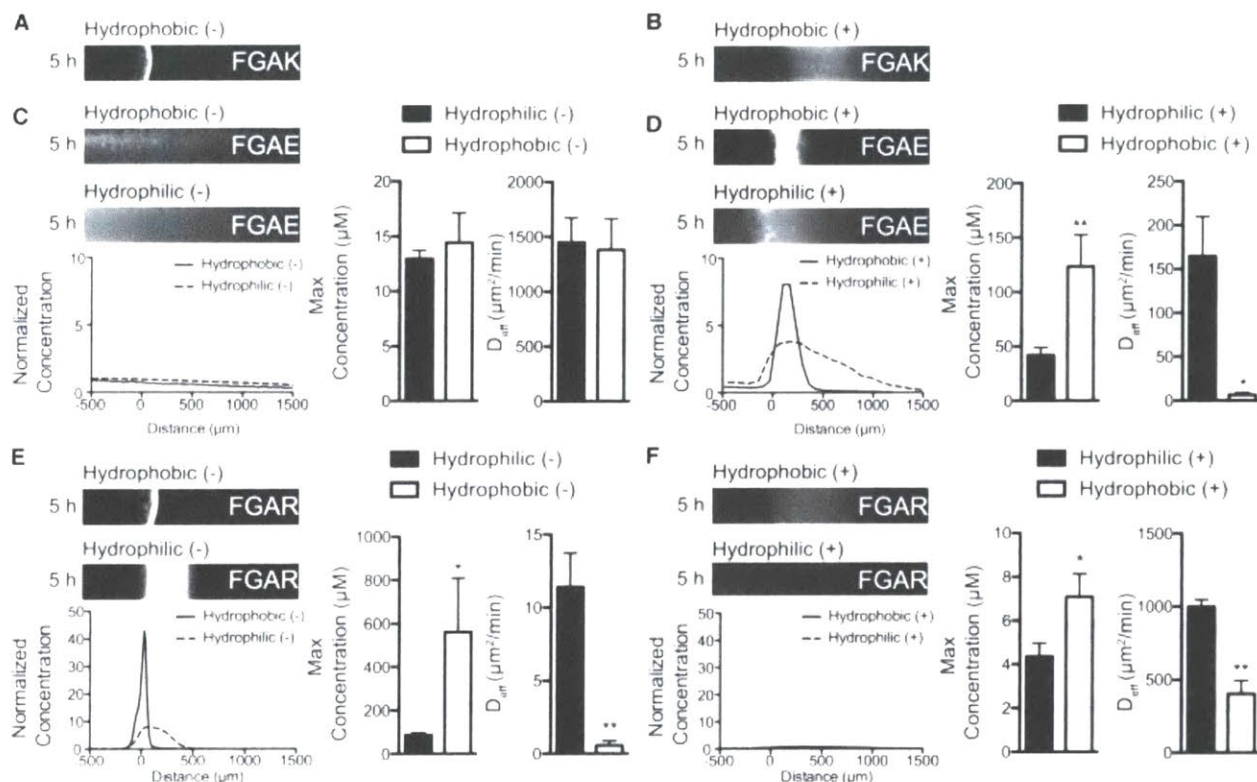
Figure 4.6 Substituting K with E reverses selectivity of FGAK gels, whereas replacement of K with R maintains selectivity similar to that of FGAK gels. Shown here are representative (A) hydrophobic (−) and (B) hydrophobic (+) fluorescence images in FGAK gels. (C) Shown here are representative fluorescence images of hydrophobic (−) and hydrophilic (−) reporters in an FGAE gel with the corresponding concentration profiles showing the reversal of selective recognition when compared to FGAK gels with the same reporters. In FGAE gels, anionic reporters do not interact at the gel interface. (D) Shown here are representative fluorescence images of hydrophobic (+) and hydrophilic (+) reporters in an FGAE gel, showing the reversal of selective recognition when compared to FGAK gels. In FGAE gels, the hydrophobic (+) reporter binds more significantly at the interface compared to the hydrophilic (+) reporter. (E) Depicted here are representative fluorescence images of hydrophobic (−) and hydrophilic (−) reporters in an FGAR gel, showing the similarity of selective recognition to that of FGAK gels. (F) Given here are representative fluorescence images of hydrophobic (+) and hydrophilic (+) reporters in an FGAR gel, showing the similarity of selective recognition to that of FGAK gels. Error bars are SDs of at least three independent replicates. $*p < 0.05$ and $**p < 0.01$, unpaired Student's t-test.

123

and indeed carry a net negative charge. Taken together, the data in Fig. 4.6 and Fig. 4.S7 suggest that charge proximal to FG sequences can tune hydrophobic selectivity, and that charge is essential in determining the hydrophobic moieties recognized by FG sequences.

4.4.6 Amino acid side-chain chemistry is an additional regulator of FG function

Last, we tested how the side-chain chemistry of charged amino acids affects the self-assembly and molecular recognition of FG sequences. In Nsp1, the predominant cation is K rather than R (Fig. 4.1a), suggesting that the two cations may not function equally. Similarly, E is highly conserved at position 9, whereas D is not as well conserved throughout the repeats (Fig. 4.1a). To test the importance of side-chain chemistry, we determined in the diffusion assay that FGAR gels have uptake properties similar to those of FGAK gels (Fig. 4.6a and e). The hydrophobic (−) reporter accumulated (561.1 ± 247.7 µM) and interacted with the interface of the FGAR gels (0.6 ± 0.3 µm^2/min), whereas the hydrophobic (+) reporter did not (7.1 ± 1.1 µM and 403.7 ± 90.2 µm^2/min, respectively) (Fig. 4.6e and a). These data suggest that from an electrostatic standpoint, R is just as capable as K in helping an FG sequence differentiate between hydrophobic substrates. However, mechanically, FGAR forms gels with an approximate stiffness of 2000 Pa throughout the frequency sweep (Fig. 4.S8a), fourfold more compliant than FGAK gels (Fig. 4.2a). The differences in the mechanical properties of the FGAK and FGAR gels suggest that the K and R residues may predominantly affect the structural self-assembly of FG domains. Transmission electron

microscopy revealed that FGAR peptides (Fig. 4.S8b) form different structures than FGAK peptides (Fig. 4.2a).

FGAD did not form a gel at the standard 2% (w/v), formed no repeating structures (Fig. 4.S8c and d), and inverted when flowed (data not shown), suggesting
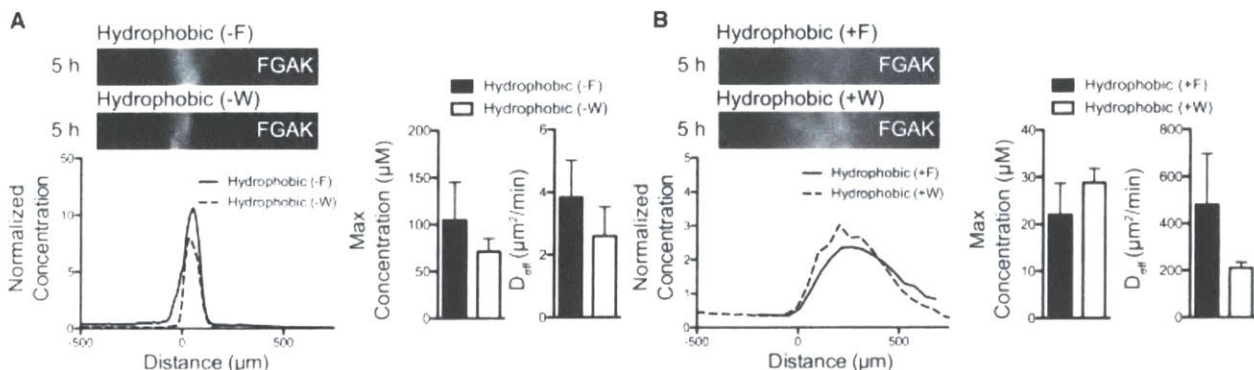
Figure 4.7 Electrostatic interactions modify hydrophobic recognition of W-containing reporters similar to F-containing reporters in FGAK gels. Comparisons between (A) hydrophobic (−F) and hydrophobic (−W) reporters and (B) hydrophobic (+F) and hydrophobic (+W) reporters in FGAK gels show that replacing F with W does not alter hydrophobic molecular recognition for FGAK gels. Representative fluorescence images and corresponding concentration profiles are shown. Error bars are SDs of at least three independent replicates. No significant differences were detected via Student's t-test.

that the small chemical differences between E and D transition the material from a selective gel to a viscous solution. Taken together, these data show that in addition to net charge, amino acid side-chain properties such as their inherent hydrophobicity, side-chain sterics, and charge distribution all may be important parameters that determine hydrophobic FG-mediated interactions.

4.4.7 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 natural aromatic amino acids other than F. In native NTRs such as NTF2 and other transport receptors, tryptophan (W) is commonly

used for the transient hydrophobic interactions required for transport. To test whether an electrostatic dependence persists when F is converted to W, we constructed the reporters hydrophobic (+W) and hydrophobic (−W) from the original hydrophobic F reporters (termed +F and −F, respectively, for the remainder of the article). The W-containing reporters were modulated by electrostatic interactions in a similar manner to F-containing reporters. The maximum accumulated fluorescence at the interface and diffusion did not significantly differ between the hydrophobic −F and −W reporters (Fig. 4.7a). The cationic reporters exhibited similar maximum interface properties, but the +W reporter showed a slight and nonsignificant trend toward slower diffusion than the +F reporter (+F: 479.9 ± 217.2 μm^2/min; +W: 210.5 ± 24.6 μm^2/min) (Fig. 4.7b), indicating that although both aromatic residues can be modulated by electrostatic interactions, the identity of the hydrophobic residue has a small effect on transport in the gel. These results reveal that in the NPC, the phenomenon of electrostatic interactions regulating hydrophobic interactions may be generalizable to aromatic residues other than F; moreover, selective transport may also be tuned by the identity of the nonpolar amino acid.

4.5 DISCUSSION

This investigation demonstrates how the environment surrounding FG sequences tunes the hydrophobic interactions necessary for correct molecular recognition of hydrophobic substrates. Here, we have explicitly shown how the presence, placement, and type of charged amino acid are all essential parameters for FG function in gels. Translating concepts that are well established in the peptide field enabled us to establish how even simplified variants of the Nsp1 consensus sequence can encode

126

complex molecular recognition. The data reported here support previous theoretical predictions that electrostatic interactions may be an important regulator of NPC selectivity [100] and sequence observations that the bias in the net charge of NTRs may be important in transport through the NPC [99]. Moreover, recent reports on how charge influences interactions with hydrophobic surfaces [96,197] emphasize that the interplay between electrostatics and hydrophobicity may be fundamental for various biological functions such as molecular recognition and DNA packing [198,199].

For the nuclear pore field, this investigation shifts the spotlight from how hydrophobic effects determine NPC selectivity to how electrostatics contribute to this selectivity [98,174,175,200]. The results described here provide insight into how individual hydrophobic FG sequences may distinguish a particular subset of hydrophobic substrates using a combination of electrostatic and hydrophobic domains that are in close proximity to each other (Fig. 4.8). The closer the charged residues are to FG sequences, the greater the influence of charge on FG self-assembly (Fig. 4.8a) and recognition of diffusing substrates (Fig. 4.8b). As the charged residue is situated greater than the theoretical Debye length at physiological conditions, the hydrophobic and electrostatic motifs are able to act independently for molecular recognition.

We recognize that although the study of engineered peptides can answer questions on how hydrophobic interactions are affected by the precise placement of charged amino acids, short-peptide gels do not recapitulate all of the unique facilitated diffusion properties of intact FG nucleoporins. This disconnect likely arises from structural differences, as short peptides form fibers and rods (Fig. 4.2 and Fig. 4.S8) that are not the dominant structures observed or predicted in the native nuclear pore

[172,201]. Moreover, the interactions between diffusing reporter peptides and the self-assembled gel may cause structural changes in the gel and also alter the diffusivity of the reporter, which has also been reported for Impβ in native NPCs [202]. Nevertheless, this investigation of peptides provides strong evidence (Figs. 4.4, 4.5, and 4.6) that
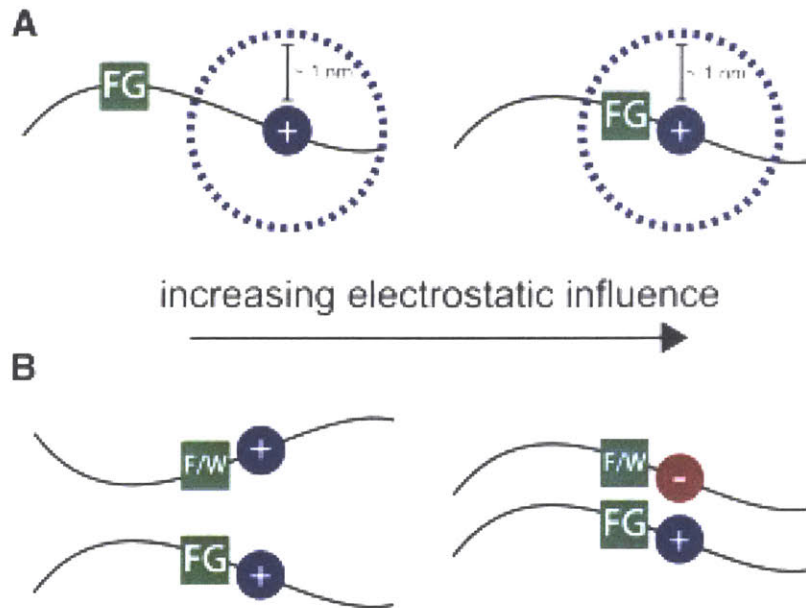
Figure 4.8 Schematic of the influence of electrostatics on hydrophobic interactions mediated by FG domains. (*A*) At physiologically relevant salt concentrations, electrostatic interactions typically have a range of ~1 nm (Debye length, *dotted lines*), which determines how much a charged residue influences the FG sequence's recognition of and binding to hydrophobic substrates. When charged residues are moved further away from the FG sequence, electrostatic interactions become less significant for FG-mediated self-assembly and selective transport. (*B*) Complementary charge and hydrophobic interactions enable the strongest interactions between a substrate and FG domains. Modulating charge types and the presence of hydrophobic domains can tune interactions from strong binding to free diffusion within the gel and in in vivo systems.

charge is an essential regulator of transport through the NPC by tuning the essential first step of recognizing specific hydrophobic substrates. Moreover, engineered FG-containing peptides do capture the interplay between hydrophobic and electrostatic

interactions at single amino acid resolution (Fig. 4.5), which is not possible with currently available in vivo techniques.

Our results build on previous studies of intact nucleoporins containing multiple repeats of FG domains where charge primarily plays a structural role in the cohesion of the self-assembled matrix [201]. Our results also support recent work in which Impβ was recognized to have varying binding capacities with synthetic FG-containing polyacrylamide gels depending on the charged state of the material[203]. We believe that our strategy, which was inspired and informed by peptide engineering, complements other well-established methods used to understand the selective transport mechanism of NPCs, such as the in vivo minimal NPC [181,182], gel and selectivity analysis of individual nucleoporins [98,173,174,204], and binding interactions with surface-grafted nucleoporin films [201,205–207]. We expect that this rational approach from the peptide field, which focuses on conserved repeating sequences, can be extended to other biological gel systems such as mucus, byssal threads, and cartilage, where complex disordered proteins based on repeat units and reversible cross linking constitute a significant proportion of the material and dictate its function [208–210].

4.6 SUPPLEMENTAL INFORMATION

4.6.1 Supplementary Data S1

Purity of synthesized peptides assessed by analytical HPLC and mass spectrometry

FGAK Peptide

Analytical HPLC BgcD[USf[a`
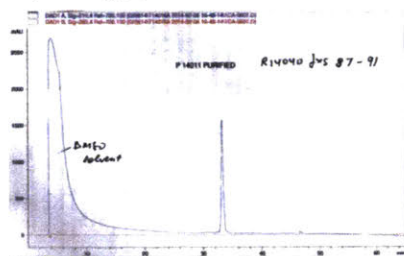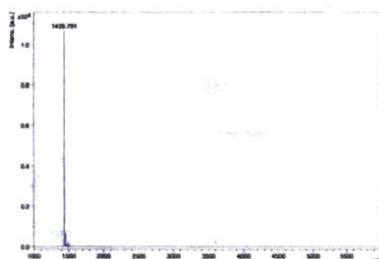
R14040 JvS 87-91

Mass spectrometry analysis

Laser repetition rate in Hz: 60 Hz
Linear detector voltage: 2.691 kV
Reflector detector voltage: 1.59 kV
Ion source voltage 1: 20 kV
Ion source voltage 2: 18.55 kV
Ion source lens voltage: 6.5 kV
Number of shots: 400

BOL
Boston Open Labs

Analysis Report

Inj. Date: 11/16/2015   12:58:58
Product Name: P155934
Mobile Phase: A: 0.05% TFA in H2O
                B: 0.05% TFA in 100% ACN
Grads: 21%-31%B in 20 min   Flow : 1.0 ml/min
Column :   Agilent ZORBAX 300SB-C18 5um 4.6*150mm 220nm

Operator: Aiqin Wang
Lot: QP102815KZ1E
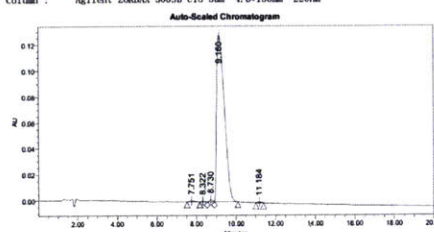
Peak Results

	Retention Time (min)	Int Type	Area (µV*sec)	Height (µV)	Width (sec)	% Area
1	7.751	BB	12262	590	39.000	0.34
2	8.322	BV	6094	572	22.000	0.17
3	8.730	VV	18379	1727	22.000	0.52
4	9.160	VB	3515511	128499	72.000	98.78
5	11.184	BB	6614	609	23.000	0.19

BOL
Boston Open Labs

Product Name: P155934
Lot: QP102815KZ1E
MW: 1435.64
Date: 2015-11-16

130
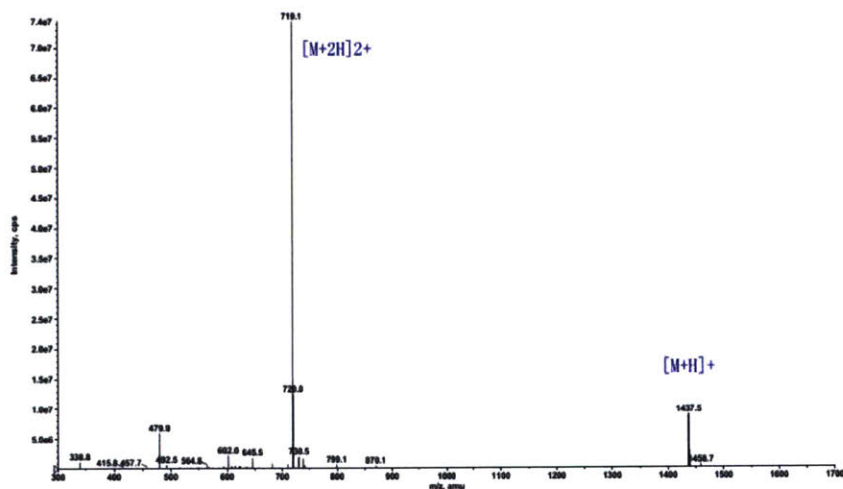
Analysis Report

Inj. Date: 11/16/2015   11:28:06   Operator: Aiqin Wang
Product Name: P155935   Lot: QP102815KZ1F
Mobile Phase: A: 0.05% TFA in H2O
              B: 0.05% TFA in 100% ACN
Grads: 23%-33%B in 20 min   Flow : 1.0 ml/min
Column :   Agilent ZORBAX 300SB-C18 5um 4.6*150mm 220nm
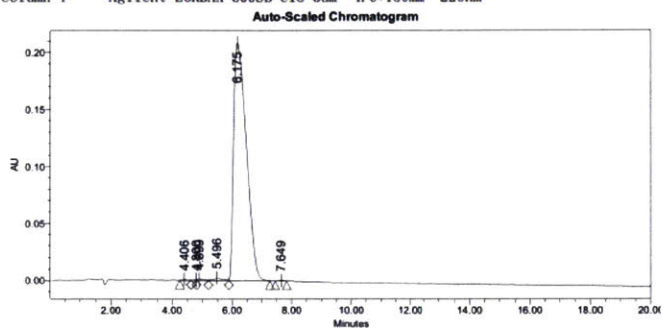
Peak Results

	Retention Time (min)	Int Type	Area (μV*sec)	Height (μV)	Width (sec)	% Area
1	4.406	BV	5127	574	22.000	0.08
2	4.800	VV	2555	496	11.000	0.04
3	4.899	VV	15792	1193	24.000	0.25
4	5.496	Vv	45962	2333	41.000	0.73
5	6.175	vB	6256502	208887	83.000	98.83
6	7.649	BB	4769	436	23.000	0.08

Product Name: P155935
Lot: QP102815KZ1F
MW: 1435.64
Date:2015-11-16

131
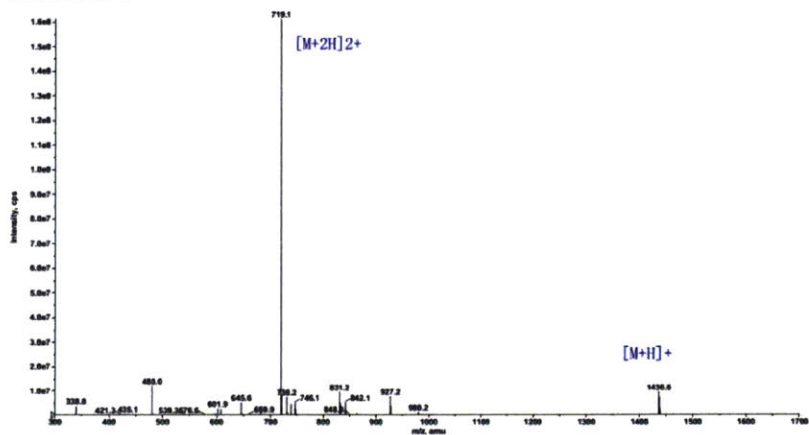
FGAR Peptide

Analytical HPLC
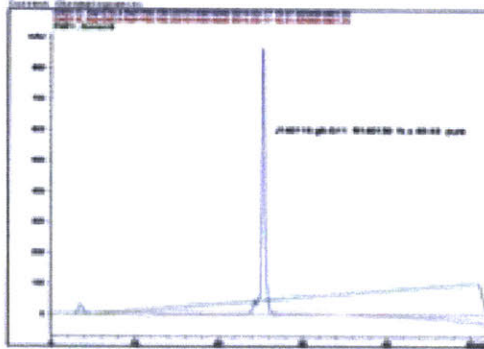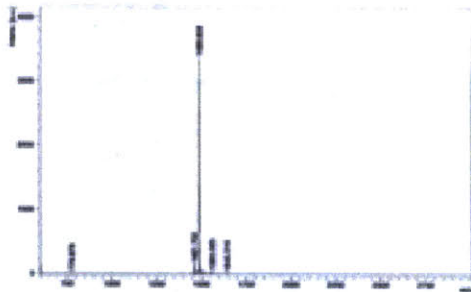
Mass spectrometry analysis

Laser repetition rate in Hz: 60 Hz
Linear detector voltage: 2.691 kV
Reflector detector voltage: 1.59 kV
Ion source voltage 1: 20 kV
Ion source voltage 2: 18.55 kV
Ion source lens voltage: 6.5 kV
Number of shots: 400

FGAE Peptide

Analytical HPLC

Mass spectrometry analysis

Laser repetition rate in Hz: 60 Hz
Linear detector voltage: 2.691 kV
Reflector detector voltage: 1.59 kV
Ion source voltage 1: 20 kV
Ion source voltage 2: 18.55 kV
Ion source lens voltage: 6.5 kV
Number of shots: 400

132

FGAS Peptide

Analytical HPLC
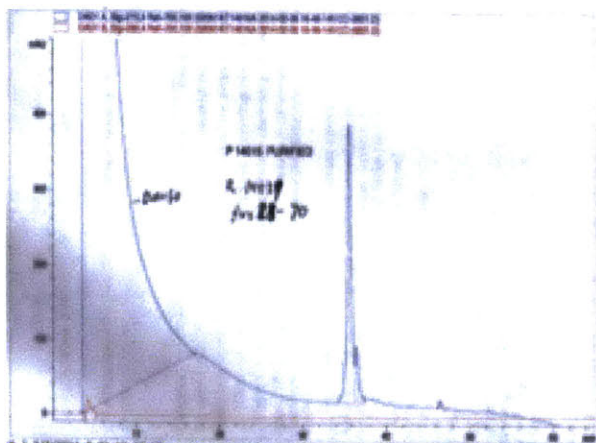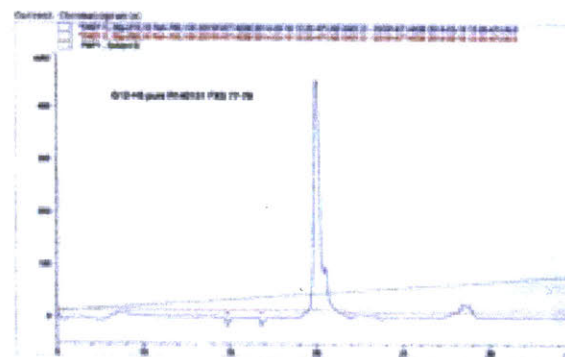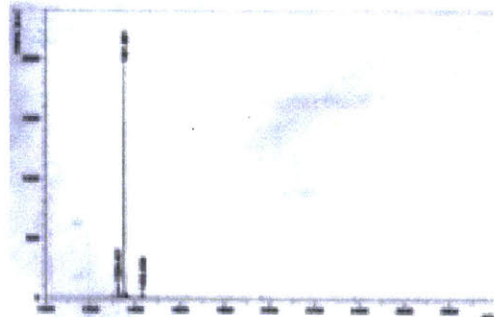
Mass spectrometry analysis

Laser repetition rate in Hz: 60 Hz
Linear detector voltage: 2.691 kV
Reflector detector voltage: 1.59 kV
Ion source voltage 1: 20 kV
Ion source voltage 2: 18.55 kV
Ion source lens voltage: 6.5 kV
Number of shots: 400

FGAD peptide

Analytical HPLC

Mass spectrometry analysis
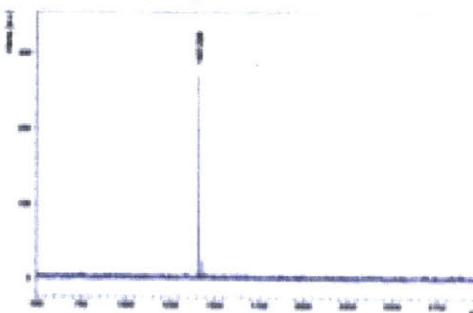
Laser repetition rate in Hz: 60 Hz
Linear detector voltage: 2.691 kV
Reflector detector voltage: 1.59 kV
Ion source voltage 1: 20 kV
Ion source voltage 2: 18.55 kV
Ion source lens voltage: 6.5 kV
Number of shots: 400

Nsp1 Consensus Peptide

Analytical HPLC
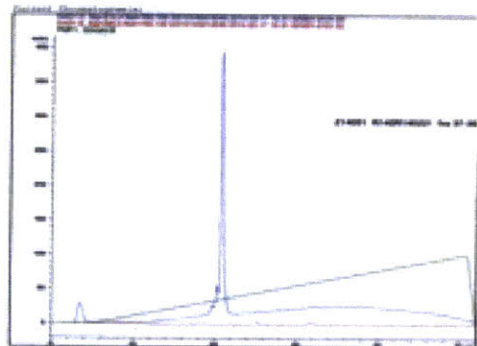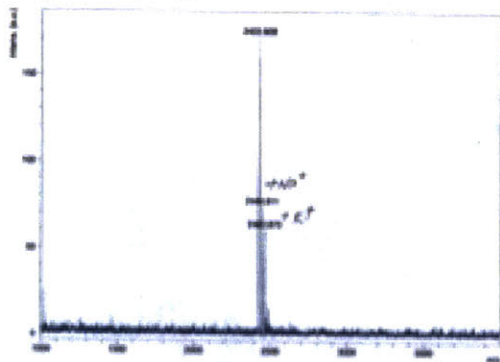
Mass spectrometry analysis

Laser repetition rate in Hz: 60 Hz
Linear detector voltage: 2.691 kV
Reflector detector voltage: 1.59 kV
Ion source voltage 1: 20 kV
Ion source voltage 2: 18.55 kV
Ion source lens voltage: 6.5 kV
Number of shots: 400

SGAK Peptide

Analytical HPLC
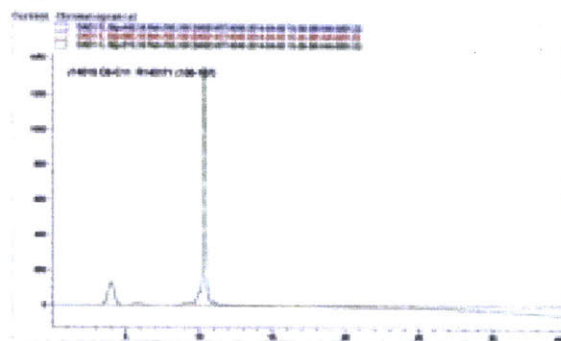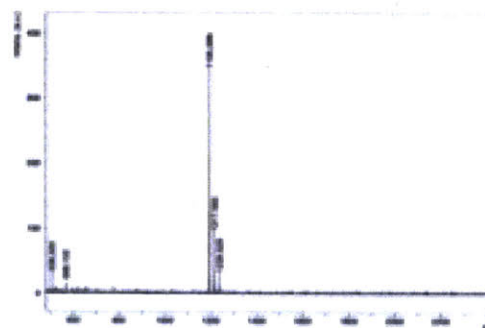
Mass spectrometry analysis

Laser repetition rate in Hz: 60 Hz
Linear detector voltage: 2.691 kV
Reflector detector voltage: 1.59 kV
Ion source voltage 1: 20 kV
Ion source voltage 2: 18.55 kV
Ion source lens voltage: 6.5 kV
Number of shots: 400

134

Hydrophilic (+) Peptide with N-terminal FAM labeling

Analytical HPLC

Mass spectrometry analysis

Laser repetition rate in Hz: 60 Hz
Linear detector voltage: 2.691 kV
Reflector detector voltage: 1.59 kV
Ion source voltage 1: 20 kV
Ion source voltage 2: 18.55 kV
Ion source lens voltage: 6.5 kV
Number of shots: 400
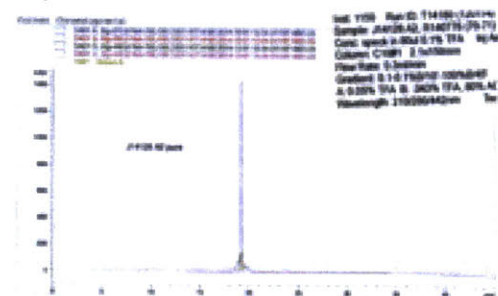
Hydrophilic (-) peptide reporter with N-terminal FAM label

Analytical HPLC
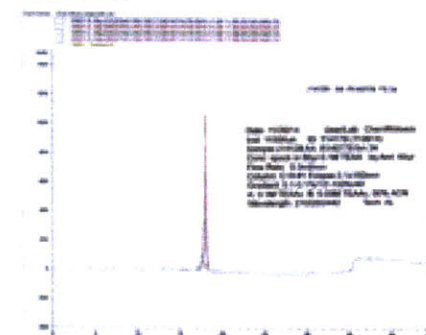

Mass spectrometry analysis
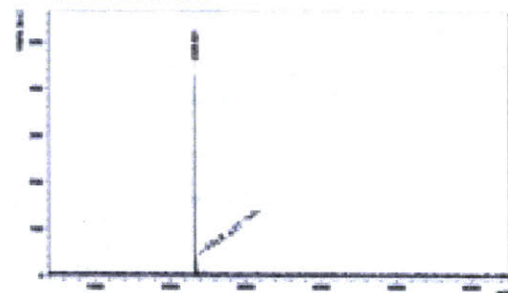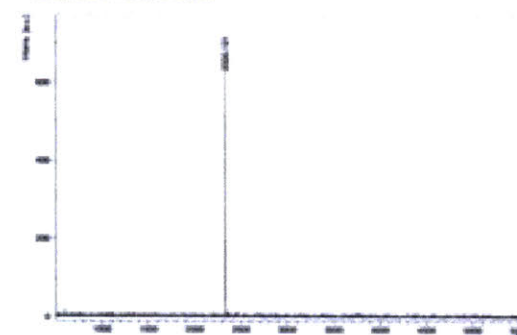
Laser repetition rate in Hz: 60 Hz
Linear detector voltage: 2.691 kV
Reflector detector voltage: 1.59 kV
Ion source voltage 1: 20 kV
Ion source voltage 2: 18.55 kV
Ion source lens voltage: 6.5 kV
Number of shots: 400

Hydrophobic (+) Reporter Peptide with N-terminal FAM

Analytical HPLC

J14128 A1 pure

Sample: J14128 A1, R140775 (53
Conc: speck in 60ul 0.1% TFA
Column: C18#1 2.1x150mm
Flow Rate: 0.3ml/min
Gradient: 0.1-0.1%B/10'-100%B/4
A: 0.05% TFA B: .043% TFA, 60%
Wavelength: 210/280/442nm
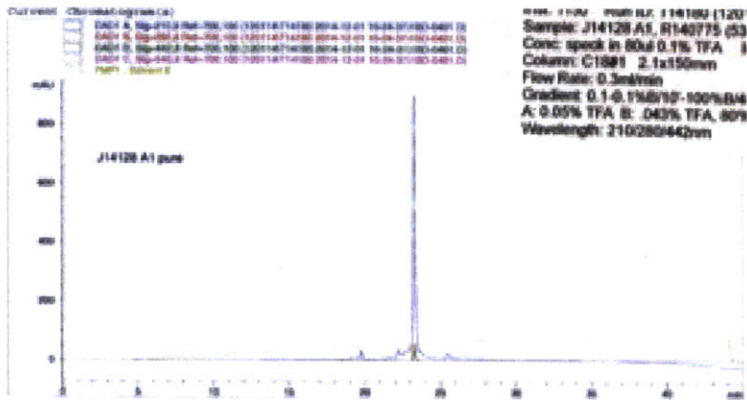
Mass spectrometry analysis
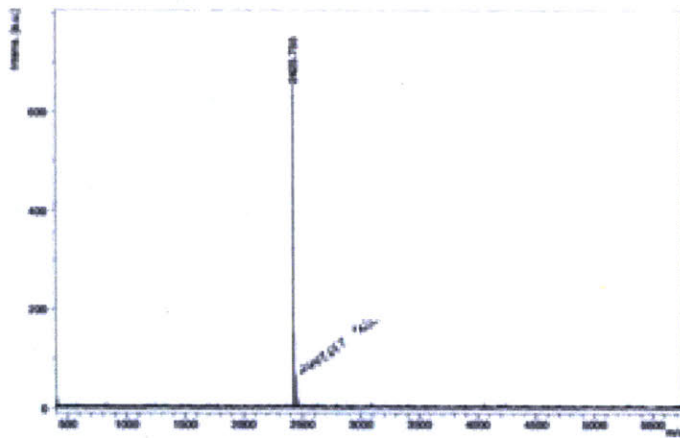
Hydrophobic (−) Reporter Peptides with N-terminal FAM

Analytical HPLC

Mass spectrometry analysis
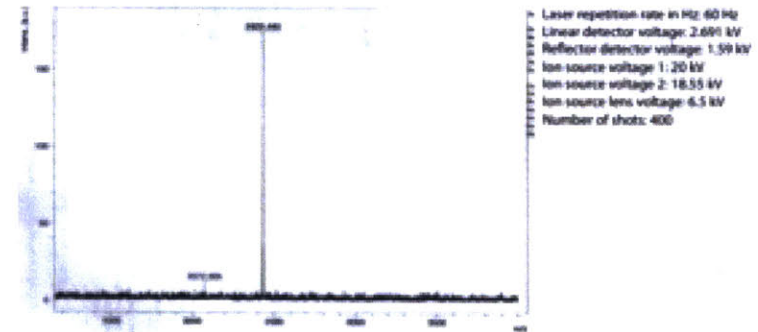
Laser repetition rate in Hz: 60 Hz
Linear detector voltage: 2.69/1 kV
Reflector detector voltage: 1.59 kV
Ion source voltage 1: 20 kV
Ion source voltage 2: 18.55 kV
Ion source lens voltage: 6.5 kV
Number of shots: 400

Hydrophobic (+W) peptide with N-terminal FAM labeling

Analytical HPLC

Mass spectrometry analysis

Laser repetition rate in Hz: 60 Hz
Linear detector voltage: 2.691 kV
Reflector detector voltage: 1.59 kV
Ion source voltage 1: 20 kV
Ion source voltage 2: 18.55 kV
Ion source lens voltage: 6.5 kV
Number of shots: 400

Hydrophobic (-W) peptide with N-terminal FAM labeling

HPLC Purification. Fraction 35 was collected for mass spectrometry and diffusion analysis.
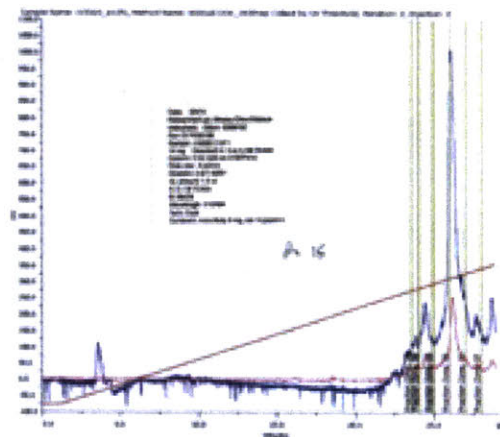
Mass spectrometry analysis
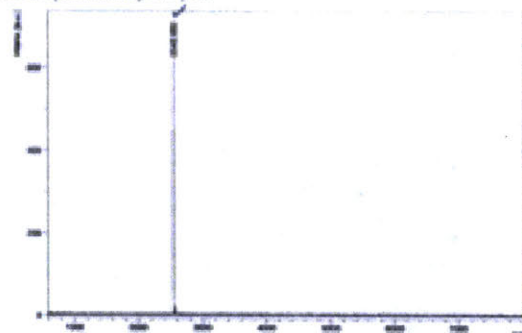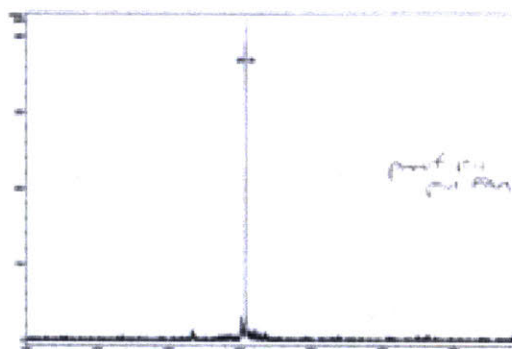
Laser repetition rate in Hz: 60 Hz
Linear detector voltage: 2.691 kV
Reflector detector voltage: 1.59 kV
Ion source voltage 1: 20 kV
Ion source voltage 2: 18.55 kV
Ion source lens voltage: 6.5 kV
Number of shots: 400
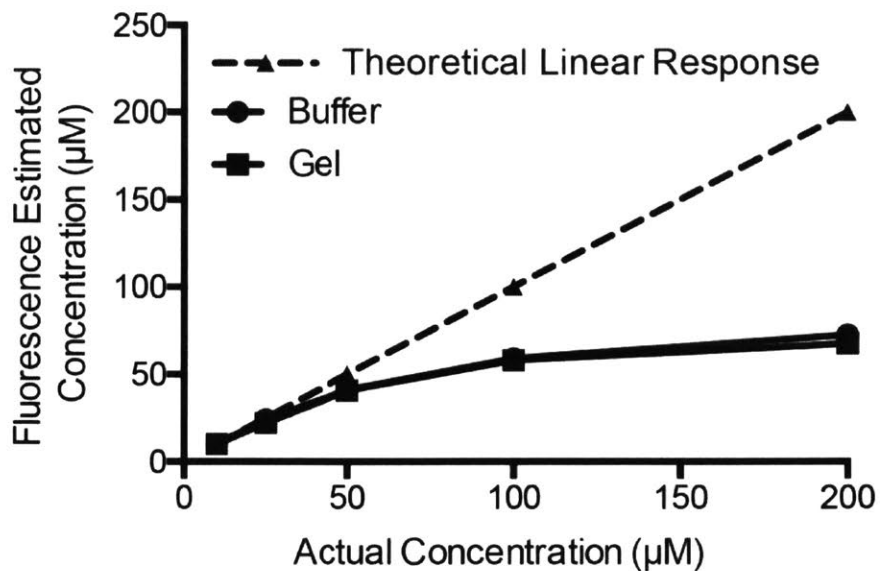
Figure 4.S2 Quantification of fluorescence signal as a function of fluorophore concentration.Fluorescence signal is approximately linear up to 50 µM and saturates by 100 µM in buffer and gel conditions. Dashed lines represent the theoretical linear response of fluorescence as a function of concentration. The gels and buffer calibration curves overlap in their associations. All concentrations are reported according to the experimental curve developed and represent lower values of the actual concentrations for values >100 µM.
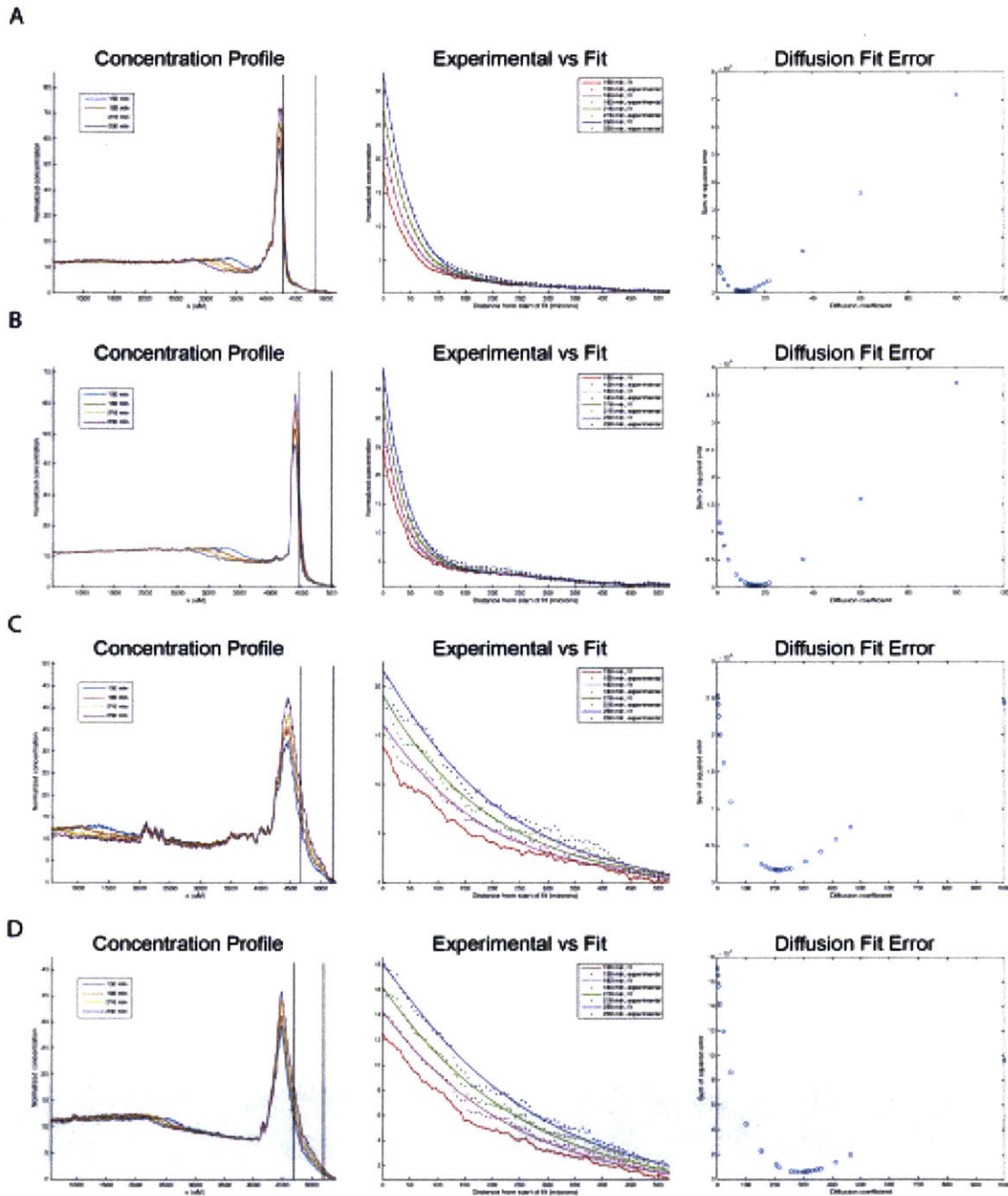
Figure 4.S3 Analytical process for calculating effective diffusion coefficients. Examples are given for A) NTF2 diffusion into FGAK, B) W7A diffusion into FGAK, C) NTF2 diffusion into FGAE, and D) W7A diffusion into FGAE to show the reliability of the analytical process across multiple gels. The first column represents the region of the concentration profile where the fitting is implemented. The second column contains the actual data (circles) vs. fit (solid line) at four evenly spaced time points. The third column contains the error of the fit as a function of iterated effective diffusion coefficients.

139

A

SSSGAKAAKASSSG
(no gel)

B
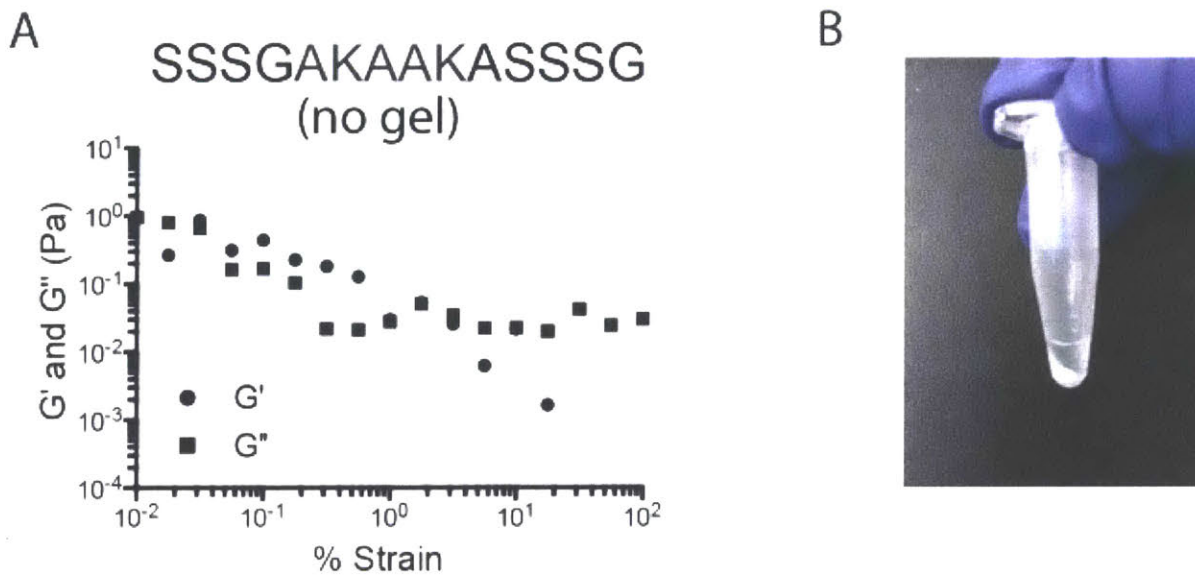
Figure 4.S4 Verification of F as essential amino acid for self-assembly in FGAK peptides. A) Frequency sweep of the F!S substitution (FGAK ! SGAK) to determine the effect of F on the self-assembly of peptides. The elastic modulus (G') and loss modulus (G") are reported. Note that the measured values are below the sensitivity of the rheometer using the specified cone-plate geometry due to the viscous nature of SGAK peptide solutions. B) Precipitated FGAS peptides in 20 mM NaCl, 20 mM HEPES [pH 7] after gentle centrifugation.

Figure 4.S5 Verification of hydrophobic domain availability in FG peptide gels. Transport of Nile Red into FGAK and FGAE gels after 0 h and 3 h of incubation. Fluorescence indicates that the dye is able to access hydrophobic environments created by FG domains within the gels. Images are of representative gels from three independent replicates.

Normalized Anionic Reporters

Normalized Cationic Reporters
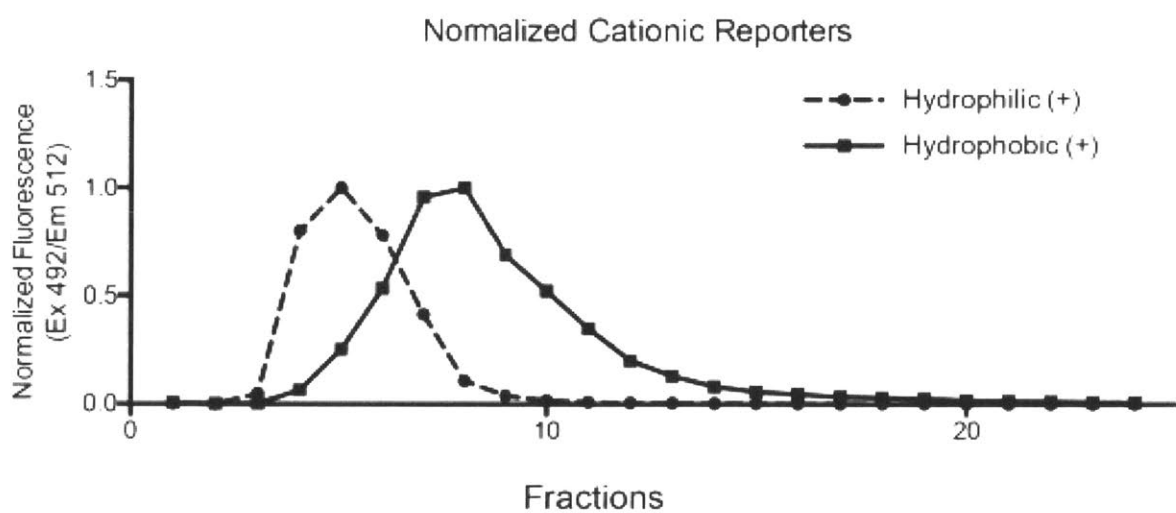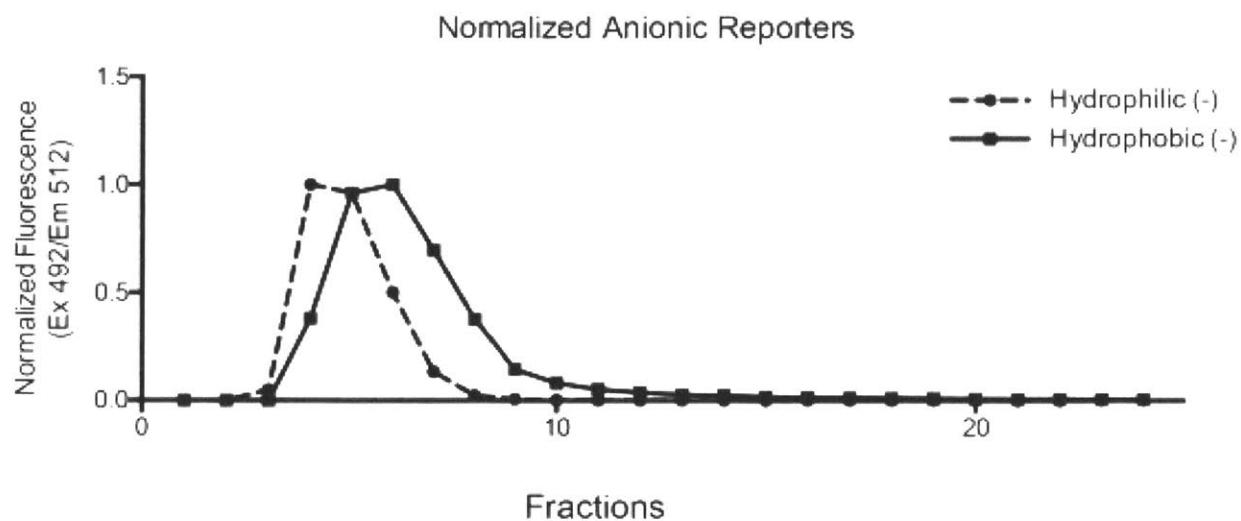
Figure 4.S6 Fractionation of hydrophilic reporters and their hydrophobic counterparts in phenyl-sepharose columns. Fluorescence signals from each fraction were collected and normalized to the signal with the highest intensity of emission. For both cationic and anionic reporters, the hydrophobic reporters eluted later. This increased retention time reflects stronger binding to phenyl-sepharose beads.
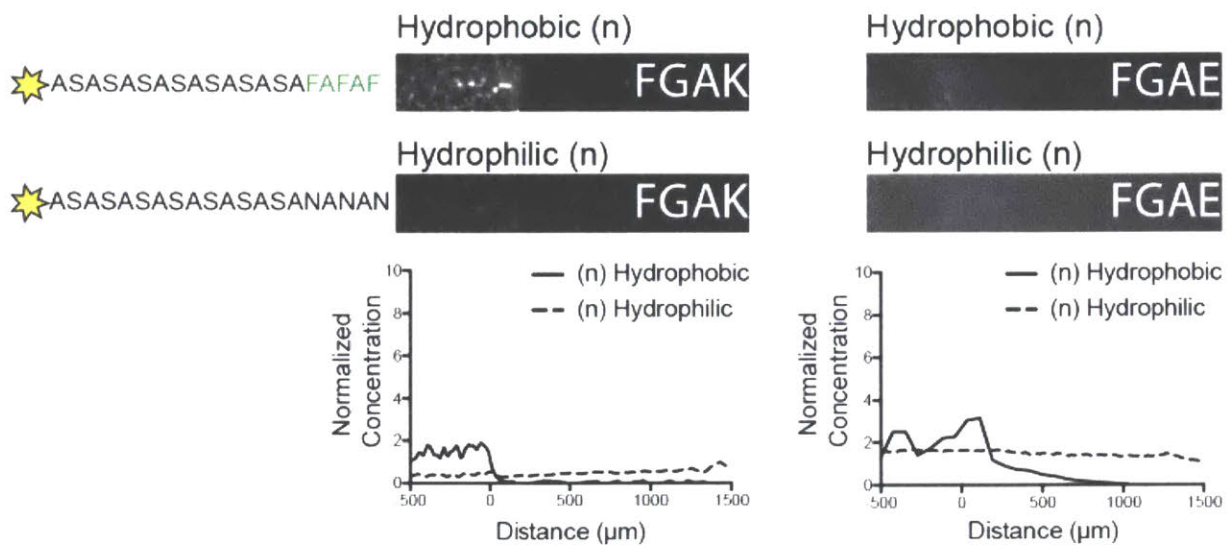
Figure 4.S7 Diffusion of neutrally charged Hydrophilic (n) and Hydrophobic (n) reporters into cationic FGAK and anionic FGAE gels. Purely neutral reporters interact minimally with the FGAK and FGAE gels regardless of overall hydrophobicity.

A

FGAR
(gel)

B

FGAR

C

FGAD
(no gel)

D

FGAD
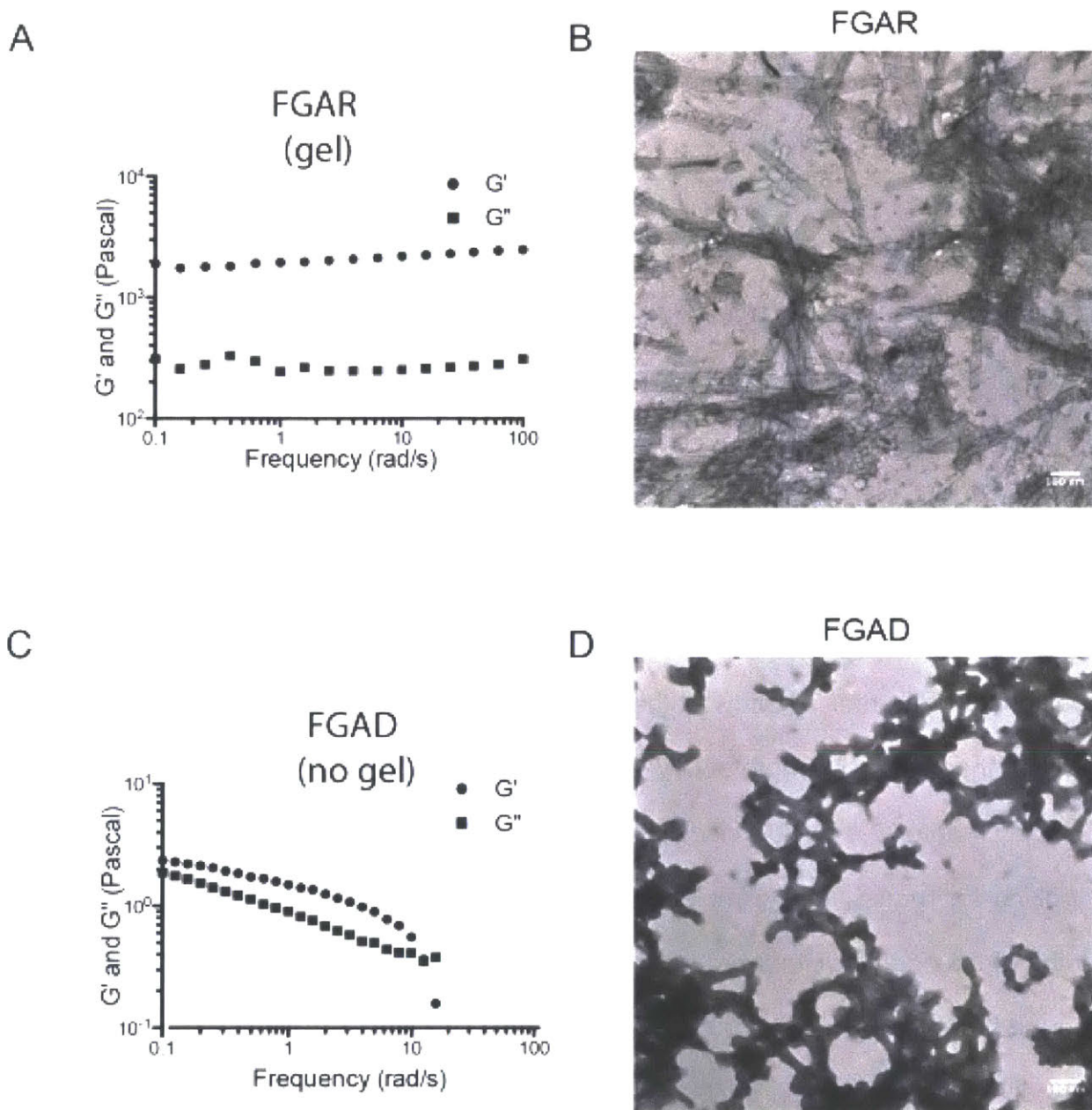
Figure 4.S8 Affect of amino acid sidechain chemistry on self-assembly and mechanical properties of FG-containing peptides. A) Frequency sweep of FGAR gel with G' (storage) and G" (loss) moduli reported at 2% (w/v) showing the stable self-assembled matrix is maintained when converting from K to R. B) Corresponding image from transmission electron microscopy showing the structural variation of FGAR peptide self-assembly when compared to that of FGAK peptides. C) Frequency sweep of FGAD peptide solution with G' (storage) and G" (loss) moduli reported at 2% (w/v) showing that FGAD does not form a gel. D) Corresponding image from transmission electron microscopy showing the amorphous structure of FGAD peptide aggregates.

143

5 CONCLUSIONS

In this thesis, I took multiple approaches to analyze questions related to the selective permeability and binding properties of mucus, mucins, and the nuclear pore. In Chapter 2, I focused on small molecule binding to MUC5AC, DNA, and alginate. MUC5AC is a constitutive component of gastric and respiratory mucus, present at particularly high concentrations in muco-obstructive diseases, particularly asthma,[211] while DNA and alginate are found in CF respiratory mucus. I used an SMM to identify DAP as a MUC5AC binding motif and then combined equilibrium dialysis with qLCMS to confirm that DAP is associated with mucin binding, at least under relatively low salt conditions. Furthermore, I showed that under these conditions increased hydrophobicity amplified the binding strength of DAP-containing molecules.

In Chapter 3, I applied my ED technique to the MUC5B rich HBE mucus. Instead of looking for mucin binding molecules without any particular function, I examined the binding of the *P. aeruginosa*-produced toxin pyocyanin and numerous drugs to mucus. I confirmed that the highly cationic polymyxin antibiotics bind to HBE mucus (and show in Appendix A that polymyxin antibiotics bind to MUC2 and a different source of MUC5B as well), while most other antibiotics did not show any noticeable binding. Two exceptions are aztreonam, for reasons that are unclear, and iclaprim. Iclaprim's binding to HBE mucus is particularly striking because it contains the DAP more motif and is more hydrophobic than its analogue trimethoprim (which did not bind mucus). Further investigation into iclaprim's interactions with mucus does appear warranted. I next considered inhaled non-antibiotic drugs, where I did not observed particularly strong

binding though there are hints of interesting differences between different bronchodilators.

I then modeled the pharmacokinetics of the lung environment in order to predict how potentially significant diffusion may be for changing the duration of action of an inhaled drug. I concluded that the diffusion, and by extension diffusion slowed by mucus binding, may be important where large, immobile mucus plugs are present.

The clear next steps for this project center around attempting to perform rational design on drugs to optimize mucus binding properties. This concept has been previously attempted, to my knowledge, for two different potential therapies and in both cases the goal was to reduce mucus binding because dosage frequency was not likely limiting. First, tobramycin is a cationic antibiotic used as an inhaled anti-*P. aeruginosa* therapy in CF, but it is inhibited by CF lung mucus.[75] The Smyth group has made a PEGylated form of tobramycin, which reduces tobramycin's activity where no polyanions are present but also reduces inhibition by polyanions. The resulting net effect appears to be slightly improved activity against biofilms and in the presence of mucus.[21,212] A similar effort has taken place for lysozyme, a cationic antimicrobial protein which is inhibited by mucus. Mutagenesis of lysozyme resulted in a protein with reduced inhibition by polyanions and promising *in vivo* activity against *P. aeruginosa* lung infections.[213–217]

Following these proofs of concept, drug modification to reduce mucus binding may also be useful for polymyxin antibiotics and potentially even iclaprim. For polymyxin antibiotics, reducing the positive charge would undoubtedly reduce both activity and polyanion binding. However, it is possible that other modifications would reduce binding

145

without reducing activity. Our results from Chapter 2 suggest that making iclaprim more hydrophilic would reduce its mucus binding. Unfortunately, hydrophobic interactions between iclaprim and its target, dihydrofolate reductase (DHFR), are critical to iclaprim's activity.[218] However, it may be possible to make it more hydrophilic in a location that would not interfere with iclaprim's interactions with DHFR.

We also discussed in Chapter 3 how increasing lung mucus binding might be desirable if four conditions are met: a problematically short lung residence time but still long enough to be substantial, an intracellular drug target, a muco-obstructive disease, and a mechanism that does not result in immediate mucociliary clearance of the mucus that is binding the drug. Furthermore, while we have primarily discussed respiratory mucus in this dissertation, similar considerations regarding mucus binding and drug activity, and diffusion and dosing frequency, apply to other mucus layers such as intestinal and cervical mucus.[136] Applying our techniques to study these niches may show similar promise and challenges and is an interesting prospect for future work.

In Chapter 4, I examined a different biopolymer system, the nuclear pore. Transport through the nuclear pore is a biochemically complex problem but appears centered around interactions with nucleoporins, intrinsically disordered proteins with hydrophobic FG domains that act as binding sites for nuclear transport factors and potentially as crosslinkers. Previously, the contribution of electrostatics to interactions with FG motifs has been underappreciated and we designed a peptide-based hydrogel system to dissect the possible contribution of charge interactions. Specifically, we showed that electrostatic and hydrophobic interactions can be cooperative or antagonistic where the charged and hydrophobic groups are close, which can affect

gelation and also diffusion of charged probes. I helped show this by modeling diffusion profiles in order to fit diffusion coefficients, and using these diffusion coefficients to infer transport properties and binding strengths. As we discuss in Chapter 4, future work could include extension of the peptide approach to other biological hydrogels such as mucus and cartilage. Alternatively, increasing the length of the synthetic peptides to match the length of the Nsp1 consensus repeat may yield further insights. More ambitiously, it would be a technical challenge but very scientifically rewarding to genetically encode, express, and purify what could be called artificial nucleoporins: proteins consisting of many simplified FG repeats. Systematically changing the repeat sequence may build a fuller understanding of how peptide properties scale up to proteins.

References

1. Witten, J. & Ribbeck, K. The particle in the spider's web: transport through biological hydrogels. *Nanoscale* **9**, 8080–8095 (2017).
2. Orell, A., Fröls, S. & Albers, S.-V. Archaeal biofilms: the great unexplored. *Annu. Rev. Microbiol.* **67**, 337–354 (2013).
3. Flemming, H.-C. & Wingender, J. The biofilm matrix. *Nat. Rev. Microbiol.* **8**, 623–633 (2010).
4. Flemming, H.-C., Neu, T. R. & Wozniak, D. J. The EPS matrix: the "house of biofilm cells". *J. Bacteriol.* **189**, 7945–7947 (2007).
5. Strambio-De-Castillia, C., Niepel, M. & Rout, M. P. The nuclear pore complex: bridging nuclear transport and gene regulation. *Nat. Rev. Mol. Cell Biol.* **11**, 490–501 (2010).
6. Schmidt, H. B. & Görlich, D. Transport Selectivity of Nuclear Pores, Phase Separation, and Membraneless Organelles. *Trends Biochem. Sci.* **41**, 46–61 (2016).
7. Johansson, M. E. V., Sjövall, H. & Hansson, G. C. The gastrointestinal mucus system in health and disease. *Nat. Rev. Gastroenterol. Hepatol.* **10**, 352–361 (2013).
8. Thornton, D. J., Rousseau, K. & McGuckin, M. A. Structure and function of the polymeric mucins in airways mucus. *Annu. Rev. Physiol.* **70**, 459–486 (2008).
9. Knowles, M. R. & Boucher, R. C. Mucus clearance as a primary innate defense mechanism for mammalian airways. *J. Clin. Invest.* **109**, 571–577 (2002).
10. Netsomboon, K. & Bernkop-Schnürch, A. Mucoadhesive vs. mucopenetrating particulate drug delivery. *Eur. J. Pharm. Biopharm. Off. J. Arbeitsgemeinschaft Pharm. Verfahrenstechnik EV* **98**, 76–89 (2016).
11. Lai, S. K., Wang, Y.-Y. & Hanes, J. Mucus-penetrating nanoparticles for drug and gene delivery to mucosal tissues. *Adv. Drug Deliv. Rev.* **61**, 158–171 (2009).
12. Critchfield, A. S. *et al.* Cervical mucus properties stratify risk for preterm birth. *PloS One* **8**, e69528 (2013).
13. Lewis, K. Persister cells, dormancy and infectious disease. *Nat. Rev. Microbiol.* **5**, 48–56 (2007).
14. Schuster, B. S., Suk, J. S., Woodworth, G. F. & Hanes, J. Nanoparticle diffusion in respiratory mucus from humans without lung disease. *Biomaterials* **34**, 3439–3446 (2013).
15. Abdulkarim, M. *et al.* Nanoparticle diffusion within intestinal mucus: Three-dimensional response analysis dissecting the impact of particle surface charge, size and heterogeneity across polyelectrolyte, pegylated and viral particles. *Eur. J. Pharm. Biopharm.* 1–9 (2015).
16. Duncan, G. A., Jung, J., Hanes, J. & Suk, J. S. The Mucus Barrier to Inhaled Gene Therapy. *Mol. Ther.* **24**, 2043–2053 (2016).
17. Witten, J., Samad, T. & Ribbeck, K. Selective permeability of mucus barriers. *Curr. Opin. Biotechnol.* **52**, 124–133 (2018).
18. Maisel, K., Ensign, L., Reddy, M., Cone, R. & Hanes, J. Effect of surface chemistry on nanoparticle interaction with gastrointestinal mucus and distribution in the

gastrointestinal tract following oral and rectal administration in the mouse. *J. Controlled Release* **197**, 48–57 (2015).

19. Ensign, L. M. *et al.* Mucus-penetrating nanoparticles for vaginal drug delivery protect against herpes simplex virus. *Sci. Transl. Med.* **4**, 138ra79–138ra79 (2012).

20. Sigurdsson, H. H., Kirch, J. & Lehr, C.-M. Mucus as a barrier to lipophilic drugs. *Int. J. Pharm.* **453**, 56–64 (2013).

21. Bahamondez-Canas, T. F., Zhang, H., Tewes, F., Leal, J. & Smyth, H. D. C. PEGylation of Tobramycin Improves Mucus Penetration and Antimicrobial Activity against Pseudomonas aeruginosa Biofilms in Vitro. *Mol. Pharm.* (2018). doi:10.1021/acs.molpharmaceut.8b00011

22. Stewart, P. S. A review of experimental measurements of effective diffusive permeabilities and effective diffusion coefficients in biofilms. *Biotechnol. Bioeng.* **59**, 261–272 (1998).

23. Stewart, P. S. Diffusion in biofilms. *J. Bacteriol.* **185**, 1485–1491 (2003).

24. Khanvilkar, K., Donovan, M. D. & Flanagan, D. R. Drug transfer through mucus. *Adv. Drug Deliv. Rev.* **48**, 173–193 (2001).

25. Darch, S. E. *et al.* Spatial determinants of quorum signaling in a Pseudomonas aeruginosa infection model. *Proc. Natl. Acad. Sci.* **115**, 4779–4784 (2018).

26. Li, L., Lieleg, O., Jang, S., Ribbeck, K. & Han, J. A microfluidic in vitro system for the quantitative study of the stomach mucus barrier function. *Lab. Chip* **12**, 4071–4079 (2012).

27. Eibauer, M. *et al.* Structure and gating of the nuclear pore complex. *Nat. Commun.* **6**, 7532 (2015).

28. Moussavi-Baygi, R. & Mofrad, M. R. K. Rapid Brownian Motion Primes Ultrafast Reconstruction of Intrinsically Disordered Phe-Gly Repeats Inside the Nuclear Pore Complex. *Sci. Rep.* **6**, 29991 (2016).

29. Li, C., Goryaynov, A. & Yang, W. The selective permeability barrier in the nuclear pore complex. *Nucleus* **7**, 430–446 (2016).

30. James, G. A. *et al.* Biofilms in chronic wounds. *Wound Repair Regen.* **16**, 37–44 (2008).

31. Moreau-Marquis, S., Stanton, B. A. & O'Toole, G. A. Pseudomonas aeruginosa biofilm formation in the cystic fibrosis airway. *Pulm. Pharmacol. Ther.* **21**, 595–599 (2008).

32. Jennings, L. K. *et al.* Pel is a cationic exopolysaccharide that cross-links extracellular DNA in the Pseudomonas aeruginosa biofilm matrix. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 11353–11358 (2015).

33. Byrd, M. S. *et al.* Genetic and biochemical analyses of the Pseudomonas aeruginosa Psl exopolysaccharide reveal overlapping roles for polysaccharide synthesis enzymes in Psl and LPS production. *Mol. Microbiol.* **73**, 622–638 (2009).

34. Ghafoor, A., Hay, I. D. & Rehm, B. H. A. Role of exopolysaccharides in Pseudomonas aeruginosa biofilm formation and architecture. *Appl. Environ. Microbiol.* **77**, 5238–5246 (2011).

35. Lai, S. K., Wang, Y.-Y., Wirtz, D. & Hanes, J. Micro- and macrorheology of mucus. *Adv. Drug Deliv. Rev.* **61**, 86–100 (2009).

36. Bansil, R. & Turner, B. S. Mucin structure, aggregation, physiological functions and biomedical applications. *Curr. Opin. Colloid Interface Sci.* **11**, 164–170 (2006).

37. Boucher, R. C. Cystic fibrosis: a disease of vulnerability to airway surface dehydration. *Trends Mol. Med.* **13**, 231–240 (2007).

38. Kater, A., Henke, M. O. & Rubin, B. K. The role of DNA and actin polymers on the polymer structure and rheology of cystic fibrosis sputum and depolymerization by gelsolin or thymosin beta 4. *Ann. N. Y. Acad. Sci.* **1112**, 140–153 (2007).

39. Sutherland, I. W. Biofilm exopolysaccharides: a strong and sticky framework. *Microbiology* **147**, 3–9 (2001).

40. Thornton, D. J. From Mucins to Mucus: Toward a More Coherent Understanding of This Essential Barrier. *Proc. Am. Thorac. Soc.* **1**, 54–61 (2004).

41. Becher, N., Adams Waldorf, K., Hein, M. & Uldbjerg, N. The cervical mucus plug: structured review of the literature. *Acta Obstet. Gynecol. Scand.* **88**, 502–513 (2009).

42. Lieleg, O. & Ribbeck, K. Biological hydrogels as selective diffusion barriers. *Trends Cell Biol.* **21**, 543–551 (2011).

43. Netti, P. A., Berk, D. A., Swartz, M. A., Grodzinsky, A. J. & Jain, R. K. Role of Extracellular Matrix Assembly in Interstitial Transport in Solid Tumors. *Cancer Res.* **60**, 2497–2503 (2000).

44. Chauhan, V. P., Stylianopoulos, T., Boucher, Y. & Jain, R. K. Delivery of molecular and nanoscale medicine to tumors: transport barriers and strategies. *Annu. Rev. Chem. Biomol. Eng.* **2**, 281–298 (2011).

45. Ernsting, M. J., Murakami, M., Roy, A. & Li, S.-D. Factors controlling the pharmacokinetics, biodistribution and intratumoral penetration of nanoparticles. *J. Control. Release Off. J. Control. Release Soc.* **172**, 782–794 (2013).

46. Jackson, A. & Gu, W. Transport Properties of Cartilaginous Tissues. *Curr. Rheumatol. Rev.* **5**, 40–50 (2009).

47. Chew, S. C. *et al.* Dynamic remodeling of microbial biofilms by functionally distinct exopolysaccharides. *mBio* **5**, e01536–14 (2014).

48. Crater, J. S. & Carrier, R. L. Barrier Properties of Gastrointestinal Mucus to Nanoparticle Transport. *Macromol. Biosci.* **10**, 1473–1483 (2010).

49. Forier, K. *et al.* Transport of nanoparticles in cystic fibrosis sputum and bacterial biofilms by single-particle tracking microscopy. *Nanomed.* **8**, 935–949 (2013).

50. Lai, S. K., Wang, Y.-Y., Hida, K., Cone, R. & Hanes, J. Nanoparticles reveal that human cervicovaginal mucus is riddled with pores larger than viruses. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 598–603 (2010).

51. Birjiniuk, A. *et al.* Single particle tracking reveals spatial and dynamic organization of the Escherichia coli biofilm matrix. *New J. Phys.* 1–14 (2014).

52. Nance, E. A. *et al.* A dense poly(ethylene glycol) coating improves penetration of large polymeric nanoparticles within brain tissue. *Sci. Transl. Med.* **4**, 149ra119–149ra119 (2012).

53. Khutoryanskiy, V. V. Beyond PEGylation: Alternative surface-modification of nanoparticles with mucus-inert biomaterials. *Adv. Drug Deliv. Rev.* (2017). doi:10.1016/j.addr.2017.07.015

54. Wühr, M. *et al.* The Nuclear Proteome of a Vertebrate. *Curr. Biol.* **25**, 2663–2671 (2015).

55. Mohr, D., Frey, S., Fischer, T., Güttler, T. & Görlich, D. Characterisation of the passive permeability barrier of nuclear pore complexes. *EMBO J.* **28**, 2541–2553 (2009).

56. Duncan, G. A. *et al.* Microstructural alterations of sputum in cystic fibrosis lung disease. *JCI Insight* **1**,

57. Ehre, C. *et al.* An Improved Inhaled Mucolytic to Treat Airway Muco-obstructive Diseases. *Am. J. Respir. Crit. Care Med.* **199**, 171–180 (2018).

58. Yuan, S. *et al.* Oxidation increases mucin polymer cross-links to stiffen airway mucus gels. *Sci. Transl. Med.* **7**, 276ra27-276ra27 (2015).

59. Johansson, M. E. *et al.* Bacteria penetrate the normally impenetrable inner colon mucus layer in both murine colitis models and patients with ulcerative colitis. *Gut* gutjnl–2012 (2013).

60. Yildiz, H. M., Carlson, T. L., Goldstein, A. M. & Carrier, R. L. Mucus Barriers to Microparticles and Microbes are Altered in Hirschsprung's Disease. *Macromol. Biosci.* **15**, 712–718 (2015).

61. Smith-Dupont, K. B. *et al.* Probing the potential of mucus permeability to signify preterm birth risk. *Sci. Rep.* **7**, 10302 (2017).

62. Shaikh, R., Singh, T. R. R., Garland, M. J., Woolfson, A. D. & Donnelly, R. F. Mucoadhesive drug delivery systems. *J. Pharm. Bioallied Sci.* **3**, 89–100 (2011).

63. Olmsted, S. S. *et al.* Diffusion of Macromolecules and Virus-Like Particles in Human Cervical Mucus. *Biophys. J.* **81**, 1930–1937 (2001).

64. Grodzinsky, A. *Forces, Fields and Flows in Biological Systems*. (Garland Science, 2011).

65. Tseng, B. S. *et al.* The extracellular matrix protects Pseudomonas aeruginosa biofilms by limiting the penetration of tobramycin. *Environ. Microbiol.* **15**, 2865–2878 (2013).

66. Byun, S., Sinskey, Y. L., Lu, Y. C. S., Frank, E. H. & Grodzinsky, A. J. Transport and binding of tumor necrosis factor-α in articular cartilage depend on its quaternary structure. *Arch. Biochem. Biophys.* **540**, 1–8 (2013).

67. Dowd, C. J., Cooney, C. L. & Nugent, M. A. Heparan sulfate mediates bFGF transport through basement membrane by diffusion with rapid reversible binding. *J. Biol. Chem.* **274**, 5236–5244 (1999).

68. Mammen, M., Choi, S.-K. & Whitesides, G. M. Polyvalent interactions in biological systems: implications for design and use of multivalent ligands and inhibitors. *Angew. Chem. Int. Ed.* **37**, 2754–2794 (1998).

69. Cohen, M. *et al.* Influenza A penetrates host mucus by cleaving sialic acids with neuraminidase. *Virol. J.* **10**, 321 (2013).

70. Yang, X. *et al.* A beneficiary role for neuraminidase in influenza virus penetration through the respiratory mucus. *PloS One* **9**, e110026 (2014).

71. Ehre, C. *et al.* Overexpressing mouse model demonstrates the protective role of Muc5ac in the lungs. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 16528–16533 (2012).

72. Bernkop-Schnürch, A. Thiomers: a new generation of mucoadhesive polymers. *Adv. Drug Deliv. Rev.* **57**, 1569–1582 (2005).

73. Li, L., Lieleg, O., Jang, S., Ribbeck, K. & Han, J. A microfluidic in vitro system for the quantitative study of the stomach mucus barrier function. *Lab. Chip* **12**, 4071–9 (2012).

74. Alipour, M., Suntres, Z. E. & Omri, A. Importance of DNase and alginate lyase for enhancing free and liposome encapsulated aminoglycoside activity against Pseudomonas aeruginosa. *J. Antimicrob. Chemother.* **64**, 317–325 (2009).

75. Hunt, B. E., Weber, A., Berger, A., Ramsey, B. & Smith, A. L. Macromolecular mechanisms of sputum inhibition of tobramycin activity. *Antimicrob. Agents Chemother.* **39**, 34–39 (1995).

76. Cao, B. *et al.* Antibiotic penetration and bacterial killing in a Pseudomonas aeruginosa biofilm model. *J. Antimicrob. Chemother.* **70**, 2057–2063 (2015).

77. Nichols, W. W., Dorrington, S. M., Slack, M. P. & Walmsley, H. L. Inhibition of tobramycin diffusion by binding to alginate. *Antimicrob. Agents Chemother.* **32**, 518–523 (1988).

78. Nichols, W. W., Evans, M. J., Slack, M. P. & Walmsley, H. L. The penetration of antibiotics into aggregates of mucoid and non-mucoid Pseudomonas aeruginosa. *J. Gen. Microbiol.* **135**, 1291–1303 (1989).

79. Anderl, J. N., Franklin, M. J. & Stewart, P. S. Role of antibiotic penetration limitation in Klebsiella pneumoniae biofilm resistance to ampicillin and ciprofloxacin. *Antimicrob. Agents Chemother.* **44**, 1818–1824 (2000).

80. Ciofu, O., Tolker-Nielsen, T., Jensen, P. Ø., Wang, H. & Høiby, N. Antimicrobial resistance, respiratory tract infections and role of biofilms in lung infections in cystic fibrosis patients. *Adv. Drug Deliv. Rev.* **85**, 7–23 (2015).

81. Walters, M. C., Roe, F., Bugnicourt, A., Franklin, M. J. & Stewart, P. S. Contributions of antibiotic penetration, oxygen limitation, and low metabolic activity to tolerance of Pseudomonas aeruginosa biofilms to ciprofloxacin and tobramycin. *Antimicrob. Agents Chemother.* **47**, 317–323 (2003).

82. Billings, N. *et al.* The extracellular matrix Component Psl provides fast-acting antibiotic defense in Pseudomonas aeruginosa biofilms. *PLoS Pathog.* **9**, e1003526 (2013).

83. Benincasa, M. *et al.* Activity of antimicrobial peptides in the presence of polysaccharides produced by pulmonary pathogens. *J. Pept. Sci.* **15**, 595–600 (2009).

84. Batoni, G., Maisetta, G. & Esin, S. Antimicrobial peptides and their interaction with biofilms of medically relevant bacteria. *Biochim. Biophys. Acta* **1858**, 1044–1060 (2016).

85. Herasimenka, Y. *et al.* Interaction of antimicrobial peptides with bacterial polysaccharides from lung pathogens. *Peptides* **26**, 1127–1132 (2005).

86. Chan, C., Burrows, L. L. & Deber, C. M. Alginate as an auxiliary bacterial membrane: binding of membrane-active peptides by polysaccharides. *J. Pept. Res.* **65**, 343–351 (2005).

87. Yin, L. M., Lee, S., Mak, J. S. W., Helmy, A. S. & Deber, C. M. Differential Binding of L- vs. D-isomers of Cationic Antimicrobial Peptides to the Biofilm Exopolysaccharide Alginate. (2013).

88. Chan, C., Burrows, L. L. & Deber, C. M. Helix induction in antimicrobial peptides by alginate in biofilms. *J. Biol. Chem.* **279**, 38749–38754 (2004).

89. Kuo, H. H., Chan, C., Burrows, L. L. & Deber, C. M. Hydrophobic Interactions in Complexes of Antimicrobial Peptides with Bacterial Polysaccharides. *Chem. Biol. Amp Drug Des.* **69**, 405–412 (2007).

90. Bucki, R., Sostarecz, A. G., Byfield, F. J., Savage, P. B. & Janmey, P. A. Resistance of the antibacterial agent ceragenin CSA-13 to inactivation by DNA or F-actin and its activity in cystic fibrosis sputum. *J. Antimicrob. Chemother.* **60**, 535–545 (2007).

91. Bucki, R., Namiot, D. B., Namiot, Z., Savage, P. B. & Janmey, P. A. Salivary mucins inhibit antibacterial activity of the cathelicidin-derived LL-37 peptide but not the cationic steroid CSA-13. *J. Antimicrob. Chemother.* **62**, 329–335 (2008).

92. Bucki, R., Byfield, F. J. & Janmey, P. A. Release of the antimicrobial peptide LL-37 from DNA/F-actin bundles in cystic fibrosis sputum. *Eur. Respir. J.* **29**, 624–632 (2007).

93. Weiner, D. J., Bucki, R. & Janmey, P. A. The Antimicrobial Activity of the Cathelicidin LL37 Is Inhibited by F-actin Bundles and Restored by Gelsolin. *Am. J. Respir. Cell Mol. Biol.* **28**, 738–745 (2003).

94. Cone, R. A. Barrier properties of mucus. *Adv. Drug Deliv. Rev.* **61**, 75–85 (2009).

95. Li, L. D. *et al.* Spatial configuration and composition of charge modulates transport into a mucin hydrogel barrier. *Biophys. J.* **105**, 1357–1365 (2013).

96. Ma, C. D., Wang, C., Acevedo-Vélez, C., Gellman, S. H. & Abbott, N. L. Modulation of hydrophobic interactions by proximally immobilized ions. *Nature* **517**, 347–350 (2015).

97. Breeuwer, M. & Goldfarb, D. S. Facilitated nuclear transport of histone H1 and other small nucleophilic proteins. *Cell* **60**, 999–1008 (1990).

98. Frey, S. & Görlich, D. A Saturated FG-Repeat Hydrogel Can Reproduce the Permeability Properties of Nuclear Pore Complexes. *Cell* **130**, 512–523 (2007).

99. Colwell, L. J., Brenner, M. P. & Ribbeck, K. Charge as a selection criterion for translocation through the nuclear pore complex. *PLoS Comput. Biol.* **6**, e1000747 (2010).

100. 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. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 3363–3368 (2013).

101. Raveh, B. *et al.* Slide-and-exchange mechanism for rapid and selective transport through the nuclear pore complex. *Proc. Natl. Acad. Sci.* **113**, E2489–97 (2016).

102. Witten, J., Samad, T. & Ribbeck, K. Molecular characterization of mucus binding. *Biomacromolecules* (2019). doi:10.1021/acs.biomac.8b01467

103. Leal, J., Smyth, H. D. C. & Ghosh, D. Physicochemical properties of mucus and their impact on transmucosal drug delivery. *Int. J. Pharm.* **532**, 555–572 (2017).

104. Murgia, X., Loretz, B., Hartwig, O., Hittinger, M. & Lehr, C.-M. The role of mucus on drug transport and its potential to affect therapeutic outcomes. *Adv. Drug Deliv. Rev.* (2017). doi:10.1016/j.addr.2017.10.009

105. Kirkham, S., Sheehan, J. K., Knight, D., Richardson, P. S. & Thornton, D. J. Heterogeneity of airways mucus: variations in the amounts and glycoforms of the major oligomeric mucins MUC5AC and MUC5B. *Biochem. J.* **361**, 537–546 (2002).

106. Fahy, J. V. & Dickey, B. F. Airway Mucus Function and Dysfunction. *N. Engl. J. Med.* **363**, 2233–2247 (2010).

107. Duffy, C. V., David, L. & Crouzier, T. Covalently-crosslinked mucin biopolymer hydrogels for sustained drug delivery. *Acta Biomater.* **20**, 51–59 (2015).

108. Shraga, N., Belgorodsky, B. & Gozin, M. Organic Reactions Promoted by Mucin Glycoproteins. *J. Am. Chem. Soc.* **131**, 12074–12075 (2009).

109. Gordon, C. A., Hodges, N. A. & Marriott, C. Antibiotic interaction and diffusion through alginate and exopolysaccharide of cystic fibrosis-derived Pseudomonas aeruginosa. *J. Antimicrob. Chemother.* **22**, 667–674 (1988).

110. Larhed, A. W., Artursson, P. & Björk, E. The Influence of Intestinal Mucus Components on the Diffusion of Drugs. *Pharm. Res.* **15**, 66–71 (1998).

111. Huang, J. X. *et al.* Mucin binding reduces colistin antimicrobial activity. *Antimicrob. Agents Chemother.* **59**, 5925–5931 (2015).

112. Schneider-Futschik, E. K. *et al.* Sputum Active Polymyxin Lipopeptides: Activity against Cystic Fibrosis Pseudomonas aeruginosa Isolates and Their Interactions with Sputum Biomolecules. *ACS Infect. Dis.* (2018). doi:10.1021/acsinfecdis.7b00238

113. Levy, J., Smith, A. L., Kenny, M. A., Ramsey, B. & Schoenknecht, F. D. Bioactivity of Gentamicin in Purulent Sputum from Patients with Cystic Fibrosis or Bronchiectasis: Comparison with Activity in Serum. *J. Infect. Dis.* **148**, 1069–1076 (1983).

114. Saggers, B. A. & Lawson, D. Some observations on the penetration of antibiotics through mucus in vitro. *J. Clin. Pathol.* **19**, 313–317 (1966).

115. Gargano, A. F. G., Lämmerhofer, M., Lönn, H., Schoenmakers, P. J. & Leek, T. Mucin-based stationary phases as tool for the characterization of drug–mucus interaction. *J. Chromatogr. A* **1351**, 70–81 (2014).

116. Pontier, C., Pachot, J., Botham, R., Lenfant, B. & Arnaud, P. HT29-MTX and Caco-2/TC7 Monolayers as Predictive Models for Human Intestinal Absorption: Role of the Mucus Layer. *J. Pharm. Sci.* **90**, 1608–1619 (2001).

117. Hong, J. A., Neel, D. V., Wassaf, D., Caballero, F. & Koehler, A. N. Recent discoveries and applications involving small-molecule microarrays. *Curr. Opin. Chem. Biol.* **18**, 21–28 (2014).

118. Celli, J. *et al.* Viscoelastic Properties and Dynamics of Porcine Gastric Mucin. *Biomacromolecules* **6**, 1329–1333 (2005).

119. Sanders, N., Rudolph, C., Braeckmans, K., De Smedt, S. C. & Demeester, J. Extracellular barriers in respiratory gene therapy. *Adv. Drug Deliv. Rev.* **61**, 115–127 (2009).

120. Suk, J. S. *et al.* N-acetylcysteine Enhances Cystic Fibrosis Sputum Penetration and Airway Gene Transfer by Highly Compacted DNA Nanoparticles. *Mol. Ther.* **19**, 1981–1989 (2011).

121. Henderson, A. G. *et al.* Cystic fibrosis airway secretions exhibit mucin hyperconcentration and increased osmotic pressure. *J. Clin. Invest.* **124**, 3047–3060 (2014).

122. PEDERSEN, S. S., KHARAZMI, A. & ESPERSEN, F. Pseudomonas aeruginosa Alginate in Cystic Fibrosis Sputum and the Inflammatory Response. *INFECT IMMUN* **58**, 6 (1990).

123. Bradner, J. E., McPherson, O. M. & Koehler, A. N. A method for the covalent capture and screening of diverse small molecules in a microarray format. *Nat. Protoc.* **1**, 2344–2352 (2006).

124. Vardakas, K. Z., Voulgaris, G. L., Samonis, G. & Falagas, M. E. Inhaled colistin monotherapy for respiratory tract infections in adults without cystic fibrosis: a systematic review and meta-analysis. *Int. J. Antimicrob. Agents* **51**, 1–9 (2018).

125. Knowles, M. R. *et al.* Ion composition of airway surface liquid of patients with cystic fibrosis as compared with normal and disease-control subjects. *J. Clin. Invest.* **100**, 2588–2595 (1997).

126. Pezzulo, A. A. *et al.* Reduced Airway Surface pH Impairs Bacterial Killing in the Porcine Cystic Fibrosis Lung. *Nature* **487**, 109–113 (2012).

127. Saltzman, J. & Bendtsen, C. Modeling the Effect of Mucin Binding in the Gut on Drug Delivery. *Bull. Math. Biol.* 1–17 (2018). doi:10.1007/s11538-018-0417-3

128. Smith, A. C. *et al.* Albumin Inhibits Pseudomonas aeruginosa Quorum Sensing and Alters Polymicrobial Interactions. *Infect. Immun.* **85**, e00116-17 (2017).

129. Takács-Novák, K. *et al.* Protonation Equilibria of Quinolone Antibacterials. *J. Pharm. Sci.* **79**, 1023–1028 (1990).

130. Tang, B. C. *et al.* Biodegradable polymer nanoparticles that rapidly penetrate the human mucus barrier. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 19268–19273 (2009).

131. Mastorakos, P. *et al.* Highly compacted biodegradable DNA nanoparticles capable of overcoming the mucus barrier for inhaled lung gene therapy. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 8720–8725 (2015).

132. Holmén, J. M. *et al.* Mucins and their O-Glycans from human bronchial epithelial cell cultures. *Am. J. Physiol.-Lung Cell. Mol. Physiol.* **287**, L824–L834 (2004).

133. Matsui, H. *et al.* A physical linkage between cystic fibrosis airway surface dehydration and Pseudomonas aeruginosa biofilms. *Proc. Natl. Acad. Sci.* **103**, 18131–18136 (2006).

134. Hill, D. B. *et al.* A Biophysical Basis for Mucus Solids Concentration as a Candidate Biomarker for Airways Disease. *PLOS ONE* **9**, e87681 (2014).

135. Weers, J. Inhaled antimicrobial therapy – Barriers to effective treatment. *Adv. Drug Deliv. Rev.* **85**, 24–43 (2015).

136. Netsomboon, K. & Bernkop-Schnürch, A. Mucoadhesive vs. mucopenetrating particulate drug delivery. *Eur. J. Pharm. Biopharm.* **98**, 76–89 (2016).

137. Fulcher, M. L. & Randell, S. H. Human Nasal and Tracheo-Bronchial Respiratory Epithelial Cell Culture. in *Epithelial Cell Culture Protocols: Second Edition* (eds. Randell, S. H. & Fulcher, M. L.) 109–121 (Humana Press, 2013). doi:10.1007/978-1-62703-125-7_8

138. Haber, C. & Wirtz, D. Shear-Induced Assembly of λ-Phage DNA. *Biophys. J.* **79**, 1530–1536 (2000).

139. Hill, D. B. & Button, B. Establishment of Respiratory Air–Liquid Interface Cultures and Their Use in Studying Mucin Production, Secretion, and Function. in *Mucins: Methods and Protocols* (eds. McGuckin, M. A. & Thornton, D. J.) 245–258 (Humana Press, 2012). doi:10.1007/978-1-61779-513-8_15

140. Das, T. *et al.* Phenazine virulence factor binding to extracellular DNA is important for *Pseudomonas aeruginosa* biofilm formation. *Sci. Rep.* **5**, 8398 (2015).

141. Turner, K. H., Wessel, A. K., Palmer, G. C., Murray, J. L. & Whiteley, M. Essential genome of Pseudomonas aeruginosa in cystic fibrosis sputum. *Proc. Natl. Acad. Sci.* **112**, 4110–4115 (2015).

142. Samson, C. *et al.* Long-term effects of azithromycin in patients with cystic fibrosis. *Respir. Med.* **117**, 1–6 (2016).

143. Goss, C. H. & Muhlebach, M. S. Review: Staphylococcus aureus and MRSA in cystic fibrosis. *J. Cyst. Fibros.* **10**, 298–306 (2011).

144. Boegh, M., García-Díaz, M., Müllertz, A. & Nielsen, H. M. Steric and interactive barrier properties of intestinal mucus elucidated by particle diffusion and peptide permeation. *Eur. J. Pharm. Biopharm. Off. J. Arbeitsgemeinschaft Pharm. Verfahrenstechnik EV* **95**, 136–143 (2015).

145. Andrews, J. *et al.* Concentrations in plasma, epithelial lining fluid, alveolar macrophages and bronchial mucosa after a single intravenous dose of 1.6 mg/kg of iclaprim (AR-100) in healthy men. *J. Antimicrob. Chemother.* **60**, 677–680 (2007).

146. Kiem, S. & Schentag, J. J. Interpretation of Epithelial Lining Fluid Concentrations of Antibiotics against Methicillin Resistant Staphylococcus aureus. *Infect. Chemother.* **46**, 219–225 (2014).

147. Cao, B. *et al.* Diffusion Retardation by Binding of Tobramycin in an Alginate Biofilm Model. *PloS One* **11**, e0153616 (2016).

148. Drew, K. R. P., Sanders, L. K., Culumber, Z. W., Zribi, O. & Wong, G. C. L. Cationic Amphiphiles Increase Activity of Aminoglycoside Antibiotic Tobramycin in the Presence of Airway Polyelectrolytes. *J. Am. Chem. Soc.* **131**, 486–493 (2008).

149. Matalon, S., Bartoszewski, R. & Collawn, J. F. Role of epithelial sodium channels in the regulation of lung fluid homeostasis. *Am. J. Physiol. - Lung Cell. Mol. Physiol.* **309**, L1229–L1238 (2015).

150. Hirsh, A. J. *et al.* Evaluation of second generation amiloride analogs as therapy for cystic fibrosis lung disease. *J. Pharmacol. Exp. Ther.* **311**, 929–938 (2004).

151. Strong, P., Ito, K., Murray, J. & Rapeport, G. Current approaches to the discovery of novel inhaled medicines. *Drug Discov. Today* **23**, 1705–1717 (2018).

152. Larhed, A. W., Artursson, P., Gr\a asjö, J. & Björk, E. Diffusion of drugs in native and purified gastrointestinal mucus. *J. Pharm. Sci.* **86**, 660–665 (1997).

153. Bajpayee, A. G., Wong, C. R., Bawendi, M. G., Frank, E. H. & Grodzinsky, A. J. Avidin as a model for charge driven transport into cartilage and drug delivery for treating early stage post-traumatic osteoarthritis. *Biomaterials* **35**, 538–549 (2014).

154. Edward, J. T. Molecular volumes and the Stokes-Einstein equation. *J. Chem. Educ.* **47**, 261 (1970).

155. Braeckmans, K., Peeters, L., Sanders, N. N., De Smedt, S. C. & Demeester, J. Three-Dimensional Fluorescence Recovery after Photobleaching with the Confocal Scanning Laser Microscope. *Biophys. J.* **85**, 2240–2252 (2003).

156. Diffusion Time Calculator - PhysiologyWeb. Available at: https://www.physiologyweb.com/calculators/diffusion_time_calculator.html. (Accessed: 2nd March 2019)

157. Verkman, A. S., Song, Y. & Thiagarajah, J. R. Role of airway surface liquid and submucosal glands in cystic fibrosis lung disease. *Am. J. Physiol.-Cell Physiol.* **284**, C2–C15 (2003).

158. Tipirneni, K. E. *et al.* Assessment of acquired mucociliary clearance defects using micro-optical coherence tomography. *Int. Forum Allergy Rhinol.* **7**, 920–925 (2017).

159. Duncan, G. A. *et al.* An Adeno-Associated Viral Vector Capable of Penetrating the Mucus Barrier to Inhaled Gene Therapy. *Mol. Ther. - Methods Clin. Dev.* **9**, 296–304 (2018).

160. Xu, Q. *et al.* Impact of Surface Polyethylene Glycol (PEG) Density on Biodegradable Nanoparticle Transport in Mucus ex Vivo and Distribution in Vivo. *ACS Nano* **9**, 9217–9227 (2015).

161. Ratjen, F. *et al.* Pharmacokinetics of inhaled colistin in patients with cystic fibrosis. *J. Antimicrob. Chemother.* **57**, 306–311 (2006).

162. Mentz, W. M. *et al.* Deposition, Clearance, and Effects of Aerosolized Amiloride in Sheep Airways. *Am. Rev. Respir. Dis.* **134**, 938–943 (1986).

163. Noone, P. G. *et al.* Airway Deposition and Clearance and Systemic Pharmacokinetics of Amiloride Following Aerosolization With an Ultrasonic Nebulizer to Normal Airways. *Chest* **112**, 1283–1290 (1997).

164. Popov, A., Schopf, L., Bourassa, J. & Chen, H. Enhanced pulmonary delivery of fluticasone propionate in rodents by mucus-penetrating nanoparticles. *Int. J. Pharm.* **502**, 188–197 (2016).

165. 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.* **113**, 2088–2099 (2017).

166. Stewart, M. Molecular mechanism of the nuclear protein import cycle. *Nat. Rev. Mol. Cell Biol.* **8**, 195–208 (2007).

167. Stewart, M. *et al.* Molecular mechanism of translocation through nuclear pore complexes during nuclear protein import. *FEBS Lett.* **498**, 145–149 (2001).

168. Görlich, D. & Kutay, U. Transport Between the Cell Nucleus and the Cytoplasm. *Annu. Rev. Cell Dev. Biol.* **15**, 607–660 (1999).

169. Ando, D. *et al.* Nuclear Pore Complex Protein Sequences Determine Overall Copolymer Brush Structure and Function. *Biophys. J.* **106**, 1997–2007 (2014).

170. Ribbeck, K. & Görlich, D. Kinetic analysis of translocation through nuclear pore complexes. *EMBO J.* **20**, 1320–1330 (2001).

171. Ghavami, A., Veenhoff, L. M., van der Giessen, E. & Onck, P. R. Probing the Disordered Domain of the Nuclear Pore Complex through Coarse-Grained Molecular Dynamics Simulations. *Biophys. J.* **107**, 1393–1402 (2014).

172. Gamini, R., Han, W., Stone, J. E. & Schulten, K. Assembly of Nsp1 Nucleoporins Provides Insight into Nuclear Pore Complex Gating. *PLOS Comput. Biol.* **10**, e1003488 (2014).

173. Frey, S., Richter, R. P. & Görlich, D. FG-Rich Repeats of Nuclear Pore Proteins Form a Three-Dimensional Meshwork with Hydrogel-Like Properties. *Science* **314**, 815–817 (2006).

174. Frey, S. & Görlich, D. FG/FxFG as well as GLFG repeats form a selective permeability barrier with self-healing properties. *EMBO J.* **28**, 2554–2567 (2009).

175. Ribbeck, K. & Görlich, D. The permeability barrier of nuclear pore complexes appears to operate via hydrophobic exclusion. *EMBO J.* **21**, 2664–2671 (2002).

176. Bayliss, R., Littlewood, T., Strawn, L. A., Wente, S. R. & Stewart, M. GLFG and FxFG Nucleoporins Bind to Overlapping Sites on Importin-β. *J. Biol. Chem.* **277**, 50597–50606 (2002).

177. Bayliss, R. *et al.* Interaction between NTF2 and xFxFG-containing nucleoporins is required to mediate nuclear import of RanGDP 11Edited by I. B. Holland. *J. Mol. Biol.* **293**, 579–593 (1999).

178. Strawn, L. A., Shen, T. & Wente, S. R. The GLFG Regions of Nup116p and Nup100p Serve as Binding Sites for Both Kap95p and Mex67p at the Nuclear Pore Complex. *J. Biol. Chem.* **276**, 6445–6452 (2001).

179. Grant, R. P., Neuhaus, D. & Stewart, M. Structural Basis for the Interaction Between the Tap/NXF1 UBA Domain and FG Nucleoporins at 1Å Resolution. *J. Mol. Biol.* **326**, 849–858 (2003).

180. Bednenko, J., Cingolani, G. & Gerace, L. Importin β contains a COOH-terminal nucleoporin binding region important for nuclear transport. *J Cell Biol* **162**, 391–401 (2003).

181. Strawn, L. A., Shen, T., Shulga, N., Goldfarb, D. S. & Wente, S. R. Minimal nuclear pore complexes define FG repeat domains essential for transport. *Nat. Cell Biol.* **6**, 197–206 (2004).

182. Patel, S. S., Belmont, B. J., Sante, J. M. & Rexach, M. F. Natively Unfolded Nucleoporins Gate Protein Diffusion across the Nuclear Pore Complex. *Cell* **129**, 83–96 (2007).

183. Ribbeck, K., Kutay, U., Paraskeva, E. & Görlich, D. The translocation of transportin–cargo complexes through nuclear pores is independent of both Ran and energy. *Curr. Biol.* **9**, 47-S1 (1999).

184. Ando, D., Colvin, M., Rexach, M. & Gopinathan, A. Physical Motif Clustering within Intrinsically Disordered Nucleoporin Sequences Reveals Universal Functional Features. *PLOS ONE* **8**, e73831 (2013).

185. Yamada, J. *et al.* A bimodal distribution of two distinct categories of intrinsically disordered structures with separate functions in FG nucleoporins. *Mol. Cell. Proteomics* **9**, 2205–2224 (2010).

186. Zhang, S. Fabrication of novel biomaterials through molecular self-assembly. *Nat. Biotechnol.* **21**, 1171–1178 (2003).

187. Hinman, M. B., Jones, J. A. & Lewis, R. V. Synthetic spider silk: a modular fiber. *Trends Biotechnol.* **18**, 374–379 (2000).

188. Xu, M. & Lewis, R. V. Structure of a protein superfiber: spider dragline silk. *Proc. Natl. Acad. Sci.* **87**, 7120–7124 (1990).

189. Meyer, D. E. & Chilkoti, A. Quantification of the Effects of Chain Length and Concentration on the Thermal Behavior of Elastin-like Polypeptides. *Biomacromolecules* **5**, 846–851 (2004).

190. Nettles, D. L., Chilkoti, A. & Setton, L. A. Applications of elastin-like polypeptides in tissue engineering. *Adv. Drug Deliv. Rev.* **62**, 1479–1485 (2010).

191. Wright, E. R. & Conticello, V. P. Self-assembly of block copolymers derived from elastin-mimetic polypeptide sequences. *Adv. Drug Deliv. Rev.* **54**, 1057–1073 (2002).

192. Crooks, G. E., Hon, G., Chandonia, J.-M. & Brenner, S. E. WebLogo: A Sequence Logo Generator. *Genome Res.* **14**, 1188–1190 (2004).

193. Hurt, E. C. A novel nucleoskeletal-like protein located at the nuclear periphery is required for the life cycle of Saccharomyces cerevisiae. *EMBO J.* **7**, 4323–4334 (1988).

194. 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.* **100**, 2450–2455 (2003).

195. Denning, D. P., Uversky, V., Patel, S. S., Fink, A. L. & Rexach, M. The Saccharomyces cerevisiae Nucleoporin Nup2p Is a Natively Unfolded Protein. *J. Biol. Chem.* **277**, 33447–33455 (2002).

196. Rathore, O. & Sogah, D. Y. Self-Assembly of β-Sheets into Nanostructures by Poly(alanine) Segments Incorporated in Multiblock Copolymers Inspired by Spider Silk. *J. Am. Chem. Soc.* **123**, 5231–5239 (2001).

197. Chen, S. *et al.* Ionic interactions. Subnanoscale hydrophobic modulation of salt bridges in aqueous media. *Science* **348**, 555–559 (2015).

198. Migliori, A. D., Smith, D. E. & Arya, G. Molecular Interactions and Residues Involved in Force Generation in the T4 Viral DNA Packaging Motor. *J. Mol. Biol.* **426**, 4002–4017 (2014).

199. Migliori, A. D. *et al.* Evidence for an electrostatic mechanism of force generation by the bacteriophage T4 DNA packaging motor. *Nat. Commun.* **5**, 4173 (2014).

200. Hülsmann, B. B., Labokha, A. A. & Görlich, D. The Permeability of Reconstituted Nuclear Pores Provides Direct Evidence for the Selective Phase Model. *Cell* **150**, 738–751 (2012).

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

202. Lowe, A. R. *et al.* Importin-β modulates the permeability of the nuclear pore complex in a Ran-dependent manner. *eLife* **4**, (2015).

203. Friedman, A. K. & Baker, L. A. Synthetic hydrogel mimics of the nuclear pore complex display selectivity dependent on FG-repeat concentration and electrostatics. *Soft Matter* **12**, 9477–9484 (2016).

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

205. Eisele, N. B., Labokha, A. A., Frey, S., Görlich, D. & Richter, R. P. Cohesiveness tunes assembly and morphology of FG nucleoporin domain meshworks – Implications for nuclear pore permeability. *Biophys. J.* **105**, 1860–1870 (2013).

206. Eisele, N. B., Frey, S., Piehler, J., Görlich, D. & Richter, R. P. Ultrathin nucleoporin phenylalanine–glycine repeat films and their interaction with nuclear transport receptors. *EMBO Rep.* **11**, 366–372 (2010).

207. Zahn, R. *et al.* A physical model describing the interaction of nuclear transport receptors with FG nucleoporin domain assemblies. *eLife* **5**, (2016).

208. Holten-Andersen, N., Zhao, H. & Waite, J. H. Stiff Coatings on Compliant Biofibers: The Cuticle of Mytilus californianus Byssal Threads,. *Biochemistry* **48**, 2752–2759 (2009).

209. Waller, K. A. *et al.* Role of lubricin and boundary lubrication in the prevention of chondrocyte apoptosis. *Proc. Natl. Acad. Sci.* **110**, 5852–5857 (2013).

210. Rose, M. C. & Voynow, J. A. Respiratory Tract Mucin Genes and Mucin Glycoproteins in Health and Disease. *Physiol. Rev.* **86**, 245–278 (2006).

211. Lachowicz-Scroggins, M. E. *et al.* Abnormalities in MUC5AC and MUC5B Protein in Airway Mucus in Asthma. *Am. J. Respir. Crit. Care Med.* **194**, 1296–1299 (2016).

212. Du, J. *et al.* Improved Biofilm Antimicrobial Activity of Polyethylene Glycol Conjugated Tobramycin Compared to Tobramycin in Pseudomonas aeruginosa Biofilms. *Mol. Pharm.* **12**, 1544–1553 (2015).

213. Gill, A., Scanlon, T. C., Osipovitch, D. C., Madden, D. R. & Griswold, K. E. Crystal structure of a charge engineered human lysozyme having enhanced bactericidal activity. *PloS One* **6**, e16788 (2011).

214. Griswold, K. E. *et al.* Bioengineered lysozyme in combination therapies for Pseudomonas aeruginosa lung infections. *Bioengineered* **5**, 143–147 (2014).

215. Sanders, L. K. *et al.* Control of electrostatic interactions between F-actin and genetically modified lysozyme in aqueous media. *Proc. Natl. Acad. Sci.* **104**, 15994–15999 (2007).

216. Scanlon, T. C. *et al.* Enhanced antimicrobial activity of engineered human lysozyme. *ACS Chem. Biol.* **5**, 809–818 (2010).

217. Teneback, C. C. *et al.* Bioengineered lysozyme reduces bacterial burden and inflammation in a murine model of mucoid Pseudomonas aeruginosa lung infection. *Antimicrob. Agents Chemother.* **57**, 5559–5564 (2013).

218. Oefner, C. *et al.* Increased hydrophobic interactions of iclaprim with Staphylococcus aureus dihydrofolate reductase are responsible for the increase in affinity and antibacterial activity. *J. Antimicrob. Chemother.* **63**, 687–698 (2009).

*A*Appendix A: Mucus and Mucin Environments Reduce the Efficacy of Polymyxin and Fluoroquinolone Antibiotics Against *Pseudomonas Aeruginosa*

The work presented in this appendix was published in the journal *ACS Biomaterials Science & Engineering* on February 22, 2019. The authors were Tahoura Samad, Julia Co, Jacob Witten, and Katharina Ribbeck.[1]

My contributions: I performed the ED experiments (Fig. A.3C) and assisted in the writing of the manuscript.

A.1 ABSTRACT

Mucus, a biopolymer hydrogel that covers all wet epithelia of the body, is a potential site for infection by pathogenic bacteria. Mucus can bind small molecules and influence bacterial physiology, two factors that may affect the efficacy of antibiotics. Inspite of this, the impact of mucus on antibiotic activity has not been thoroughly characterized. We examined the activity of polymyxin and fluoroquinolone antibiotics against the opportunistic pathogen Pseudomonas aeruginosa in native mucus and purified mucin biopolymer environments. We found that mucus reduces the effectiveness of polymyxins and fluoroquinolones against P. aeruginosa. Mucin biopolymers MUC5AC, MUC2, and MUC5B are primary contributors to this reduction. Our findings highlight that the biomaterial environmental context should be considered when evaluating antibiotics in vitro.

A.2 INTRODUCTION

Treatment of antimicrobial-resistant infections is a major public health challenge which makes accurate in vitro evaluation of antibiotic activity critically important. Many studies have investigated how alterations in microbial genomes and transcriptomes reduce the activity of antibiotics.[2,3] However, fewer studies have explored the contributions of the microbial environment to antimicrobial activity. While previous studies have examined the impact of environmental factors like oxygen availability,[4] microbial byproducts,[5] negatively charged bacterial polysaccharides[6] and culture media[7] on efficacy, we focus on the fact that antibiotics used to treat bacterial infections in the human body must frequently act in the mucus layer.

Mucus is a hydrogel that coats and protects all wet epithelia including the eyes and the respiratory, gastrointestinal, and cervicovaginal tracts (Fig. A.1A). It is a selectively permeable layer that permits the passage of some substrates while restricting others, protecting the underlying epithelial surface. The primary gel-forming components of mucus are high-molecular weight polyanionic glycoprotein polymers called mucins (Fig. A.1B). Purified mucins exhibit important features of native mucus, including characteristic viscoelastic and selective barrier properties[8,9] and specific interactions with mucosal microbes.[10,11] Hence, purified mucins may serve as a simplified model environment for the study of mucus. Mucins present an abundance of potential electrostatic and hydrophobic binding sites for small molecules;[12,13] mucin-antibiotic binding may reduce the activity of antibiotics by sequestering them. Mucin can also modulate the physiology of bacteria,[11,14–16] which may alter their susceptibility to antibiotics (Fig. A.1C). These considerations suggest that including mucins in the
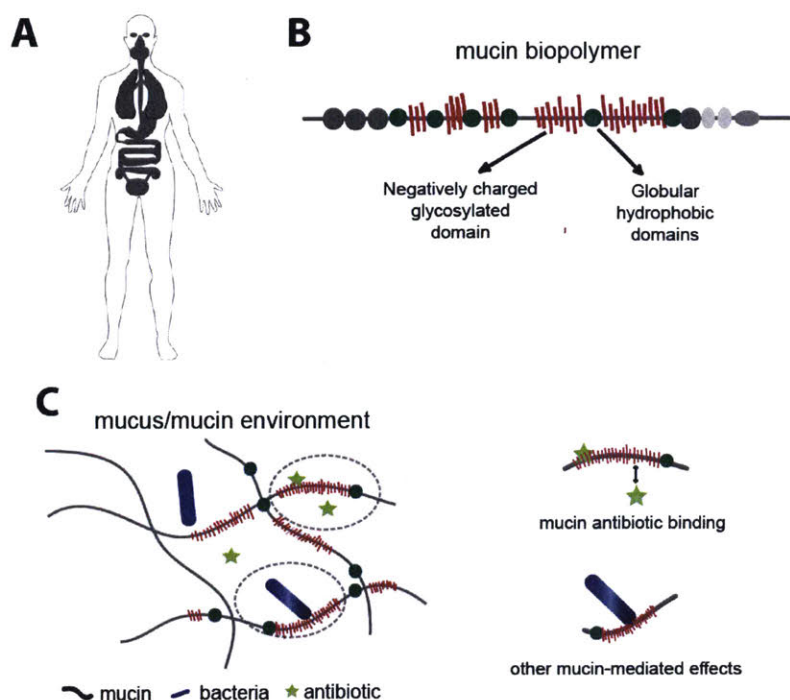
Figure A.1 (A) Mucus covers all wet epithelia including the eyes, respiratory tract, gastrointestinal tract and female reproductive tract, and is one of the primary arenas for microbes in the body. (B) Mucus contains mucins: glycoprotein polymers composed of negatively charged, densely glycosylated domains interspersed with globular hydrophobic domains (C) that may interact with antibiotics or bacteria to alter antibiotic effectiveness .

evaluation of antimicrobial activity may improve the predictability of in vitro analysis of antibiotic function.

Reduction of antimicrobial activity in sputum[17,18] or mucin,[19] and antibiotic-mucus/mucin binding[19–22] have been observed for a limited number of antibiotics. These previous studies provide an important but incomplete picture of the role played by mucus and mucin on antibiotic efficacy, and that of mucin-antibiotic binding. In order to properly evaluate the relationship between mucin-antibiotic binding and antibiotic activity in mucus, the two must be tested in the same conditions. In the work presented

here we examined the impact of native mucus from mucosal layers across the body on the activity of antibiotics against the model mucosal bacterium *Pseudomonas aeruginosa*. *P. aeruginosa* is an opportunistic pathogen that colonizes the respiratory and gastrointestinal mucosa of immune-compromised individuals as well as the mucosa of individuals with diseases such as chronic obstructive pulmonary disorder and cystic fibrosis.[23] We focused on two clinically relevant anti-pseudomonal antibiotic classes: fluoroquinolones and polymyxins. After observing that mucus inhibited the efficacy of the antibiotics, we evaluated the ability of purified mucins to mimic the effects of native mucus by comparing antibiotic activity in both environments. Our experimental design enabled direct comparison of native mucus with mucins purified from the same native mucus source. Finally, we explored the mechanism by which mucus influences antibiotic efficacy of polymyxin and fluoroquinolone antibiotics by examining mucin-antibiotic binding in the same experimental conditions.

A.3 MATERIALS AND METHODS

A.3.1 Bacterial strains and growth conditions

P. aeruginosa laboratory strain PAO1 was used for all experiments. Strains were maintained in Luria Broth (Difco, BD Falcon) and grown at 37 °C with shaking at 250 rpm.

A.3.2 Collection and Preparation of Porcine Intestinal and Gastric Mucus

Intestinal mucus was collected from porcine intestines within 24 h of slaughter and kept on ice before and during collection. The intestines were cut open and the mucus was gently scraped out using sterile glass slides. Native gastric mucus was collected from

165

porcine stomachs, which were frozen until use and kept on ice during collection. Stomachs were thawed, cut open, emptied of contents before mucus was gently scraped out using sterile glass slides. All mucus was stored at -80 °C until use.

A.3.3 Collection and Preparation of Human Saliva

Saliva was collected from human volunteers as previously described.[14] Briefly, saliva was collected from the sublingual gland using a vacuum, and kept on ice during collection. Saliva was then kept on ice until use, and used within an hour or less of collection.

A.3.4 Collection and Preparation of Human Cervical Mucus

Mucus samples were collected from 3 ovulating non-pregnant individuals as previously described,[24] as part of the same study. Briefly, after informed consent, cervical mucus was collected from three patients aged 18–45 years, during a sterile speculum exam with a 1-mL insulin syringe. Samples were collected from the external cervical os after clearing the area of vaginal discharge. Mucus was immediately snap-frozen in liquid nitrogen and stored at -80 °C.

A.3.5 Purification of MUC5AC and MUC2 from native mucus

MUC5AC and MUC2 were purified as previously described.[15] Briefly, mucus was gently scraped from pig stomachs and intestines, as described above. Native mucus samples were then solubilized in sodium chloride with sodium azide (Sigma Aldrich), benzamidine HCl (Sigma Aldrich), dibromoacetophenone (Sigma Aldrich), phenylmethlysulfonylfluoride (Sigma Aldrich), and ethylenediaminetetracetic acid (Sigma Aldrich), to prevent bacterial growth and protease degradation. After

solubilization with stirring at 4 °C, the solution was ultracentrifuged to remove insoluble contents. MUC5AC and MUC2 were then purified by gel filtration chromatography on a Sepharose (CL-2B) column. Fractions containing mucin were identified using Periodic Acid Schiff (PAS) assay, and UV_{280} measurements. Mucin fractions were dialyzed, concentrated and lyophilized. Lyophilized mucins were stored at -80 °C until use. For experiments using purified MUC5AC and MUC2, lyophilized mucins were added to PBS and solubilized at 4 °C overnight with gentle shaking.

A.3.6 Purification of MUC5B from saliva

MUC5B was purified from saliva as previously described.[14] Saliva was collected as described above and then solubilized at 4°C with stirring in sodium chloride with sodium azide (Sigma Aldrich), benzamidine HCl (Sigma Aldrich), dibromoacetophenone (Sigma Aldrich), phenylmethlysulfonylfluoride (Sigma Aldrich), and ethylenediaminetetracetic acid (Sigma Aldrich), to prevent bacterial growth and protease degradation. The solution was then centrifuged to remove insoluble contents. MUC5B was then purified by gel filtration chromatography on a Sepharose (CL-2B) column. Fractions containing mucin were identified using Periodic Acid Schiff (PAS) assay and UV_{280} measurements. Mucin fractions were dialyzed, concentrated and lyophilized. Lyophilized mucins were stored at -80 °C until use. For experiments using purified MUC5B, lyophilized mucins were added to PBS and solubilized at 4 °C overnight with gentle shaking.

A.3.7 Antibiotic Efficacy Assay

To accommodate the small volumes of native mucus, especially from human volunteers, a single antibiotic concentration was tested for comparisons between

efficacy in mucus/mucin and buffer alone. This antibiotic concentration was determined using a modified MBC-P protocol.[25] PAO1 was grown overnight in LB, diluted to 1:100 into fresh LB and then grown to mid-log (OD_{600} 0.7-1). Cells were gently pelleted to remove supernatant, (removing media, salts, and secreted microbial products), washed once with phosphate buffered saline (PBS), (Lonza, Rockland, ME, USA) and then suspended in PBS. Cells were exposed to a two-fold dilution series of each antibiotic for 2 h at 37C in static conditions and surviving cells were quantified by plating serial dilutions and counting colony forming units (CFUs). Two-fold less than the MBC-P, defined as the minimum concentration to kill all cells, (8 µg/ml colistin, 8 µg/ml polymyxin B, 1 µg/ml ofloxacin, 0.25 µg/ml ciprofloxacin, or 0.25 µg/ml levofloxacin), was used in the antibiotic efficacy assay, so that quantitative comparisons between the number of cells surviving treatment in mucin compared to buffer could be made.

To assess antibiotic efficacy in mucin and mucus gels compared to mucus/mucin free conditions, PAO1 was grown overnight in LB, diluted 1:100 into fresh LB and then grown to mid-log (OD_{600} 0.7-1). Cells were gently pelleted to remove supernatant, (removing media, salts, and secreted microbial products), washed once with phosphate buffered saline (PBS),(Lonza, Rockland, ME, USA) and then suspended in PBS. The OD_{600} of the washed cells was measured and cells were diluted into PBS. 12.5 µl of cells were added to 25 µl of native mucus, purified mucin (MUC5AC, MUC5B or MUC2; 1% w/v, dissolved in PBS) or methylcellulose (1% w/v, dissolved in PBS) (Sigma Aldrich, 15cP). 12.5 µl of antibiotics diluted into PBS were then added to mucus and cell mixture such that the final antibiotic concentration was 8 µg/ml colistin, 8 µg/ml polymyxin B, 1 µg/ml ofloxacin, 0.25 µg/ml ciprofloxacin, or 0.25 µg/ml levofloxacin. The

final concentration of mucin and methylcellulose was 0.5% w/v and the total number of cells was approximately between 5×10^6 and 1×10^7. Incubating microbes within mucus in a closed system for extended periods may result in mucus degradation, and thus does not likely reflect physiological mucus, which is constantly replenished. To avoid the loss or degradation of mucus that may occur during long-term microbial culture, we evaluated antibiotic efficacy after 2 h. Bacteria were exposed to antibiotics at 37 °C in static conditions for 2 h. Surviving cells were quantified by plating serial dilutions and counting colony forming units (CFUs).

A.3.8 Equilibrium Dialysis for Evaluation of Mucin-Antibiotic binding (ED)

Equilibrium dialysis (ED) was performed with a 12 kDa cutoff Rapid Equilibrium Dialysis device (Thermo Scientific) with 100μL of biopolymer solution or gel, or buffer control, in the sample chamber and 300uL of matching buffer in the assay chamber. Each experiment began with equal concentrations of the antibiotics loaded in both chambers ("1x concentration"). 1x concentration was 8 μg/ml colistin, 20 μg/ml polymyxin B, 1 μg/ml ofloxacin, 0.5 μg/ml ciprofloxacin, or 0.5 μg/ml levofloxacin. These concentrations were kept the same as the concentrations used in the efficacy assay when possible. The concentration of polymyxin B, ciprofloxacin and levofloxacin were increased slightly compared to what was used in the efficacy assay for ease of detection by the mass spectrometer, but this change should not have an impact on the result of the ED experiment. Equilibration took place over 4 h with shaking at 235 rpm and 37 °C, at which point aliquots were collected from the assay chamber. HPLC separation took place on an Acclaim PolarAdvantage column (3μm pores, 2.1x100mm, VWR). The HPLC method is given in Table S1. The first 5 minutes of eluent were diverted to a

waste container to avoid salt contamination of the mass spectrometer (Agilent 6410 triple quadrupole), and the rest of the eluent was analyzed using multiple reaction monitoring (MRM). Source parameters were: temperature 350 °C, gas flow 10 L/min, nebulizer 25 psi, capillary voltage 4000V, with molecular ions and fragments given in Tables S2. Multiple transitions were tracked for each of the molecules as a consistency check; only one was used for analysis but they all gave almost exactly the same results (not shown). Peak areas for each molecule's chromatogram were compared to an external standard curve to measure concentrations of each molecule.

The uptake ratio R_U of a compound was calculated by (see following section for derivation):

$$R_U = 1 + 4\left(\frac{1}{c_E} - \frac{1}{c_C}\right) \tag{S1}$$

where c_E is the concentration in the assay chamber of the experimental dialysis system and c_C is the concentration in assay chamber of the buffer-buffer dialysis control. c_E and c_C were both scaled with respect to the 1x concentration. c_E and c_C were each averaged over at least 6 wells. To avoid bias, the experimenter was blinded to which wells contained mucin and which contained only buffer for the duration of the ED experiment. Error bars for the uptake ratios shown in Fig. A.3C were calculated by propagating the standard errors for measurements of c_E and c_C to R_U.

A.3.8.1 Derivation of Equation S1

First, let us define our terms. We consider an ED system with some number of molecules of interest, which can be in one of three states: in the assay chamber (state A), in the sample chamber (state S), and nonspecifically bound to the ED system walls and thus inaccessible to measurement (state B). Furthermore, we have two ED systems

(see Fig. A.1): a buffer-buffer control (BB) and a polymer-buffer experiment (PB). We therefore name the 6 variables of interest:

x: fraction of molecules in the assay chamber in the buffer-buffer control

y: fraction of molecules non-specifically bound to the ED system in the buffer-buffer control

z: fraction of molecules in the sample chamber in the buffer-buffer control

a: fraction of molecules in the assay chamber in the polymer-buffer experiment

b: fraction of molecules non-specifically bound to the ED system in the polymer-buffer experiment

c: fraction of molecules in the sample chamber in the polymer-buffer experiment

Thus, what we seek, the uptake ratio, is given by:

$$R_U = 3c/a \qquad \text{(S2)}$$

The factor of 3 corrects for the fact that the assay chamber contains 3 times the volume of the sample chamber, and we are interested in the ratio of concentrations, not total number of molecules.

We also assume that we have reached equilibrium, and that the biopolymer does not affect equilibration between states A and B, since the assay chamber always only contains buffer and the ED chamber has a constant architecture. Therefore, we have the relation:

$$b/a = y/x \qquad \text{(S3)}$$

In the buffer-buffer control, clearly the concentrations of molecule in the sample and assay chambers will be equal, therefore:

$$x = 3\,z \qquad \text{(S4)}$$

because the assay chamber has 3 times the volume of the sample chamber. Finally, we have the normalization conditions:

$x + y + z = 1$ (S5)

$a + b + c = 1$ (S6)

Returning to Equation S2, we use these equations to get:

$R_U = 3c/a$

$\quad = 3(1\text{-}a\text{-}b)/a$

$\quad = 3/a - 3 - 3b/a$

$\quad = 3/a - 3 - 3y/x$

$\quad = 3/a - 3 - 3(1 - x - z)/x$

$\quad = 3/a - 3 - 3/x + 3 + 3z/x$

$\quad = 3/a - 3/x + 1$

$\quad = 1 + 3(1/a - 1/x)$ \qquad (S7)

Now let us consider the relationship between a and c_E and between x and c_B. Since the molecules were all added at 1x working concentration to both the assay and sample chambers, that means that 1x corresponds to having ¾ of the molecules in the assay chamber. Therefore, we have:

$c_E = 4/3\ a$ \quad (S8)

$c_B = 4/3\ x$ \quad (S9)

Substituting these relations into equation S7 gives us:

$R_U = 1 + 3(1/(\tfrac{3}{4}\ c_E) - 1/(\tfrac{3}{4}\ c_B))$

$R_U = 1 + 4(1/c_E - 1/c_B)$ \qquad (S10)

As desired.

A.3.9 Two-Chamber Antibiotic Efficacy Assay

Two-chamber antibiotic efficacy assays were conducted with slight modifications to the method described above for the antibiotic efficacy assay. Chambers for evaluating antibiotic efficacy were assembled by placing a 12-14 kDa membrane (Spectrum Labs, Catalog No. 132480, Rancho Dominguez, CA, USA) between two PSA-coated, adhesive silicone isolator wells (Grace BiosLabs Catology No. GBL666105; Sigma-Aldrich, St. Louis, MO, USA) and then placing the membrane-separated chambers between glass slides (VWR, Catolog No. 16004-422, Radnor, PA, USA). PSA coating on the silicone isolators was used to attach the isolators to the glass slides. The membrane was attached to silicone isolators using a thin layer of Gorilla glue adhesive (the Gorilla Glue Company, Cincinnati, OH, USA). Devices were cured for approximately 24 h before use. PAO1 cells were prepared as described for the original antibiotic assay.

To assess antibiotic efficacy in this split chamber, PBS or MUC5AC dissolved in PBS was first added to chambers as depicted in Supplemental Fig. A.2A. Antibiotic was added to chambers on both sides of the membrane, such that the concentration of antibiotic in both chambers was initially equal (Supplemental Fig. A.2B). For no antibiotic controls, PBS rather than antibiotic was added. Antibiotics were allowed to equilibrate between the chambers over 4 h with shaking at 250 rpm and 37°C. After 4 h, PAO1 cells or PBS were added (Supplemental Fig. A.2C). Cells, mucin and antibiotics were combined such that the final concentration of cells and mucins were identical to the original efficacy assay and such that the final concentration of antibiotic was 64 μg/ml colistin, and 4 μg/ml for ciprofloxacin. Bacteria were exposed to antibiotic at 37 °C

in static conditions for 2 h. Surviving cells were quantified by plating serial dilutions and counting colony-forming units (CFUs).

A.4 RESULTS

We first evaluated antibiotic activity in native mucus using a modified minimum bactericidal concentration for planktonic bacteria (MBC-P) assay.[25] For each antibiotic tested, the MBC-P was determined in phosphate-buffered saline (PBS) after 2h of antibiotic exposure. P. aeruginosa cells were exposed to one-half the MBC-P in phosphate-buffered saline (PBS) alone or in native mucus, and the number of surviving cells was quantified by colony forming units (CFUs) after 2 h. Porcine intestinal and gastric mucus, which are easy to source and routinely employed in studies of drug delivery,[26,27] were used to model the human gastrointestinal environment. Native human
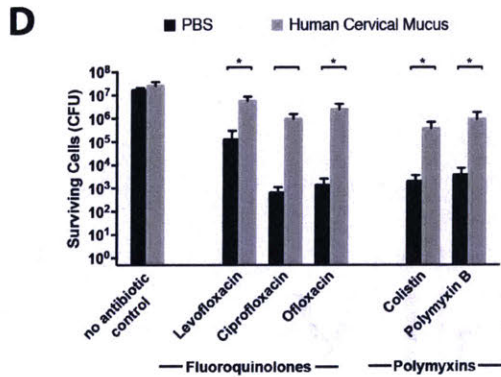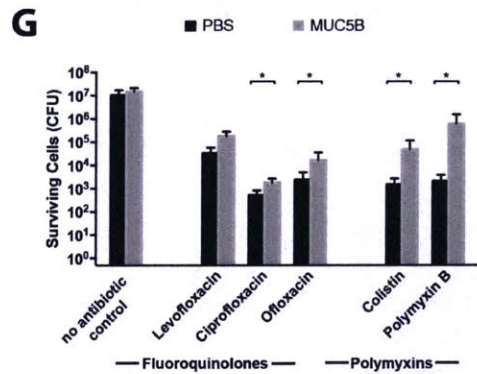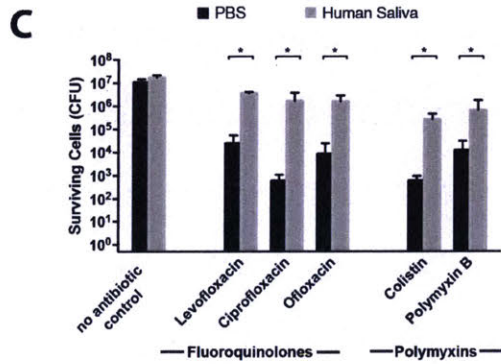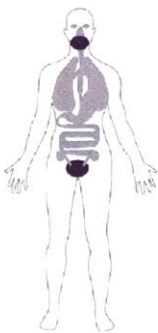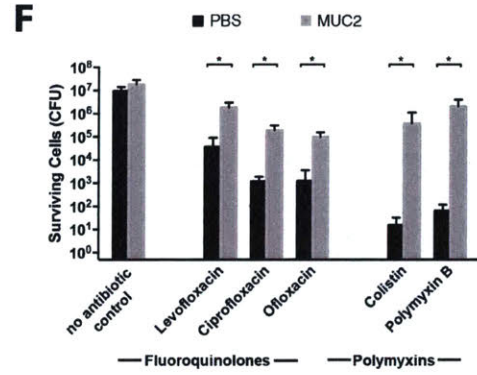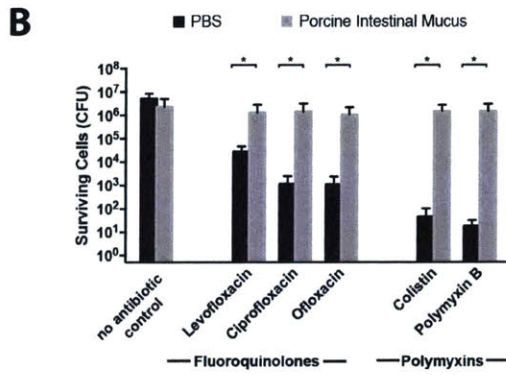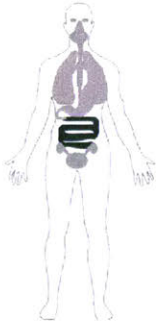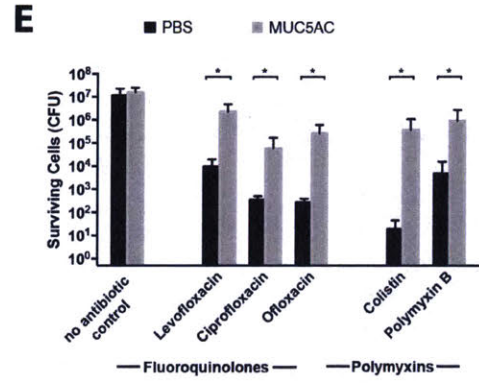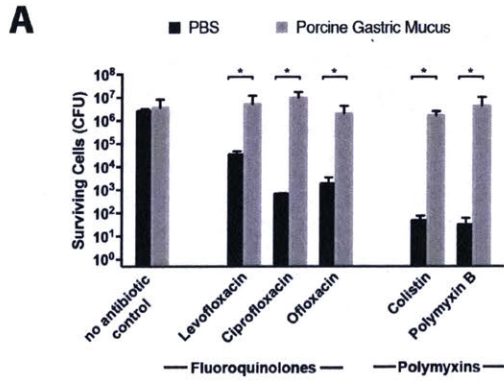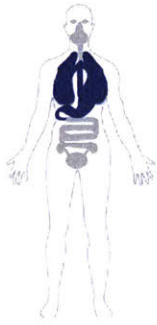
Figure A.2 Mucus and mucins reduce the efficacy of polymyxin and fluoroquinolone antibiotics against P. aeruginosa. PAO1 was exposed to polymyxin or fluoroquinolone antibiotics (polymyxin B, 8 µg/ml ; colistin, 8 µg/ml; ofloxacin, 1 µg/ml; ciprofloxacin, 0.25 µg/ml,; or levofloxacin, 0.25 µg/ml) in mucin-free buffer (PBS), and in native mucus from different surfaces in the body: (A) native gastric mucus, (B) native intestinal mucus, (C) saliva and (D) native cervical mucus. Cells were also exposed to antibiotics in purified mucins (0.5% w/v dissolved in PBS): (E) MUC5AC, the primary mucin present in the human lungs and stomach (F) MUC2, the primary mucin found in the intestines and (G) MUC5B, a mucin primarily found in saliva and the cervix. Cells were exposed to antibiotics for 2 h at 37°C and surviving cells were quantified by serial dilution and plating. All mucus and mucin samples increased the number of surviving cells. No antibiotic control condition for each mucus and mucin type demonstrated that the presence of mucus or mucin does not lead to a substantial increase in PAO1 growth. All error bars represent standard deviation of biological replicates (n ≥ 3). (*) indicate a significant increase in the number of surviving cells after antibiotic exposure in mucus/mucin compared to that in mucus/mucin-free buffer (PBS), as determined by the t-test ($P < 0.05$).

saliva and cervical mucus were tested in order to measure the effects of these environments on antibiotic activity. All mucus types significantly reduced the effectiveness of polymyxin and fluoroquinolone antibiotics relative to the buffer-only controls (Fig. A.2A-D). When cells were exposed to polymyxin antibiotics in gastric (Fig. A.2A) or intestinal (Fig. A.2B) mucus, we observed at least a 49,000-fold increase in cell survival compared to antibiotic exposure in PBS alone. For fluoroquinolones, gastric (Fig. A.2A) and intestinal (Fig. A.2B) mucus provided a 75-fold or greater increase in cells surviving antibiotic exposure versus PBS alone. Native human saliva (Fig. A.2C) and cervical mucus (Fig. A.2D) reduced P. aeruginosa killing by all antibiotics tested relative to buffer alone, with at least 50-fold and 180-fold increases in cell survival for polymyxin and fluoroquinolone antibiotics respectively. For all mucus samples tested, cells incubated for 2 h in mucus alone (Fig. A.2A-D, no antibiotic control condition) showed no difference in cell counts between the buffer-only and mucus conditions, confirming that the increase in bacterial survival observed with mucus present was due to a reduction in antibiotic activity rather than an increase in bacterial growth stimulated

by mucus. These data suggest that the reduction in antibiotic efficacy in its presence may be a general property of mucus, independent of anatomical source.

We hypothesized that the increased survival of bacteria after antibiotic exposure in mucus environments may be due to a component of mucus common across all these niches: mucin proteins. We tested whether three distinct mucins, MUC2, MUC5AC, and MUC5B also reduced the activity of polymyxins and fluoroquinolones against *P. aeruginosa*, as observed in native mucus samples. MUC2 is the predominant mucin in the intestines and thus may serve as a simplified model of the human intestinal mucus environment.[28] MUC5AC is predominant in the stomach and lungs[28] and while MUC5B is the primary mucin found in the oral cavity and female reproductive tract, and is also found in the lungs.[28,29] MUC2 and MUC5AC were first purified from the native porcine intestinal and gastric mucus samples which were used to evaluate antibiotic efficacy. Similarly, MUC5B was purified from human saliva. Purification of mucins directly from these sources—rather than using commercially available mucins—enabled meaningful comparisons with the results from native mucus. Additionally, it has been reported that compared to commercially available mucins, proteins purified in this manner better recapitulate the properties of native mucus.[30]

Purified MUC5AC (Fig. A.2E), MUC2 (Fig. A.2F), and MUC5B (Fig. A.2G) significantly reduced the efficacy of all antibiotics tested relative to mucus-free buffer controls. When cells were exposed to polymyxin antibiotics in MUC5AC (Fig. A.2E) or MUC2 (Fig. A.2G), we observed a 10,000-fold or greater increase in cell survival compared to antibiotic exposure in PBS alone and a 140-fold increase for fluoroquinolone antibiotics (Fig. 2E-G). In MUC5B, we observed a 30-fold or greater

increase in cells surviving polymyxin antibiotic exposure (Fig. A.2G) versus buffer. For fluoroquinolones in MUC5B, we measured 30-fold, 4-fold, and 12-fold increases in cells surviving exposure to ofloxacin, ciprofloxacin, or levofloxacin, respectively (Fig. A.2G). Cell counts after incubation for 2 h without antibiotics in buffer or mucin solution were not substantially different (Fig. A.2E-G, no antibiotic control condition), demonstrating that the presence of mucins does not increase cell growth but reduces antibiotic activity. Our data suggest that mucins are primary, though not necessarily singular, contributors to the reduction in antibiotic efficacy observed in mucus. These data also highlight the potential for purified mucins as a three-dimensional model biomaterial for more accurate evaluation of antibiotic activity.

We were further interested in understanding the mechanism by which mucins and mucus affect antibiotic efficacy. Methylcellulose, a polymer often used as a mucin mimetic due to similarities in its viscoelastic properties,[31] did not impact antibiotic activity to the same degree for all antibiotics tested (Fig. A.3A), suggesting that the mucin-mediated reduction in antibiotic efficacy is not due to macromolecular crowding. Alternatively, we hypothesized that mucin-antibiotic binding could reduce the free concentration of antibiotic available to kill bacterial cells. We performed equilibrium dialysis to evaluate whether mucins were binding antibiotics. We prepared two chambers separated by a 12 kDa molecular weight cut-off membrane that permitted the passage of antibiotics, but not mucins (Fig. A.3B). One chamber contained mucins (0.5% w/v dissolved in PBS) and the other PBS; the initial concentration of antibiotic was equal in both. The system equilibrated for 4 h before the final concentration of free antibiotic in the buffer chamber was quantified by mass spectrometry. Buffer

A

Legend: PBS (black), MUC5AC (light), Methylcellulose (gray)

Y-axis: Surviving Cells (CFU), 10^0 to 10^8

X-axis categories: no antibiotic control, Levofloxacin, Ciprofloxacin, Ofloxacin, Colistin, Polymyxin B

— Fluoroquinolones —   — Polymyxins —

B

Key
★ antibiotic
〜 mucin

Buffer   Mucin

membrane

C

Legend: MUC5AC (black), MUC2 (gray), MUC5B (dark gray), Methylcellulose (light)

Y-axis: Uptake ratio, 0 to 4

X-axis categories: Colistin, Polymyxin B, Levofloxacin, Ciprofloxacin, Ofloxacin

Figure A.3 Evaluating the mechanism of mucus/mucin reduction of antibiotic efficacy. Although MUC5AC (0.5% w/v) protected bacterial cells from killing compared to buffer alone, (A) methylcellulose (0.5% w/v) did not, suggesting that neither mechanical cues nor macromolecular crowding are responsible for the protection observed. All error bars represent standard deviation of biological replicates (n ≥ 3). (*) indicates a significant increase in the number of surviving cells after antibiotic exposure in mucus compared to that in mucin-free buffer (PBS), as determined by the t-test ($P < 0.01$) (B) Schematic illustration of experimental set up for equilibrium dialysis experiments to evaluate polymyxin and fluoroquinolone binding to mucin (C) Uptake ratio from equilibrium dialysis for each antibiotic with MUC5AC, MUC2, MUC5B and methylcellulose, in the conditions used to evaluate antibiotic efficacy. Ratios greater than 1 (red dotted line) indicate binding. Colistin and polymyxin B both bind to MUC5AC, MUC2, and MUC5B but not to methylcellulose. Uptake ratios of fluoroquinolone antibiotics to biopolymers are close to 1, consistent with weak or no binding. Error bars represent standard error of biological replicates (n ≥ 3) (see Methods for standard error calculation).

composition and mucin concentration were maintained between the antibiotic efficacy experiments and these assays in order to ensure consistency in ionic strength, pH and the number of potential mucin binding sites, which are experimental conditions expected to affect binding. It was expected that if an antibiotic bound to mucin, the relative concentration of antibiotic would increase in the mucin chamber and decrease in the buffer chamber. This was quantified by calculating the uptake ratio, defined as the mucin chamber:buffer chamber antibiotic concentration ratio. Uptake ratios greater than one indicate binding. For MUC5AC and MUC2, we found that the average uptake ratios for colistin and polymyxin were greater than or equal to 2.4, and for MUC5B, 1.4 or more. These data suggest that polymyxin antibiotics bind mucins in the conditions used to evaluate antibiotic efficacy (Fig. A.3C), with stronger binding to MUC5AC and MUC2 compared with MUC5B. This binding is likely electrostatic in nature (cationic polymyxins and polyanionic mucins), and may account for the reduction in polymyxin activity in mucus and mucins (Fig. A.2).

Despite the reduction in activity observed for fluoroquinolone antibiotics in mucus and mucins (Fig. A.2), we found that the uptake ratios for levofloxacin, ciprofloxacin and ofloxacin were close to 1, suggesting weak or no binding of fluoroquinolones to mucins in the conditions used to evaluate antimicrobial efficacy (Fig. A.3C). This is consistent with the fact that the fluoroquinolones, with no net charge and moderate hydrophobicity, are unlikely to bind to mucin through electrostatic or hydrophobic means. These results suggest that mechanisms other than mucin-antibiotic binding contribute to the reduction in fluoroquinolone activity observed in mucus/mucins. To pursue this hypothesis we performed a modified antibiotic efficacy assay in which we exposed *P. aeruginosa* to antibiotics in a system with two connected chambers separated by a 12 kDa MWCO membrane that allowed the selective diffusion of only antibiotics between them. As in the equilibrium dialysis experiments, we first added mucin, buffer, and antibiotic to the chambers and allowed the system to equilibrate for 4 h before adding the bacteria. Cells were exposed to antibiotic for 2 h and then quantified by serial dilution and CFU counting (Fig A.S1).

This modified assay enabled us to compare antibiotic efficacy in two conditions with mucins: one in which cells and mucin were in the same chamber (Fig. A.4, ii), and one in which cells were separated from mucins by a membrane (Fig. A.4, iii). Fig. 4 shows that exposure to ciprofloxacin with mucin and cells in the same chamber (Fig. 4, ii) significantly increased bacterial survival relative to a mucus-free buffer control (Fig. 4, i), consistent with our initial efficacy experiments (Fig. A.2). However, when cells were separated from mucin by a membrane and exposed to ciprofloxacin, no such increase was observed (Fig. A.4, iii). For a mucin-binding antibiotic we would expect to see a
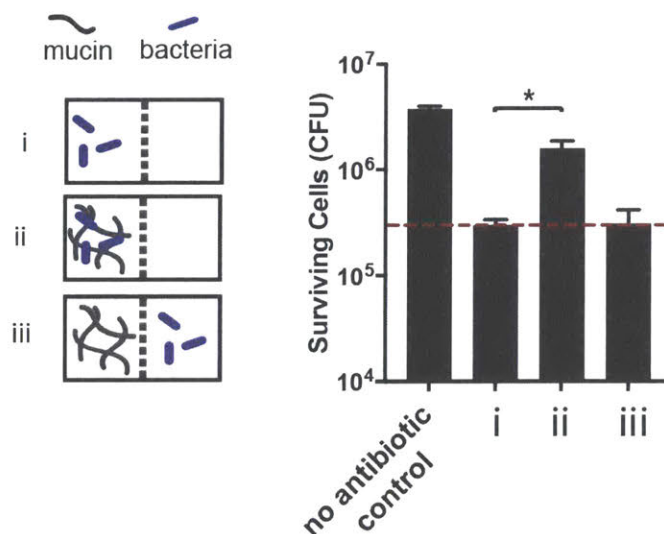
Figure A.4 Two chambers were separated by a 12 kDa MWCO membrane that allowed for the free diffusion of antibiotics, but not mucin or cells, between the chambers. As in the equilibrium dialysis experiments, mucin, buffer, and antibiotic were added to the chambers (Fig. S1) and the system was left to equilibrate for 4 h before adding P. aeruginosa cells. Bacteria were exposed to ciprofloxacin for 2 h in three different conditions: (i) buffer with no mucin, (ii) mucin, and (iii) buffer with mucin in the connected chamber, separated from the cells by the membrane. In condition iii, mucin was not in direct contact with the bacteria, but could still interact with the antibiotic. Ciprofloxacin efficacy was reduced relative to the mucin-free control in condition ii, recapitulating the result from Fig. 2 in this modified system. However, in condition iii (bacteria were separated from mucin by a membrane and exposed to ciprofloxacin in buffer), no increase in surviving cells was observed. This suggests that mucin is not able to bind and reduce the antibiotic concentration of ciprofloxacin and that a mechanism other than mucin-antibiotic binding accounts for the observed reduction in ciprofloxacin efficacy. Dotted red line marks the number of surviving cells in the mucin free control (i) . All error bars represent standard deviation of biological replicates ($n \geq 3$). (*) indicates a significant increase in the number of surviving cells after antibiotic exposure compared to that in mucin-free buffer, as determined by the t-test ($P < 0.0001$).

reduction in antibiotic activity even when mucins are separate from the cells, since mucin-antibiotic binding would reduce the free antibiotic concentration. Indeed, for colistin, which binds mucin strongly, this is exactly what was observed (Fig A.S2). These data support our hypothesis that for fluoroquinolone antibiotics like ciprofloxacin, mucin may be mediating the reduction in antibiotic efficacy via a mechanism other than

mucin-antibiotic binding. We hypothesize that possible mechanisms include mucin modulation of bacterial physiology or mucin reduction of antibiotic uptake by the cells, however further studies will be required to elucidate the exact mechanism(s).

A.5 CONCLUSION

We examined the impact of native mucus from surfaces throughout the body on the efficacy of antibiotics against *Pseudomonas aeruginosa*, a formidable opportunistic pathogen. Our data show that mucus substantially diminishes the activity of polymyxin and fluoroquinolone antibiotics against *P. aeruginosa*. Mucin biopolymers, the gel-forming components of mucus, are primary contributors to this effect and are strong candidates for the construction of a model environment that would enable more accurate in vitro antimicrobial evaluation. We determined that antibiotic binding by mucin likely plays a role in the reduced effectiveness of polymyxin antibiotics, but that mucin may reduce the activity of fluoroquinolones via alternative mechanisms. While this work focused on two relevant classes of antibiotics and P. aeruginosa, mucus and mucins are likely to impact other antibiotic classes and other mucosal microbes. Our findings highlight the importance of considering the biomaterial environment for the in vitro evaluation of antibiotics, particularly for mucosal pathogens. Understanding the mechanism(s) by which mucus reduces antibiotic activity is important for devising strategies to maximize therapeutic effect. Thinking broadly, this work highlights important considerations for the future design of antimicrobial releasing polymers and materials, as our data suggest that researchers should consider not only the ability of these polymers and materials to bind and release antimicrobials, but also whether

183

interactions between bacteria and the polymer/material may reduce the efficacy of the antimicrobial.

A.6 REFERENCES

1. Samad, T., Co, J. Y., Witten, J. & Ribbeck, K. Mucus and Mucin Environments Reduce the Efficacy of Polymyxin and Fluoroquinolone Antibiotics against Pseudomonas aeruginosa. *ACS Biomater. Sci. Eng.* (2019). doi:10.1021/acsbiomaterials.8b01054
2. Kapoor, G., Saigal, S. & Elongavan, A. Action and resistance mechanisms of antibiotics: A guide for clinicians. *J. Anaesthesiol. Clin. Pharmacol.* **33**, 300 (2017).
3. Breidenstein, E. B., de la Fuente-Núñez, C. & Hancock, R. E. Pseudomonas aeruginosa: all roads lead to resistance. *Trends Microbiol.* **19**, 419–426 (2011).
4. Schaible, B., Taylor, C. T. & Schaffer, K. Hypoxia increases antibiotic resistance in Pseudomonas aeruginosa through altering the composition of multidrug efflux pumps. *Antimicrob. Agents Chemother.* **56**, 2114–2118 (2012).
5. Radlinski, L. *et al.* Pseudomonas aeruginosa exoproducts determine antibiotic efficacy against Staphylococcus aureus. *PLoS Biol.* **15**, e2003981 (2017).
6. Tseng, B. S. *et al.* The extracellular matrix protects Pseudomonas aeruginosa biofilms by limiting the penetration of tobramycin. *Environ. Microbiol.* **15**, 2865–2878 (2013).
7. Crabbé, A. *et al.* Antimicrobial efficacy against Pseudomonas aeruginosa biofilm formation in a three-dimensional lung epithelial model and the influence of fetal bovine serum. *Sci. Rep.* **7**, 43321 (2017).
8. Celli, J. *et al.* Viscoelastic Properties and Dynamics of Porcine Gastric Mucin. *Biomacromolecules* **6**, 1329–1333 (2005).
9. Li, L. D. *et al.* Spatial configuration and composition of charge modulates transport into a mucin hydrogel barrier. *Biophys. J.* **105**, 1357–1365 (2013).
10. Celli, J. P. *et al.* Helicobacter pylori moves through mucus by reducing mucin viscoelasticity. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 14321–14326 (2009).
11. Caldara, M. *et al.* Mucin Biopolymers Prevent Bacterial Aggregation by Retaining Cells in the Free-Swimming State. *Curr. Biol.* **22**, 2325–2330 (2012).
12. Bansil, R. & Turner, B. S. Mucin structure, aggregation, physiological functions and biomedical applications. *Curr. Opin. Colloid Interface Sci.* **11**, 164–170 (2006).
13. Witten, J., Samad, T. & Ribbeck, K. Selective permeability of mucus barriers. *Curr. Opin. Biotechnol.* **52**, 124–133 (2018).
14. Frenkel, E. S. & Ribbeck, K. Salivary mucins protect surfaces from colonization by cariogenic bacteria. *Appl. Environ. Microbiol.* AEM.02573-14 (2014). doi:10.1128/AEM.02573-14
15. Kavanaugh, N. L., Zhang, A. Q., Nobile, C. J., Johnson, A. D. & Ribbeck, K. Mucins Suppress Virulence Traits of Candida albicans. *mBio* **5**, e01911-14 (2014).
16. Frenkel, E. S. & Ribbeck, K. Salivary mucins promote the coexistence of competing oral bacterial species. *ISME J.* **11**, 1286 (2017).

17. Mendelman, P. M. *et al.* Aminoglycoside penetration, inactivation, and efficacy in cystic fibrosis sputum. *Am. Rev. Respir. Dis.* **132**, 761–765 (1985).
18. King, P., Lomovskaya, O., Griffith, D. C., Burns, J. L. & Dudley, M. N. In vitro pharmacodynamics of levofloxacin and other aerosolized antibiotics under multiple conditions relevant to chronic pulmonary infection in cystic fibrosis. *Antimicrob. Agents Chemother.* **54**, 143–148 (2010).
19. Huang, J. X. *et al.* Mucin binding reduces colistin antimicrobial activity. *Antimicrob. Agents Chemother.* **59**, 5925–5931 (2015).
20. Bhat, P. G., Flanagan, D. R. & Donovan, M. D. The limiting role of mucus in drug absorption: Drug permeation through mucus solution. *Int. J. Pharm.* **126**, 179–187 (1995).
21. Niibuchi, J.-J., Aramaki, Y. & Tsuchiya, S. Binding of antibiotics to rat intestinal mucin. *Int. J. Pharm.* **30**, 181–187 (1986).
22. Ramphal, R., Lhermitte, M., Filliat, M. & Roussel, P. The binding of anti-pseudomonal antibiotics to macromolecules from cystic fibrosis sputum. *J. Antimicrob. Chemother.* **22**, 483–490 (1988).
23. Klockgether, J. & Tümmler, B. Recent advances in understanding Pseudomonas aeruginosa as a pathogen. *F1000Research* **6**, (2017).
24. Smith-Dupont, K. B. *et al.* Probing the potential of mucus permeability to signify preterm birth risk. *Sci. Rep.* **7**, 10302 (2017).
25. Mah, T.-F. *et al.* A genetic basis for Pseudomonas aeruginosa biofilm antibiotic resistance. *Nature* **426**, 306–310 (2003).
26. Crater, J. S. & Carrier, R. L. Barrier Properties of Gastrointestinal Mucus to Nanoparticle Transport. *Macromol. Biosci.* **10**, 1473–1483 (2010).
27. Boegh, M. & Nielsen, H. M. Mucus as a barrier to drug delivery – understanding and mimicking the barrier properties. *Basic Clin. Pharmacol. Toxicol.* **116**, 179–186 (2015).
28. McGuckin, M. A., Thornton, D. J. & Whitsett, J. A. Mucins and mucus. in *Mucosal Immunology* 231–250 (Elsevier, 2015).
29. Andersch-Björkman, Y., Thomsson, K. A., Larsson, J. M. H., Ekerhovd, E. & Hansson, G. C. Large scale identification of proteins, mucins, and their O-glycosylation in the endocervical mucus during the menstrual cycle. *Mol. Cell. Proteomics* **6**, 708–716 (2007).
30. Celli, J. P. *et al.* Rheology of gastric mucin exhibits a pH-dependent sol- gel transition. *Biomacromolecules* **8**, 1580–1586 (2007).
31. Ivic, A. *et al.* Critical evaluation of methylcellulose as an alternative medium in sperm migration tests. *Hum. Reprod.* **17**, 143–149 (2002).

A.7 SUPPLEMENTAL INFORMATION
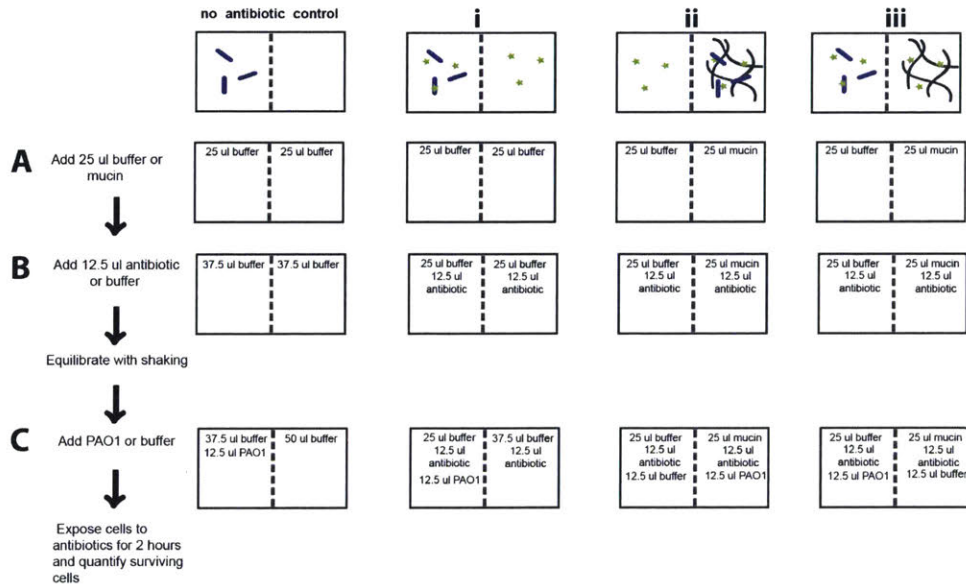
Supplemental Figure A.S1 Schematic of the two-chamber antibiotic efficacy assay. (A) Mucin or buffer were added to the chamber depending on the different conditions, which are illustrated in Fig. 4. (No antibiotic control, condition i, condition ii, and condition iii). (B) Antibiotic or buffer were then added on each side of the membrane. The system was allowed to equilibrate for 4 h at 37°C with shaking. (C) Bacteria, prepared as for the original efficacy assay, or buffer were added to each side of the membrane. Once the bacteria were added, the chambers were incubated at 37°C for 2 h and then surviving cells were quantified.

186
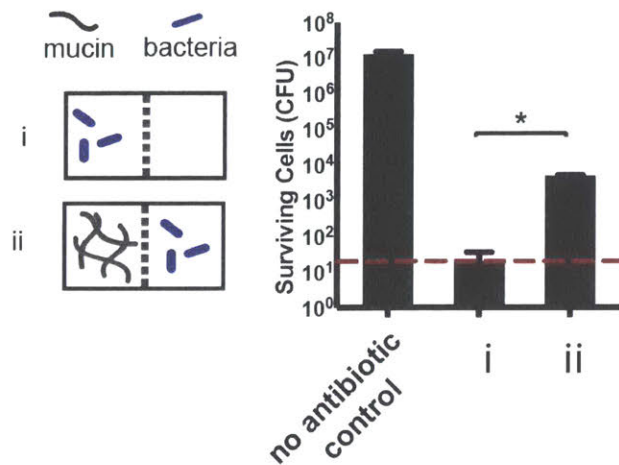
Supplemental Figure A.S2 Two-chamber antibiotic efficacy assay with colistin. Bacteria were exposed to antibiotics in two different conditions: (i) in buffer with no mucin, (ii) in buffer, with mucin in the connected chamber, separated from the cells by the membrane. In condition ii, mucin is not in direct contact with the bacteria, but still should have the ability to bind and reduce the free concentration of antibiotic available. When bacteria were separated from mucin by a membrane and exposed to colistin in buffer, an increase in surviving cells relative to the mucin-free control is observed, suggesting that mucin is able to bind and reduce the antibiotic concentration of colistin. All error bars represent standard deviation of biological replicates ($n \geq 3$). (*) indicates a significant increase in the number of surviving cells after antibiotic exposure compared to that in mucin-free buffer, as determined by the t-test ($P < 0.0001$).

Table A.S1 HPLC method. Solvent A: ultrapure water with 0.1% v/v glacial acetic acid (VWR), solvent B: HPLC grade acetonitrile (VWR) with 0.1% v/v glacial acetic acid (VWR).

Time (min)	% A	% B	Flow rate (mL/min)
0	100	0	0.3
5	100	0	0.3
20	0	100	0.3
25	0	100	0.3
25.1	100	0	0.3
35.1	100	0	0.3

Table A.S2 MS settings for antibiotics

Molecule	Precursor ion	Product ion
Polymyxin B	434.6	402
Polymyxin B	434.6	396.1
Colistin	385.9	380.1
Colistin	385.9	101.1
Levofloxacin/Ofloxacin	362.2	318.1
Levofloxacin/Ofloxacin	362.2	261.1
Ciprofloxacin	332.1	314.1
Ciprofloxacin	332.1	288.2
Ciprofloxacin	332.1	231

B Appendix B: Probing the potential of mucus permeability to signify preterm birth risk

The work presented in this appendix was published in the journal *Scientific Reports*, Volume 7 (1) on September 4, 2017. The authors were K. B. Smith-Dupont, C. E. Wagner, J. Witten, K. Conroy, H. Rudoltz, K. Pagidas, V. Snegovskikh, M. House & K. Ribbeck.[1]

My contributions: I derived and implemented the permeability metric described in Fig. B.3b and shown in Fig. B.5b, consulted on the statistics, and collaborated in the writing of the manuscript.

B.1 ABSTRACT

Preterm birth is the leading cause of neonatal mortality, and is frequently associated with intraamniotic infection hypothesized to arise from bacterial ascension across a dysfunctional cervical mucus plug. To study this dysfunction, we assessed the permeability of cervical mucus from non-pregnant ovulating (n=20) and high- (n=9) and low-risk (n=16) pregnant women to probes of varying sizes and surface chemistries. We found that the motion of negatively charged, carboxylated microspheres in mucus from pregnant patients was significantly restricted compared to ovulating patients, but not significantly different between high- and low-risk pregnant women. In contrast, charged peptide probes small enough to avoid steric interactions, but sensitive to the biochemical modifications of mucus components exhibited significantly different transport profiles through mucus from high- and low-risk patients. Thus, although both microstructural rearrangements of the components of mucus as well as biochemical modifications to their adhesiveness may alter the overall permeability of the cervical mucus plug, our findings suggest that the latter mechanism plays a dominant role in the impairment of the function of this barrier during preterm birth. We expect that these probes may be readily adapted to study the mechanisms underlying disease progression on all mucosal epithelia, including those in the mouth, lungs, and gut.

B.2 INTRODUCTION

Preterm birth, defined as birth prior to 37 weeks of gestation, affects up to 18% of pregnancies world-wide, and is the leading cause of neonatal death and the second leading cause of childhood death below the age of 5 years [2]. In addition to this immediate mortality risk, children born preterm who survive beyond infancy are predisposed to short-term complications related to the incomplete development of multiple organ systems, as well as long-term developmental disorders and early death [2,3]. Altogether, the annual financial burden of preterm birth and its related complications is estimated at $26 billion in the United States [2].

Although the causes of preterm birth are complex and multifaceted, intra-amniotic infection, which facilitates preterm birth by triggering an inflammatory response resulting in cervical ripening, weakening of the chorioamniotic membrane, and myometrial contractions [4], is present in an estimated 25–40% of clinical cases [2,3]. During healthy pregnancy, cervical mucus forms a compact, protective 'plug' between the sterile uterus and the colonized vagina, selectively permitting the passage of desirable agents such as nutrients, gases, and immunological factors, while excluding potentially deleterious environmental particles and pathogens [5–8]. Since one pathway for intra-amniotic infection is ascension of pathogens across the cervical mucus plug, changes in the permeability of this barrier have long been suspected to play a critical role in the etiology of preterm birth [5,7,8]. For instance, the presence of bacteria in amniotic fluid cultures is strongly associated with a shortened cervix, which suggests that the length of the cervical barrier is important for blocking bacterial ascension from the vagina [9]. Further, the bacteria that infect the chorioamniotic membranes in association with

preterm birth are often the same species found in the vaginal flora, suggesting that ascension occurred through the cervix[4]. Indeed, every wet epithelial surface of the body is shielded by a similar mucus barrier, and it is well established in other organ systems that numerous medical conditions including asthma [10], cystic fibrosis [10–13], ulcerative colitis[14–16], and chronic obstructive pulmonary disease [17] are intimately related to local changes in mucus properties and permeability.

The primary contributor to the gel-forming properties of mucus are mucin glycoproteins, which consist of bottlebrush-like, densely glycosylated segments and bare hydrophobic regions [5–8,18]. Mucin molecules form a viscoelastic gel capable of selectively transporting foreign molecules on the basis of size, chemical interactions, or a combination of the two [4,5,9,19–23]. Consequently, modifications to the mesh size of the network or the physicochemical properties of the constituent mucin molecules would be expected to strongly influence the permeability of mucus as well as its macroscopic mechanical response [24]. In a previous study [5,7,8], we showed that the viscoelastic moduli of cervical mucus from patients at high risk for preterm birth are significantly lower than those measured in samples from low-risk pregnancies, and that mucus from high-risk pregnancies exhibits a significantly greater degree of spinnbarkeit under extension. Here, we build on this work and interrogate whether permeability measurements at both the micro- and nanoscales in cervical mucus are altered in women at high risk for preterm birth. We note that while the design of our study makes it difficult to causally link changes in mucus permeability and early delivery, it does permit for clear associations of our measured permeability markers with preterm labor and subsequent preterm birth.

B.3 MATERIALS AND METHODS

B.3.1 Mucus sample collection

B.3.1.1 Ovulating patients

Our study included 20 ovulating non-pregnant individuals aged 18–45 years, with no restrictions based on race, ethnicity, or spoken language. Our exclusion criteria were recent sexually transmitted infection, intercourse within 24 h of collection, abnormal PAP smear within the last 6 months, or cervical surgery within the last 6 months. Patients with polycystic ovarian syndrome were excluded from the current study. Patients receiving treatment for infertility were excluded. Ovulation was detected via self-administration of a urine luteinizing hormone test kit (recommended kit: Clear Blue Easy, SPD Swiss Precision Diagnostics GmbH, Switzerland) prior to a procedure (i.e., intrauterine insemination). The number of days of the menstrual cycles did not define "ovulation": the window of "ovulation" for collecting ovulatory mucus was defined as 36 h from a positive ovulation result. After informed consent was obtained from 20 patients, specimens were collected during a sterile speculum exam with a 1-mL insulin syringe immediately before any other procedure. Specimens were gathered directly from the external cervical os after clearing the area of vaginal discharge. Vaginal discharge, if present, was removed by wiping with a large-tip swab (Scopette). Cervical mucus was not removed by this process because cervical mucus is adherent to the cervical canal. Mucus was immediately snap-frozen in liquid nitrogen and stored at −80 °C prior to analysis.

B.3.1.2 Pregnant patients

We enrolled pregnant women (18–50 years old) with singleton pregnancies at gestational ages of 24–34 weeks. We enrolled two patient groups: (i) 22 healthy, low-risk women being followed in an outpatient obstetric clinic, and (ii) 16 high-risk women admitted to the hospital following symptomatic preterm labor, defined as cervical dilation of at least 1 cm in the setting of uterine contractions. High-risk patients were invited to participate after contractions had stopped and only when no cervical exam had been performed for 48 h. Patients were not in active labor at the time of enrollment, but were in a state of arrested preterm labor. High-risk patients who gave birth after 37 weeks gestation (n = 7) as well as low-risk patients who gave birth prior to 37 weeks gestation (n = 2) were not included in the results of this study.

Pregnant patients were excluded for the following reasons: maternal medical conditions predisposing the patient to preterm delivery, placenta previa, active vaginal infection at the time of sample collection, pelvic exam or sexual intercourse within 48 h of sample collection, and the presence of labor, rupture of membranes, or active vaginal bleeding at the time of sample collection. Pregnant patients receiving intra-muscular or vaginal progesterone doses were also excluded.

Cervical mucus was sampled via sterile speculum examination. Prior to mucus collection, the area surrounding the external cervical os was cleared of vaginal discharge. If vaginal discharge was present on the cervical surface, then it was removed as for ovulating patients. A sterile catheter 3.1 mm in diameter (Aspirette Endocervical Aspirator, Cooper Surgical, Trumbull, CT, USA) was placed at the external cervical os and used to aspirate a distal specimen of cervical mucus. If cervical mucus was visible,

then a cervical cytobrush was used to collect the specimen. Mucus samples were snap-frozen in liquid nitrogen and stored at −80 °C prior to analysis.

Out of concern for the welfare of our patients, care was taken to remove the minimum amount of mucus necessary for this study. Two previous studies reported the average weights of the entire cervical mucus plug from pregnant women after term delivery to be 7.0 g (n = 15)[25] and 6.87 g (n = 18)6. Whole plugs from women who delivered preterm weighed an average of 3.8 g (n = 4) [25]. In our cohort, on average, the estimated weight of mucus collected was 134 mg (±180 mg) from low-risk pregnant patients (n = 14) and 316 mg (±207 mg) from high-risk pregnant patients (n = 9). We estimate the average fraction (by weight) of extracted mucus to be 2–8% of the entire plug.

B.3.2 Mucus sample preparation and storage

Samples were assigned sequential numbers for cataloging according to the date of collection. Researchers were not blinded to group allocation when performing each experiment and assessing the outcome, due to the visible macroscopic differences between samples in each patient category. For example, spinnbarkeit can readily be observed during pipetting of the samples or during injection of the samples into the capillaries and microfluidic devices. Additionally, nearly every cervical mucus sample from non-pregnant ovulating women was optically clear, in contrast to the mucus from pregnant women, which was typically opaque.

All samples were delivered frozen to MIT. Upon receipt, samples were divided into aliquots of various masses to reduce the number of freeze-thaw cycles. Samples were stored at −80 °C until use. Aliquots were thawed immediately prior to use. All

samples experienced no more than two freeze-thaw cycles before use, thus normalizing the effects of freezing over all samples. The microstructure of mucus appears affected beyond two freeze-thaw cycles (Fig. B.S1a), while the adhesive properties of mucus appear less sensitive to repeated freezing/thawing (Fig. B.S1b).

B.3.3 Single particle tracking (SPT) experiments and analysis

B.3.3.1 Experimental protocol

SPT specimens were prepared by combining 15 µL of cervical mucus with 0.5 µL of a solution of fluorescent, negatively charged (carboxylated) microspheres 1 µm in diameter (Magsphere, Inc., Pasadena, CA, USA) in deionized water at a dilution ratio of 1:400, resulting in an overall dilution ratio of 1:12,000 for the microspheres. If more or less than 15 µL of mucus were available, then the volume of bead solution was adjusted to accommodate a 30:1 (mucus:microsphere solution) volume ratio. All specimens were subsequently vortexed for 30 s to ensure adequate mixing, then pipetted into borosilicate square capillaries 0.9 mm × 0.9 mm × 15 mm (#8290; Vitrocom, Mountain Lakes, NJ, USA). Capillaries were sealed on both ends using a 1:1:1 mixture of petroleum jelly, lanolin, and paraffin, and then mounted onto microscope slides for imaging.

Imaging was performed at 30.3 frames per second for 10 s and at room temperature with a Zeiss Axio Observer D.1 inverted microscope using a Zeiss LD Plan-Neofluar 20x/0.4 Corr Ph2 objective lens (Carl Zeiss Microscopy GmbH, Jena, Germany) and a Hamamatsu Flash 4.0 C11440–22CU camera (Hamamatsu Photonics, Hamamatsu City, Japan).

B.3.3.2 Analysis

An average of 150 particles were imaged for each specimen from an average of eight movies recorded at different locations within the glass capillaries. For each image frame (Fig. 1a), particles were identified using publicly available MATLAB (v8.2.0.701; Natick, MA, USA) code [26,27], which identifies candidate features using high-intensity matches and filters them using criteria such as maximum feature eccentricity and radius of gyration60. The x and y position of each approved particle in each frame is recorded by the same code as the center of mass of the localized image intensity. As a result of the apparent drift in many of the trajectories, we applied a drift-correction algorithm taken from publicly available microrheology software[27], which subtracts the center-of-mass motion of all particles in a given frame from each individual trajectory. However, significant amounts of drift remained after application of this code as a result of the non-uniformity of this relaxation process, which caused displacements of various degrees throughout the sample. When we performed de-drifting over each individual particle as opposed to using the center-of-mass motion of the entire frame (Fig. B.S2 in the Supporting Material), the ballistic motion at large delay times was largely eliminated. However, this technique is not reliable at large delay times, since it imposes a net zero displacement for every particle at all delay times, as evidenced by the plateauing or sudden change in the MSD slope at large values of τ (Fig. B.S2b). Further, the MSDs at early lag times ($\tau < 0.3$s) from which the fits and statistical comparisons between patient groups were constructed were relatively unaffected by the choice of de-drifting algorithm, as confirmed by the similar fit values obtained for α and D α for all patient groups and the invariance of the conclusions regarding statistical significance (Fig. B.S3

in the Supporting Material). As a result, de-drifting using the center of mass of each frame was applied for the analysis reported in the main text of this manuscript, since it did not lead to this artifact at larger lag times.

From these drift-corrected x and y positions, the time-averaged MSD for a sequence of N images for the k th particle is [21,28]

$$\overline{r_k^2(\tau)} = \frac{1}{N-\tau/\Delta t} \sum_{i=1}^{N-\tau/\Delta t} [(x(i\Delta t + \tau) - x(i\Delta t))^2 + (y(i\Delta t + \tau) - y(i\Delta t))^2],$$

(1)

where Δt is the time between successive frames (Fig. B.1b). The ensemble average over all K particles is [29]

$$< r^2(\tau) > = \frac{1}{K} \sum_{k=1}^{K} \overline{r_k^2(\tau)}.$$

(2)

Reproducibility of SPT results was qualitatively confirmed by comparing results in individual aliquots of mucus from individual patients (Fig. B.S4 in the Supporting Material).

Statistical analysis was performed using Prism v7.0 (Graphpad Software, Inc., CA). Grubb's test for individual outliers was performed on log-transformed MSD values (at r = 0.1s) extracted from each patient group ($\alpha = 0.05$, $P < 0.05$; Fig. B.1c, starred data). Grubb's test assumes a normal distribution of data; therefore, MSD values were log-transformed. The diffusion behavior of particles in samples of mucus from two patients (Fig. B.1c, starred) was abnormally separated from the other patients in each group: the value of the generalized diffusion coefficient D_α at early lag times in the sample from the pregnant patient (Fig. B.1c, blue) was 3 orders of magnitude higher

than the average value in mucus from all other pregnant patients. The value of D_α at early lag times in the sample from the ovulating patient (Fig. B.1c, red) was more than one order of magnitude larger than the average value in mucus from all other ovulating patients. Thus, these two patients were removed from further analysis. The Mann-Whitney test for statistically significant differences between patient groups was used to accommodate non-normal distributions of values of D_α and α. Non-normality was determined using quartile-quartile plots of the data (Fig. B.S5 in the Supporting Material)

B.3.4 Peptide diffusion experiments and analysis

B.3.4.1 Permeability-reporting peptides

Peptides were synthesized, purified, and identified by the Swanson Biotechnology Center at the Koch Institute at MIT (Cambridge, MA, USA). Two peptides, AK10 ((AK)10-NH2) and AE10 ((AE)10-NH2), were synthesized via 9H-fluoren-9-ylmethoxycarbonyl (FMOC) peptide synthesis and labeled with one 6-carboxyfluorescein at the N-terminus after synthesis. Peptides were purified using reverse-phase high performance liquid chromatography. Peptide identity and purity was confirmed via matrix-assisted laser desorption ionization mass spectrometry. Before use in the microfluidic device, peptides were dissolved in H7 buffer (20 mM HEPES, 20 mM NaCl, pH 7) at a final concentration of 4 μM.

B.3.4.2 Experimental protocol

Microfluidic polydimethylsiloxane devices were designed and fabricated as previously described21, 38, and cured at 95 °C for 48–72 h prior to use. Devices bonded to microscope slides were mounted on an inverted epifluorescence microscope (IX-71,

Olympus American, Central Valley, PA; or Zeiss Observer Z1) equipped with a 5x or 10x objective and LED light source or mercury lamp (excitation at 450–495 nm or 475 nm, respectively). All channels were washed with H7 buffer before use. The two valves were closed or opened by applying or releasing pressure with a 3-mL syringe connected to the valves via polytetrafluoroethylene tubing. All experiments were carried out at room temperature. One to two microliters of sample were injected to fill the main channel. The two valves were closed while the inlet and outlet channels were washed with H7 buffer. The top valve was opened, exposing the mucus interface. Twenty-five microliters of peptide solution were added to the reservoir of each inlet. The flow rate of the peptide solution into the inlet channel was determined by gravitational flow. An image of the channels was captured every 10 s for 20 min. At least two successful experimental replicates were obtained for each patient mucus sample, except for one sample of cervical mucus, for which only one measurement was obtained.

B.3.4.3 Analysis

Images were analyzed using ImageJ (v1.47; Wayne Rasband, National Institutes of Health, Bethesda, MD, USA; http://imagej.nih.gov/ij) and MATLAB software. The zero time point was defined as the first image when the fluorescently labeled peptide passed the inlet-outlet branch point. The distance x-axis parameter origin represents a point in the buffer solution 100 μm before the estimated mucus-buffer interface. If no mucus-buffer interface existed, as in control H7 buffer experiments, then the zero distance point was assigned to the edge of the initial wave of fluorophore at 0 s. Data were excluded if: i) the mucus sample swelled out of the main channel; ii) the mucus-buffer interface became eroded or moved; or iii) the bottom valve broke during the 15-min

timeframe. Four cervical mucus samples from pregnant patients were excluded because of the first criterion. Transport curves were generated by measuring the fluorescence intensity of each pixel in a 20 pixel-wide lane along the central axis of the mucus channel. Peptide concentration was calculated from fluorescence intensity as

$$Peptide\ concentration\ (\mu M) = c(x) = (\frac{I_F - BKG}{I_{Bath} - BKG}) * 4\ \mu M,$$

(3)

where fluorescence intensity (I_F) is the average fluorescence intensity per unit length along the axis of the channel from the origin located 100 μm before the buffer-mucus interface to the end of the channel (Fig. B.2). Bath intensity (I_{Bath}) is a 10 × 10 pixel region in the washed outlet channel beyond the mucus interface, which is assumed to be equivalent to the intensity of a 4 μM solution of the peptide in buffer. The background intensity (BKG) is a 10 × 10 pixel region not occupied by a channel.

For each type of peptide, the concentration in buffer at each point, x, along the channel axis was averaged over all runs to provide a buffer concentration function $c_{buffer}(x,t)$, where x = 0 at 100 μm before the mucus-buffer interface, collected at 10-s intervals up to 15 min. Intensity profiles for each sample run gave a function $c_{sample}(x,t)$, also at 10-s intervals for 15 min. The concentration profile of peptide in buffer was subtracted from that of the same peptide type in a given mucus sample and integrated over the entire 450 μm distance (Fig. B.2b, 900 s, the shaded area between concentration profiles of positively charged peptides in mucus and buffer). These profiles were integrated over the 15-min experiment in 10-s intervals. The diffusion metric M was calculated as

$$M = \int_0^{900} \int_0^{450} |c_{sample}(x) - c_{buffer}(x)|\, dx\, dt.$$

(4)

The upper integration limit of 450 µm was chosen to be slightly lower than the minimum profile width in all the samples. Numerical integration of the discrete intensity profiles was performed using the trapezoidal rule. The final M value for a given sample and peptide was the average M over all technical replicates. The number of channels evaluated per patient varied between 2 and 17 replicates, based on the amount of available mucus and the presence of excluding behavior, like swelling, within the microfluidic device. The mucus from one patient could only be measured in singlet. Quantitative analysis was performed using MATLAB. Reproducibility was confirmed by measuring the permeability of two separate aliquots of mucus from a single patient in two separate batches of microfluidic devices (P = 0.0085; Fig. B.S6 in the Supporting Material). In this case, the metric values were summed to create a single metric.

Statistical analysis was performed using the Real Statistics Resource Pack software (Release 4.3; Charles Zaiontz, www.real-statistics.com). Metric values were log-transformed to improve normality of the distributions (Fig. B.S7 in the Supporting Material). Significance was calculated using Hotelling's T-Square test with unequal variance to accommodate negative and positive metrics.

B.3.5 Study approval

The study protocol was approved by the Institutional Review Boards of Tufts Medical Center (#9355), the Women and Infants' Hospital of Rhode Island (WIH 15–0073), and The Massachusetts Institute of Technology (#1501006840R001). Written informed

consent was obtained from all participants prior to enrollment in the study. All experiments were performed in accordance with the relevant guidelines and regulations.

B.4 RESULTS AND DISCUSSION

We employed two classes of probes to dissect the selective permeability of mucus samples: negatively charged carboxylated microspheres 1 μm in diameter, and charged positive and negative peptides <10 nm in size. Negatively charged (carboxylated) particles were selected for the larger probes as a result of previous findings of charge-mediated impaired diffusion for positively charged (amine functionalized) particles in mucus and mucin gels [10,13,21,30]. Additionally, although further surface functionalization such as PEGylation has been shown to reduce particle-mucus interactions to an even greater degree than carboxylation alone [10-13,31-33], this effect for particles >500 nm in diameter was found to be minimal [13,31]. Although we certainly cannot altogether rule out the possibility of mucus-particle interactions, our chosen particle length scale of 1 μm (relevant for many bacteria, sperm cells, and, as noted by Hill et al. [13], large drug-delivery vehicles) is comparable to the upper reported limit of the characteristic mesh size of native cervical mucus[34-36]. As such, steric interactions with the surrounding mucin network should be the dominant mechanism for impaired diffusive motion at this length scale. Correspondingly, we hypothesized that the diffusion of the larger probes should be altered primarily by structural reorganization of the surrounding network, while the diffusion of the smaller probes, which are minimally affected by sterics, should allow us to interrogate molecular interactions between the peptide analytes and the mucus gel.

To assess the motion of the microspheres in mucus, we used single-particle

tracking (SPT), a microrheological technique for examining the biophysical properties of

complex fluids through examination of the passive motion of randomly embedded
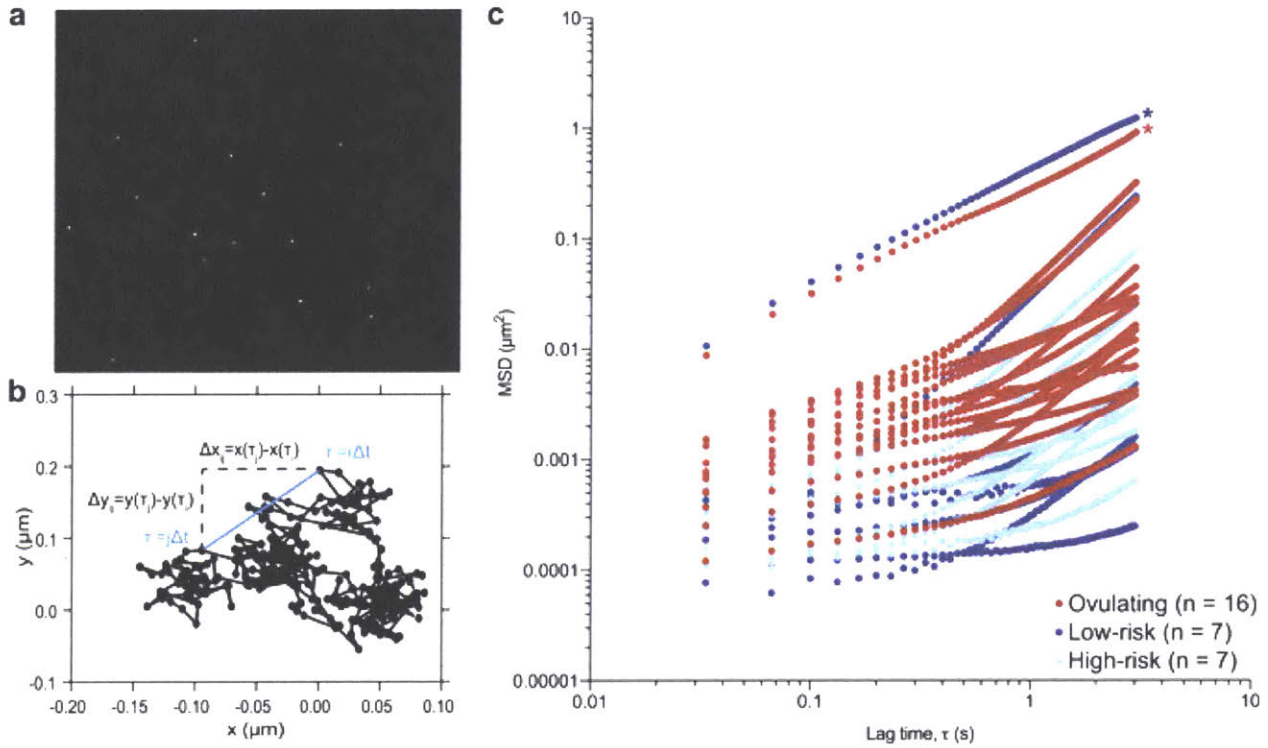
Figure B.1 SPT of cervical mucus reveals significant differences in particle mobility between pregnant and ovulating patients but not between low- and high-risk pregnant patients. (**a**) A portion of the field of view of a single frame from a typical movie of mucus containing randomly dispersed 1 μM fluorescent particles. (**b**) Sample trajectory for a single particle from (**a**). Dots denote the x and y positions of the particle recorded by the feature-finding software at each of the 300 time points separated by time Δt. The method of calculation of a single squared displacement Δr_{ij}^2 at a lag time of $\tau_{ij} = (j - i)\Delta t$ is shown, where $\Delta r_{ij}^2 = \Delta x_{ij}^2 + \Delta y_{ij}^2$. The MSD at a given lag time for one specimen is obtained by averaging this individual displacement over those from all equivalent lag times within the trajectory of a particular particle, and then over all of the tracked particles. (**c**) MSD results as a function of lag time for all cervical mucus specimens. Dark blue symbols, low-risk pregnant patients; light blue symbols, high-risk pregnant patients; red symbols, ovulating patients. The patients with the overall highest MSDs (**c**, starred) are suspected outliers from their respective patient groups (Methods).

particles (Fig. B.1a and b). This technique is routinely used in the design of drug-delivery vehicles, which must interact minimally with the various mucosal layers of the body in order to reach their target tissues prior to being cleared [32,37]. The relationships that can be drawn between the thermal fluctuations of these particles and the linear viscoelastic properties of the surrounding medium [29,38] have led to regular use of this technique in the assessment of the mechanical properties of biological fluids including mucin-based gels, such as the effect of pH on the mesh size of native cervical mucus [39], the effect of pH and salt on the microrheology of reconstituted MUC5AC gels[24], and the effect of elevated mucin concentration on the elasticity of diseased airway sputum[13]. In this work we show that SPT in cervical mucus can be used to distinguish between pregnant and non-pregnant (ovulating) women, but not between women at high and low risk for preterm birth.

For normal diffusion, as occurs in a homogeneous medium with no memory (i.e. no elastic character) and with which the microspheres do not interact, the mean squared displacement (MSD) of a particle is expected to scale linearly with the lag time, and in two dimensions the explicit form for this scaling is [28]

$$\langle r^2(\tau)\rangle = 4D\tau, \tag{5}$$

where D is the diffusion coefficient of the microsphere in the medium. When this scaling does not hold, the diffusion is termed anomalous, and the MSD is often expressed in power-law form as

$$\langle r^2(\tau)\rangle = 4D_\alpha\tau^\alpha, \tag{6}$$

where D_α is a generalized diffusion coefficient and the exponent α can be lag-time dependent. When $\alpha < 1$, the motion of the particle is subdiffusive, and when $\alpha > 1$, the

motion is superdiffusive. Anomalous diffusion has been reported in biological fluids such as lung mucus [13], F-actin gels [40], and plasma membranes [41], reflecting the complex microstructure that these media generally possess.

Table B.1 Patient characteristics associated with samples of cervical mucus used for single particle tracking. P values were calculated using a two-tailed Student's t-test. Values represented as mean (± standard deviation) or percent (%) where applicable

| Characteristics | Non-pregnant (n = 16) | Pregnant (n = 14) | | |
	Ovulatory (n = 16)	Low risk (n = 7)	High risk (n = 7)	p-value
Age (years)	35.13 (±3.44)	28.00 (±4.97)	26.83 (±6.91)	
Race (%)				
White	100.00	28.57	57.14	
Black	0.00	0.00	14.29	
Hispanic	0.00	57.14	14.29	
Other	0.00	14.29	14.29	
History (%)				
Nullipara	68.75	0.00	0.00	
Primipara	0.00	57.14	71.43	
Multipara	31.25	42.86	28.57	
Prior PTB (%)	0.00	0.00	0.00	
Gestational Age (weeks)				
Collection	NA	31.12 (±2.36)	29.84 (±2.68)	0.36
Delivery	NA	39.94 (±0.45)	34.59 (±2.12)	4.29E-04
Positive GBS (%)	NA	28.57	28.57	
Dilation (cm)	NA	0.0 (±0.0)	3.1 (±1.4)	8.24E-04

We measured the diffusion behavior of particles in cervical mucus from 16 ovulating patients, 7 low-risk pregnant patients, and 7 high-risk pregnant patients (Table B.1). The low- and high-risk pregnant patients were gestational age-matched at the time of mucus collection (P = 0.36, Student's t-test), but the high-risk patients gave birth

before 37 weeks of gestation (P = 4.29E-4, Student's t-test). To quantify our SPT results, we fit the ensemble average MSDs as a function of lag (or delay) time τ in all of the cervical mucus samples (Fig. B.1c) with power laws of the form in Eq. B.6 for τ ≤ 0.3s. These early lag times were selected for fitting as a result of the observation of drift in many of the cervical mucus samples, as evidenced by the near quadratic
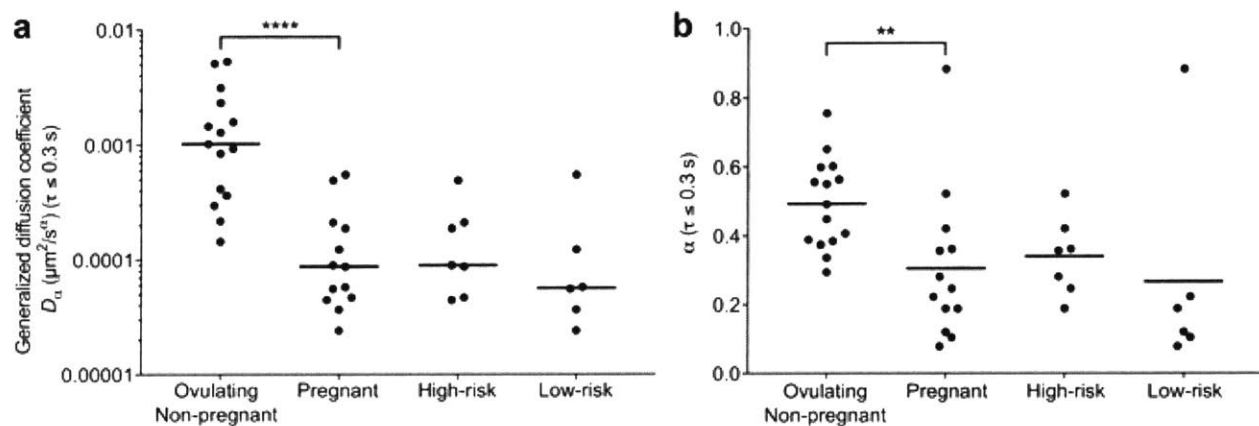
Figure B.2 The generalized diffusion coefficient D_α (a) and the exponent α (b) extracted from the MSD results at early lag times (τ ≤ 0.3 s) are significantly lower in pregnant patients compared to ovulating ones, implying that the micron-sized probe particles are significantly less mobile in these samples. Pregnant patient data are further subdivided into high- and low-risk groups, but significant differences between these same two parameters are not observed at this level. Each point represents a single patient sample from non-pregnant ovulating or pregnant patients. Bars, median values for each patient group. Significance was determined with the Mann-Whitney test (****$P < 0.0001$; **$P = 0.0021$).

dependence of the MSD on τ at delay times above 1 s (Fig. B.1c). This drift appears to be the result of relaxation within the mucus samples; detailed discussion of our drift-correction procedures appears in the Methods. Further, we identified two outliers (starred, Fig. B.1c), one in the ovulating group and one in the pregnant patient group, which we excluded from subsequent statistical analysis (Methods).

The mobility of the microspheres in mucus from ovulating patients (Fig. B.1c, red) was notably greater than that in mucus from pregnant patients (Fig. B.1c, blue). At early lag times, both the MSD exponent α (P = 2.1E-3, Mann-Whitney test) and the generalized diffusion coefficient D_α (P < 1.0E-4, Mann-Whitney test) were significantly higher in ovulating patients (Fig. B.2a and b), with α = 0.492 and D_α = 1.022 × $10^{-3} \mu m^2/s^\alpha$ the median values for ovulatory mucus and α = 0.247 and D_α = 8.698 × $10^{-5} \mu m^2/s^\alpha$ the median values for pregnancy mucus. In addition, at early delay times α was greater in samples from high-risk than low-risk patients (α = 0.356 and α = 0.154, respectively), although the difference was not statistically significant for this limited sample size (P = 7.98E-2).

The lower value of α at early lag times in pregnancy mucus compared to ovulatory mucus, and low-risk mucus compared to high-risk mucus, corresponds to more confined particle trajectories [29]. This confinement may arise from a possible combination of increased viscoelasticity and/or increased periods in which the probes are locally caged by their surroundings [29]. These results in ovulatory mucus are consistent with previous observations of marked structural reorganization in cervical mucus over the course of the menstrual cycle, which appear to maximize the permeability of the mucus to the passage of sperm during ovulation [42]. Taken together, these data reveal a significant association between pregnancy state and the physical microstructure of cervical mucus. However, the statistical ambiguity in the stratification of low- and high-risk samples using micron-sized probes and SPT suggests that structural reconfiguration alone cannot completely account for the changes in the permeability and viscoelasticity of cervical mucus from patients at high risk for preterm

birth. As such, we next assessed the impact of mucosal function and pregnancy health on the biochemical properties of all of the samples of cervical mucus using peptide analytes at the greatly reduced length scale of 10 nm.

Charge is one parameter commonly considered to influence transport through mucus, since the interactions of a particle with surrounding matrix components are mediated by surface-to-surface contacts. Consequently, we generated two short

a

(+) Positively charged peptide, (AK)$_{10}$
FAM-AKAKAKAKAKAKAKAKAKAK-NH$_2$
(-) Negatively charged peptide, (AE)$_{10}$
FAM-AEAEAEAEAEAEAEAEAEAE-NH$_2$

b
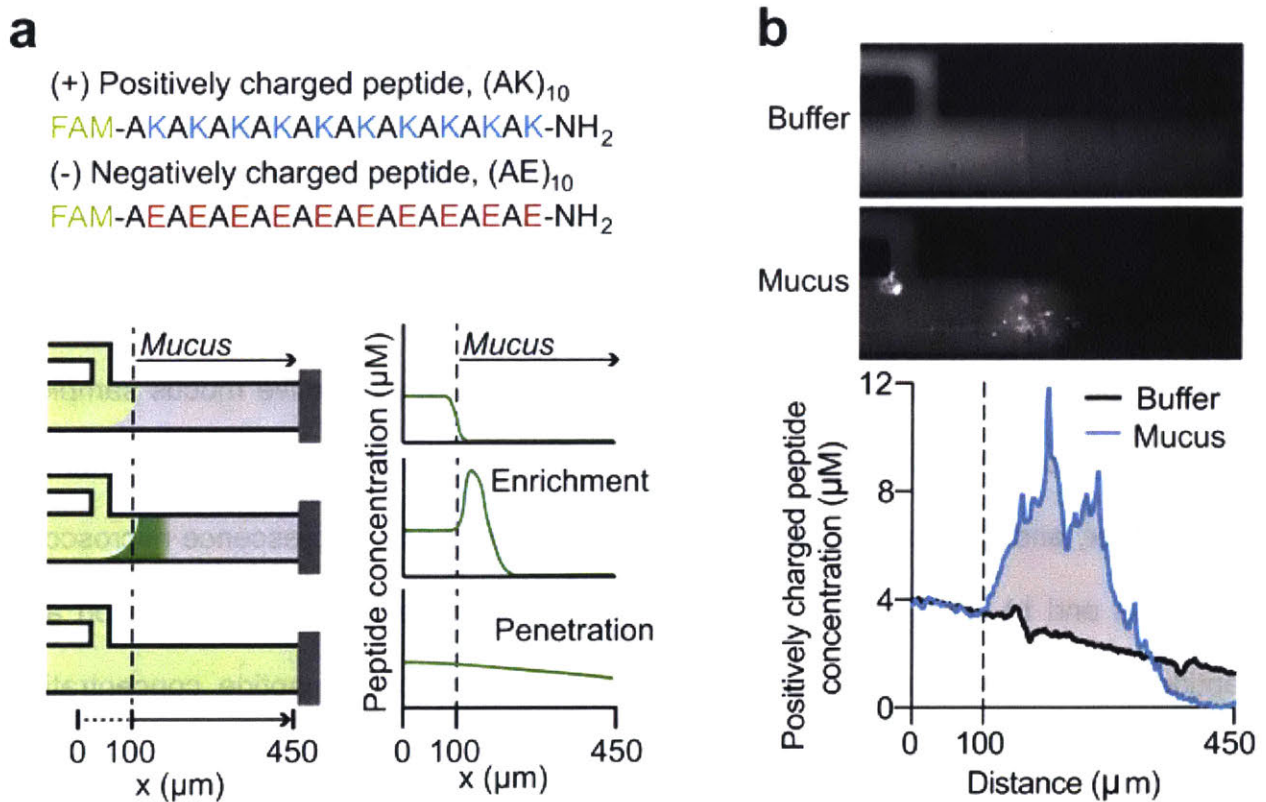
Figure B.3 Reporter peptide sequence, microfluidic design, and quantitative analysis of mucus permeability. (**a**) Positively (+) and negatively (−) charged peptide sequences. Schematic diagrams of the microfluidic device illustrate two phenomena observed in the diffusion behavior of peptides through mucus, and the resulting concentration profiles from those behaviors. In all concentration profiles and schematics, the vertical, dotted line indicates the buffer-mucus interface. (**b**) Images obtained with an epifluorescence microscope are used to calculate the concentration of peptide as a function of distance, with the mucus buffer interface located 100 μm from the origin. The images and concentration profile shown are of the positively charged peptide, 900 s after the beginning of the experiment. The gray, shaded area between the buffer and mucus curves is a graphical representation of the metric (Methods).

peptides with uniform charge: the positively charged "+" peptide consists of 10 lysine residues, each separated by one alanine residue, while the negatively charged "−" peptide consists of glutamic acid residues separated by alanines (Fig. B.3a). All peptides were covalently linked to the fluorescent dye FAM (6-carboxyfluorescein). This fluorescent tag is small, singly charged, and is not expected to affect transport into or through mucus.

We quantified partitioning and transport of the reporter peptides through mucus using a previously described microfluidics device [22,43]. With this setup, a continuous aqueous flow of fluorescent peptides is generated within a micron-scale channel at carefully controlled rates adjacent to a stable mucus layer. The small sample volume (5 μL) required for this system is ideally suited for the study of native mucus samples, which are typically not available in large volumes. The peptide solutes partition into the mucus layer, and their subsequent transport is monitored with fluorescence microscopy (Fig. B.3a and b). The rate of entry into the mucus barrier, as well as diffusion and peptide-mucin binding interactions within the mucus, result in peptide concentration profiles that can be quantified.

We quantified the transport of (+) and (−) charged peptides in mucus from 20 ovulating, 15 low-risk pregnant, and 9 high-risk pregnant patients (Table B.2; see the table caption for specific information on cohort composition). As in the SPT experiments, all pregnant patients were gestational-age matched at the time of collection (P = 0.07, Student's t-test); however, high-risk patients gave birth preterm (P = 1.31E-4, Student's t-test). Peptides at a concentration of 4 μM were introduced into

the microfluidic channel in the presence of the mucus sample or in the presence of buffer. At this concentration, the peptides do not significantly affect the ionic environment of the mucus layer. Concentration profiles of the fluorescent peptides were acquired through the center cross - section of the mucus channel (Fig. B.3a,b, Methods). Representative examples of fluorescence images of channels filled with

Table B.2 Patient characteristics associated with samples of cervical mucus used for single particle tracking. *P* values were calculated using a two-tailed Student's t-test.

| Characteristics	Non-pregnant (n = 20)	Pregnant (n = 24)		p-value
	Ovulatory (n = 20)	Low risk (n = 15)	High risk (n = 9)	
Age (years)	35.40 ($\pm$3.69)	27.80 ($\pm$5.47)	24.67 ($\pm$6.40)	
Race (%)				
White	100.00	60.00	44.44	
Black	0.00	0.00	11.11	
Hispanic	0.00	33.33	33.33	
Other	0.00	6.67	11.11	
HIstory (%)				
Nullipara	70.00	0.00	0.00	
Primipara	0.00	53.33	77.78	
Multipara	30.00	46.67	22.22	
Prior PTB	0.00	0.00	0.00	
Gestational Age (weeks)				
Collection	NA	31.66 ($\pm$1.88)	29.43 ($\pm$3.08)	0.07
Delivery	NA	39.96 ($\pm$0.60)	33.59 ($\pm$2.85)	1.31E-04
Positive GBS (%)	NA	26.67	22.22	
Dilation (cm)	NA	0.0 ($\pm$0.0)	3.1 ($\pm$1.2)	5.54E-05
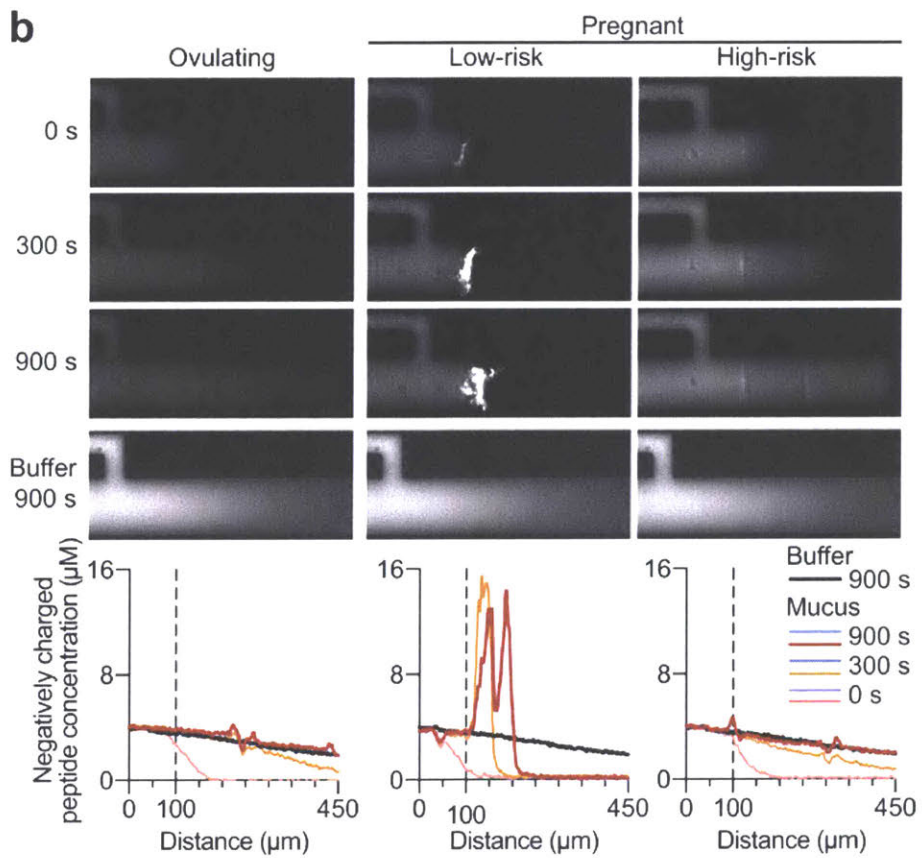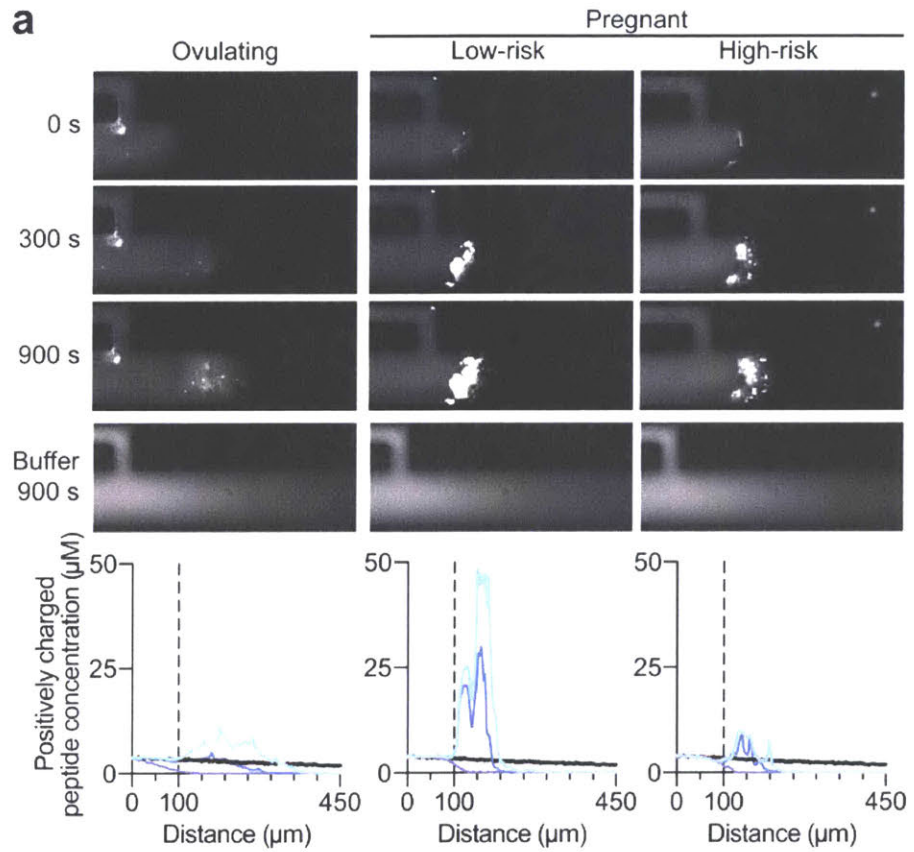
Figure B.4 Peptide diffusion through cervical mucus reveals differences in permeability and mucoadhesion between ovulating non-pregnant, low-risk, and high-risk pregnant patients. Microscope images and respective concentration profiles from a representative single patient from each of the three patient groups are shown. (**a**) Positively charged peptides (top panel, blue) enrich substantially more at the interface of mucus from pregnant patients in contrast to ovulating non-pregnant patients, indicating increased mucoadhesion. However, this enrichment is decreased in mucus from high-risk patients. (**b**) Negatively charged peptides (bottom panel, red) diffuse freely through mucus from ovulating and high-risk pregnant patients, but not through low-risk patient mucus, indicating increased adhesiveness of mucus from healthy pregnant patients.

mucus from three patients, recorded 0 s, 300 s, and 900 s after introduction of the (+) peptide and (−) peptide, appear in Fig. B.4a and b, respectively. The corresponding profiles quantifiying the peptide concentration as a function of distance through the channel are shown below each set of images. We detected a peak in the (+) peptide concentration profile at the mucus-buffer interface for all sample categories (Figs B.4a and B.5a). This peak is indicative of enrichment, which is essentially an increase in local concentration of peptide at the interface of the mucus barrier. This enrichment is likely due to upward partitioning consistent with electrostatic interactions with the negatively charged mucins, as well as possible direct binding interactions between the peptides and mucins that enhance the local peptide concentration. Even though all sample categories showed enrichment, the degree to which they were enriched differed substantially. The average maximum concentration of the (+) peptide that accumulated inside mucus from ovulating women was 10.6 (SD ± 9.7) μM (Fig. B.5a). For comparison, the same peptide accumulated to substantially higher levels in mucus from women at low risk for preterm birth, where the maximum concentration reached 36.9 (SD ± 20.9) μM (Fig. B.5a). Samples from high-risk pregnant patients had levels of enrichment (19.3 (SD ± 13.2) μM) that were more typical of ovulating women (Fig. B.5a).

a

Pregnant (n = 24)

Ovulating (n = 20)   Low-risk (n = 15)   High-risk (n = 9)

Positively charged peptide concentration (μM)

Distance (μm)

Negatively charged peptide concentration (μM)

Distance (μm)

b

Log (Negative metric)

Log (Positive metric)

Buffer   Mucus
— 900 s   — 900 s
          — 600 s
          — 300 s
          — 60 s

□ Ovulating (n = 20)
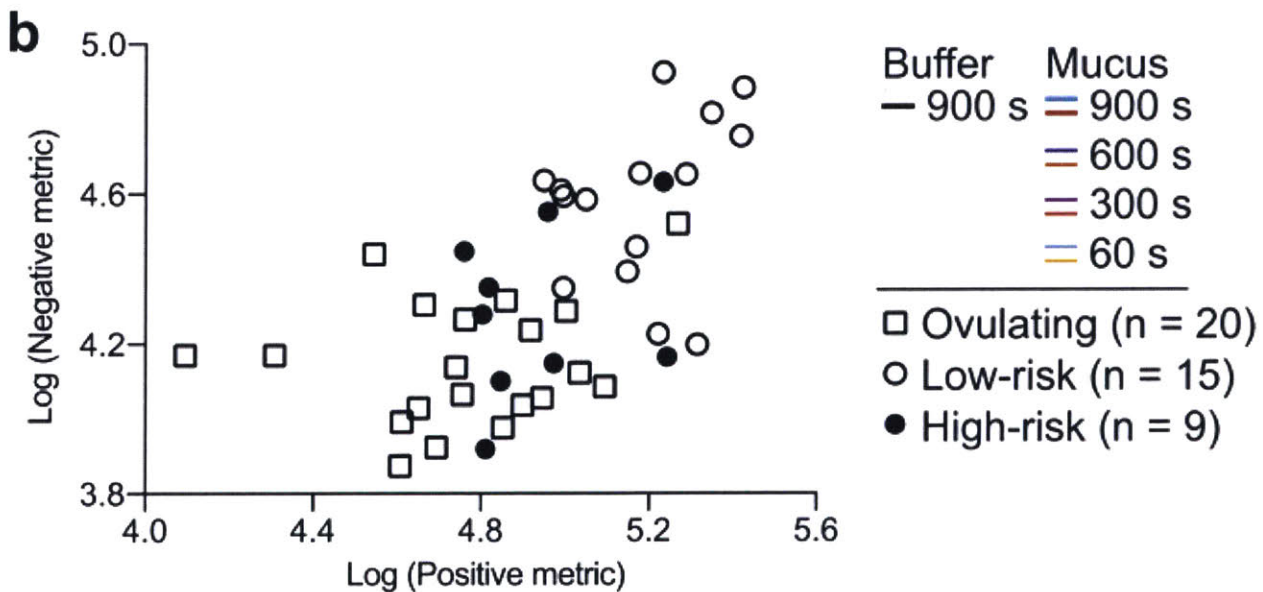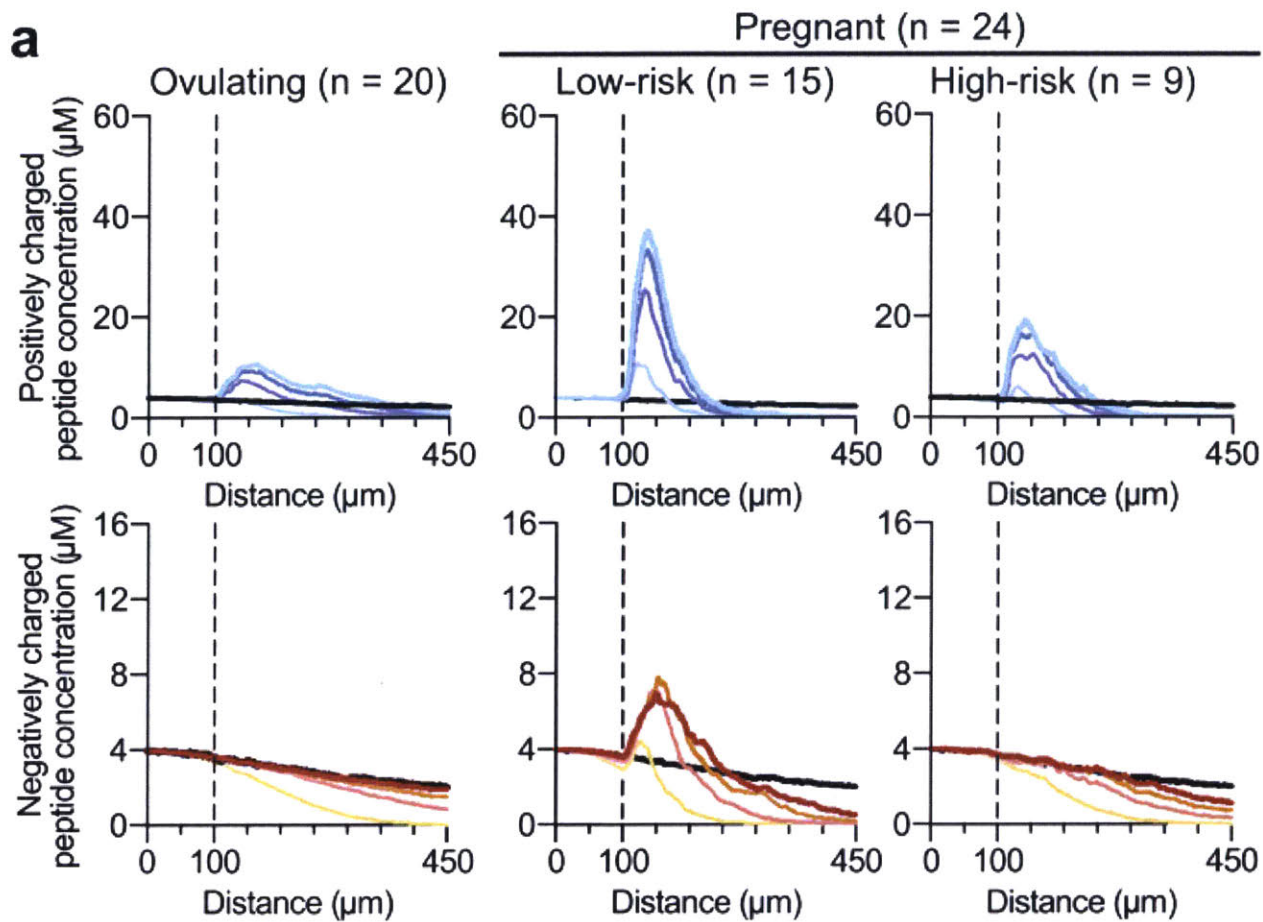○ Low-risk (n = 15)
● High-risk (n = 9)

Figure B.5 Peptide diffusion through cervical mucus enables stratification of patients at high and low risk for preterm birth. (a) Average concentration profiles of positively charged peptides (top panel, blue) show significantly increased adhesiveness in pregnant patients; however, the enrichment of positively charged peptides in mucus from high-risk patients is significantly lower compared to that of low-risk patients. Negatively charged peptide concentration profiles (bottom panel, red) differentiate between high- and low-risk patients, indicated by free diffusion through mucus from ovulating and high-risk patients, but not through mucus from low-risk patients. In all concentration profiles, the vertical, dotted line indicates the buffer-mucus interface. (b) Scatter plot of log-transformed negative and positive peptide permeability metric values reveals that high-risk patient mucus is more permeable to peptide probes. Metric values represent the difference in peptide transport behavior between mucus and buffer, integrated over the length of the channel and duration of the experiment.

For the (−) peptide, in ovulating women (Figs B.4b and B.5a), a decreasing peptide concentration gradient through the mucus layer was observed; this gradient was nearly identical with or without mucus present (Fig. B.5a), demonstrating a lack of interactions between mucus and the (−) peptide. For comparison, in mucus from low-risk pregnant patients, we consistently found a slower evolution of the flow profile, resulting in substantially lower concentrations (0.50 (SD ± 0.47) μM) of the peptide at the end of the channel (450 μm) at equivalent measurement times (Fig. B.5a). This effect could be related to reduced diffusivity or downward partitioning of the (−) peptide in the mucus. This phenomenon did not occur in mucus from high-risk pregnant patients (1.10 (SD ± 0.53) μM), where the peptide appeared to diffuse through mucus nearly as effectively as in ovulating women (Fig. B.5a). In addition, in several, but not all, low-risk pregnant individuals, we observed weak enrichment of the negative peptide at the mucus barrier interface, reflecting peptide retention by the mucus components. Retention of the peptide at the interface was not detected in any of the high-risk samples measured (Fig. B.5a). Two conclusions can be drawn from this experiment. First, the microscopic permeability of mucus from ovulating, pregnant, and high-risk

women differ significantly and can be assessed with short, charged peptides, with each charge revealing a unique binding profile characteristic of a specific mucus category. Second, the more similar permeability behavior toward both (−) and (+) peptides of mucus from high-risk pregnant patients and ovulating patients versus their low-risk pregnant counterparts strongly suggests that the barrier properties of cervical mucus are compromised in cases of elevated risk of preterm birth.

Having seen the qualitative differences in peptide transport between subject groups (Figs B.4a,b, and B.5a), we sought to test the statistical significance of these results. A quantitative metric was devised for measuring the difference in permeability of each peptide by integrating the difference between the concentration profile in buffer and in mucus over space and time. Briefly, the metric is the difference in peptide concentration between a channel filled with buffer and a channel filled with mucus, integrated over the entire length of the channel in every image taken every 10 s for the duration of the 15-min experiment (Methods, B.4; Fig. B.3b, integrated area between the concentration curve of peptide in mucus and buffer, highlighted in gray). Integration over the entire experiment accounts for the integrated area calculated from each concentration curve measured at each time point. Log-transformed metrics for positive and negative peptides in each patient mucus sample are summarized as a scatter plot in Fig. 5b. The increased enrichment and decreased penetration depth of both peptide types in mucus from low-risk patients, for example, results in higher metric values (Figs B.4a and B.5b, Methods). Conversely, peptide diffusion behavior in mucus that mimics that observed in buffer yields a lower metric value (Figs B.4a and B.5b, Methods). We performed Hotelling's T-square tests to determine the statistical significance of the

permeability differences among the three patient groups. Considering both peptides in the test, the permeability of mucus from pregnant patients was significantly reduced in comparison to ovulating non-pregnant patients (P = 7.1E-6, unequal variance Hotelling's T-square test). Further, mucus from high-risk patients was significantly more permeable to both peptides than mucus from low-risk patients (P = 5.3E-3, unequal variance Hotelling's T-square test). The metrics calculated for mucus from low-risk individuals clustered at higher values (Fig. B.5b). These data confirmed the hypothesis that charge-selective mucus permeability as measured by a set of judiciously designed peptides is a robust indicator of altered physiology of the mucus barrier.

Our permeability assays and SPT experiments revealed that cervical mucus from our three patient groups restricted the passage of charged, micron-sized particles and nanoscale peptides in significantly different ways. At the microscopic length scale where steric interactions dictate permeability, probe particles were significantly more mobile in ovulatory mucus than samples from pregnant patients. Although on average particles were more mobile in cervical mucus from high-risk than low-risk patients, these results are not significant (Fig. B.1c). The absence of statistical significance between the high- and low-risk patient groups using this technique suggests that structural reorganization alone cannot fully account for the significant rheological differences measured previously between these two patient groups[8], and that biochemical modifications to the mucins may contribute to these differences. Indeed, we found that cervical mucus from low-risk patients was significantly more effective at excluding both negatively and positively charged nanoscale peptides (Figs B.4b and B.5a), presumably due to

217

repulsive electrostatic interactions in the case of the former and electrostatically mediated binding to mucus polyanions in the case of the latter.

This study highlights that cervical mucus from pregnant women at risk for preterm birth is more permeable than that from women with healthy pregnancies. The intimate relationship between hydrogel permeability and network physicochemical properties renders this finding clinically relevant, as it suggests that risk for preterm birth is associated with altered retention of host immunological molecules and interaction of microbes with the cervical mucus barrier. In previous work we reported that mucus from high-risk patients is more translucent and has a higher spinnbarkeit than samples from low-risk patients, leading to the hypothesis that cervical mucus in high-risk patients fails to develop into the thickened and impermeable barrier typically present during healthy pregnancies[7]. The results of the current investigation support this hypothesis and additionally show that mucus from high-risk patients is more "leaky" than cervical mucus in healthy pregnancy (Fig. B.5a and b).

Our finding of significantly increased mobility of nanoscale peptides, (Fig. B.5a and b) but not micron-sized tracer beads (Fig. B.2a and b) in cervical mucus from high-risk patients suggests that this increased permeability is primarily attributable to altered adhesiveness of the components of the mucus barrier. In previous work, we showed that mucin polymers alone can regulate the diffusion of small peptides and even protons [22,43], and hence we speculate that altered mucin biochemistry can impact overall mucus permeability. As a result, it can be expected that the regulatory role of the cervical mucus barrier in selectively restricting or permitting the passage of particles and microbes may be compromised through alterations to the biochemical and/or adhesive

properties of the constituent mucin molecules. Mucins interact with and potentially regulate the activity of a plethora of non-mucin proteins in mucus, particularly ones involved in innate host defense [44]. For instance, our data predict that host antimicrobial peptides, which are generally highly cationic and contribute antimicrobial properties to the cervical mucus plug [6,45], would be less well retained in high-risk cervical mucus that is more permeable to positively charged peptides, thus making infection more common in this patient group. Modified mucoadhesive and biochemical properties of mucin molecules may also affect the movement of viral particles in cervical mucus29, and likely also regulate the physiology of microbes [46–48].

The altered mucoadhesive properties of the cervical mucus in high-risk women may be due to modification of mucin glycosylation patterns. One possible mechanism for this may be the pathogenic cleavage of sialic acid residues from mucin glycans, which could simultaneously alter the charge structure of the mucin molecules in patients at higher risk for bacterial infection as well as increase the flexibility of the mucin chains. This is consistent with our observations of reduced positively charged peptide interactions and the elevated spinnbarkeit measured in high-risk samples [8]. Additionally, bacterial vaginosis, a condition in which the normal vaginal flora are dominated by the pathogens *Gardnerella vaginalis* and *Mycoplasma hominis*, is associated with increased mucinase and sialidase activity in cervico-vaginal mucus secretions [49–51] and significantly associated with intra-amniotic infection and incidence of preterm birth [19,52–55]. While it remains uncertain whether the changes observed in cervical mucus are a cause or a consequence of preterm labor, our findings clearly demonstrate that the biophysical and biochemical properties of cervical mucus from women who deliver

preterm are significantly different from those who carry to term. Future longitudinal studies will aim to elucidate the specific physicochemical modifications to the mucin molecules that underlie these effects as they occur throughout pregnancy.

At present, there is a critical clinical need for predictive diagnostic tools and biomarkers for preterm birth. Current routine diagnostic risk screenings include measurement of cervical length, and although the correlation between a short cervix and preterm birth is significant, the positive predictive value of a short cervix for preterm delivery is low [9,56–61]. Additional known risk factors include the patient's history of preterm birth and increased fetal fibronectin levels [61–63], but many preterm births occur in women without risk factors, and many women who are identified as high-risk deliver at term, highlighting the need for reliable diagnostic biomarkers. Our finding that the permeability of cervical mucus is correlated with pregnancy risk may be exploited as a powerful diagnostic resource because of the accessibility of this biological material. Our hope is that mucus permeability could be developed into a routine test to assess preterm birth risk, and help establish therapeutic interventions for high-risk women, such as a synthetic mucus mimetic that could curb bacterial ascension by reproducing the healthy physicochemical properties of the cervical mucus barrier that are lacking in patients at high risk for preterm birth.

To conclude, the results from this study in combination with our previous rheological findings [8] establish the potential of cervical mucus as a sensitive diagnostic tool for disease. These findings may help in the development of improved, targeted treatments by spotlighting the particular importance of reinforcing the cervical mucus plug in cases of compromised barrier function. More broadly, we expect that the

identification of general biophysical patterns and the underlying biochemical changes that influence and modulate the selective barrier properties of mucus will be essential for improving our mechanistic understanding of not only preterm birth but mucosal diseases as a whole, whether they affect the cervix, lungs, gastrointestinal tract, or any surface of the body where mucus is found.

B.5 REFERENCES

1. Smith-Dupont, K. B. *et al.* Probing the potential of mucus permeability to signify preterm birth risk. *Sci. Rep.* **7**, 10302 (2017).
2. Romero, R., Dey, S. K. & Fisher, S. J. Preterm labor: one syndrome, many causes. *Science* **345**, 760–765 (2014).
3. Goldenberg, R. L., Culhane, J. F., Iams, J. D. & Romero, R. Epidemiology and causes of preterm birth. *Lancet Lond. Engl.* **371**, 75–84 (2008).
4. Goldenberg, R. L., Hauth, J. C. & Andrews, W. W. Intrauterine infection and preterm delivery. *N. Engl. J. Med.* **342**, 1500–1507 (2000).
5. Becher, N., Adams Waldorf, K., Hein, M. & Uldbjerg, N. The cervical mucus plug: structured review of the literature. *Acta Obstet. Gynecol. Scand.* **88**, 502–513 (2009).
6. Hein, M., Valore, E. V., Helmig, R. B., Uldbjerg, N. & Ganz, T. Antimicrobial factors in the cervical mucus plug. *Am. J. Obstet. Gynecol.* **187**, 137–144 (2002).
7. Bastholm, S. K., Becher, N., Stubbe, P. R., Chronakis, I. S. & Uldbjerg, N. The viscoelastic properties of the cervical mucus plug. *Acta Obstet. Gynecol. Scand.* **93**, 201–208 (2014).
8. Critchfield, A. S. *et al.* Cervical mucus properties stratify risk for preterm birth. *PloS One* **8**, e69528 (2013).
9. Gomez, R. *et al.* A short cervix in women with preterm labor and intact membranes: A risk factor for microbial invasion of the amniotic cavity. *Am. J. Obstet. Gynecol.* **192**, 678–689 (2005).
10. Rogers, D. F. Airway mucus hypersecretion in asthma: an undervalued pathology? *Curr. Opin. Pharmacol.* **4**, 241–250 (2004).
11. Yuan, S. *et al.* Oxidation increases mucin polymer cross-links to stiffen airway mucus gels. *Sci. Transl. Med.* **7**, 276ra27-276ra27 (2015).
12. Fahy, J. V. & Dickey, B. F. Airway Mucus Function and Dysfunction. *N. Engl. J. Med.* **363**, 2233–2247 (2010).
13. Hill, D. B. *et al.* A Biophysical Basis for Mucus Solids Concentration as a Candidate Biomarker for Airways Disease. *PLOS ONE* **9**, e87681 (2014).
14. Johansson, M. E. V. *et al.* Bacteria penetrate the normally impenetrable inner colon mucus layer in both murine colitis models and patients with ulcerative colitis. *Gut* **63**, 281–291 (2014).

15. Younan, F., Pearson, J., Allen, A. & Venables, C. Changes in the structure of the mucous gel on the mucosal surface of the stomach in association with peptic ulcer disease. *Gastroenterology* **82**, 827–831 (1982).

16. Atuma, C., Strugala, V., Allen, A. & Holm, L. The adherent gastrointestinal mucus gel layer: thickness and physical state in vivo. *Am. J. Physiol. Gastrointest. Liver Physiol.* **280**, G922-929 (2001).

17. Martín, R., Sánchez, B., Suárez, J. E. & Urdaci, M. C. Characterization of the adherence properties of human Lactobacilli strains to be used as vaginal probiotics. *FEMS Microbiol. Lett.* **328**, 166–173 (2012).

18. Bansil, R. & Turner, B. S. Mucin structure, aggregation, physiological functions and biomedical applications. *Curr. Opin. Colloid Interface Sci.* **11**, 164–170 (2006).

19. Gonçalves, L. F., Chaiworapongsa, T. & Romero, R. Intrauterine infection and prematurity. *Ment. Retard. Dev. Disabil. Res. Rev.* **8**, 3–13 (2002).

20. Lieleg, O. & Ribbeck, K. Biological hydrogels as selective diffusion barriers. *Trends Cell Biol.* **21**, 543–551 (2011).

21. Lieleg, O., Vladescu, I. & Ribbeck, K. Characterization of particle translocation through mucin hydrogels. *Biophys. J.* **98**, 1782–1789 (2010).

22. Li, L. D. *et al.* Spatial configuration and composition of charge modulates transport into a mucin hydrogel barrier. *Biophys. J.* **105**, 1357–1365 (2013).

23. Schneider, C. S. *et al.* Nanoparticles that do not adhere to mucus provide uniform and long-lasting drug delivery to airways following inhalation. *Sci. Adv.* **3**, e1601556 (2017).

24. Celli, J. P. *et al.* Rheology of gastric mucin exhibits a pH-dependent sol- gel transition. *Biomacromolecules* **8**, 1580–1586 (2007).

25. Becher, N., Hein, M., Danielsen, C. C. & Uldbjerg, N. Matrix metalloproteinases in the cervical mucus plug in relation to gestational age, plug compartment, and preterm labor. *Reprod. Biol. Endocrinol. RBE* **8**, 113 (2010).

26. Pelletier, V., Gal, N., Fournier, P. & Kilfoil, M. L. Microrheology of microtubule solutions and actin-microtubule composite networks. *Phys. Rev. Lett.* **102**, 188303 (2009).

27. Crocker, J. C. & Weeks, E. R. Particle tracking using IDL. (2016). Available at: www.physics.emory.edu.

28. Metzler, R., Jeon, J.-H., Cherstvy, A. G. & Barkai, E. Anomalous diffusion models and their properties: non-stationarity, non-ergodicity, and ageing at the centenary of single particle tracking. *Phys. Chem. Chem. Phys. PCCP* **16**, 24128–24164 (2014).

29. Squires, T. M. & Mason, T. G. Fluid Mechanics of Microrheology. *Annu. Rev. Fluid Mech.* **42**, 413–438 (2009).

30. Crater, J. S. & Carrier, R. L. Barrier Properties of Gastrointestinal Mucus to Nanoparticle Transport. *Macromol. Biosci.* **10**, 1473–1483 (2010).

31. Schuster, B. S., Suk, J. S., Woodworth, G. F. & Hanes, J. Nanoparticle diffusion in respiratory mucus from humans without lung disease. *Biomaterials* **34**, 3439–3446 (2013).

32. Suk, J. S., Xu, Q., Kim, N., Hanes, J. & Ensign, L. M. PEGylation as a strategy for improving nanoparticle-based drug and gene delivery. *Adv. Drug Deliv. Rev.* **99**, 28–51 (2016).

33. Suk, J. S. *et al.* The penetration of fresh undiluted sputum expectorated by cystic fibrosis patients by non-adhesive polymer nanoparticles. *Biomaterials* **30**, 2591–2597 (2009).

34. Lai, S. K., Wang, Y.-Y., Cone, R., Wirtz, D. & Hanes, J. Altering Mucus Rheology to "Solidify" Human Mucus at the Nanoscale. *PloS One* **4**, e4294 (2009).

35. Lai, S. K., Wang, Y.-Y., Hida, K., Cone, R. & Hanes, J. Nanoparticles reveal that human cervicovaginal mucus is riddled with pores larger than viruses. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 598–603 (2010).

36. Yang, M. *et al.* Nanoparticle penetration of human cervicovaginal mucus: The effect of polyvinyl alcohol. *J. Control. Release Off. J. Control. Release Soc.* **192**, 202–208 (2014).

37. Ensign, L. M., Schneider, C., Suk, J. S., Cone, R. & Hanes, J. Mucus Penetrating Nanoparticles: Biophysical Tool and Method of Drug and Gene Delivery. *Adv. Mater. Deerfield Beach Fla* **24**, 3887–3894 (2012).

38. Mason, T. G. & Weitz, D. A. Optical Measurements of Frequency-Dependent Linear Viscoelastic Moduli of Complex Fluids. *Phys. Rev. Lett.* **74**, 1250–1253 (1995).

39. Wang, Y.-Y. *et al.* The Microstructure and Bulk Rheology of Human Cervicovaginal Mucus are Remarkably Resistant to Changes in pH. *Biomacromolecules* **14**, 4429–4435 (2013).

40. Wong, I. Y. *et al.* Anomalous Diffusion Probes Microstructure Dynamics of Entangled F-Actin Networks. *Phys. Rev. Lett.* **92**, 178101 (2004).

41. Weigel, A. V., Simon, B., Tamkun, M. M. & Krapf, D. Ergodic and nonergodic processes coexist in the plasma membrane as observed by single-molecule tracking. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 6438–6443 (2011).

42. Brunelli, R. *et al.* Globular structure of human ovulatory cervical mucus. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* **21**, 3872–3876 (2007).

43. Li, L., Lieleg, O., Jang, S., Ribbeck, K. & Han, J. A microfluidic in vitro system for the quantitative study of the stomach mucus barrier function. *Lab. Chip* **12**, 4071–9 (2012).

44. Radicioni, G. *et al.* The innate immune properties of airway mucosal surfaces are regulated by dynamic interactions between mucins and interacting proteins: the mucin interactome. *Mucosal Immunol.* **9**, 1442–1454 (2016).

45. Hein, M., Helmig, R. B., Schønheyder, H. C., Ganz, T. & Uldbjerg, N. An in vitro study of antibacterial properties of the cervical mucus plug in pregnancy. *Am. J. Obstet. Gynecol.* **185**, 586–592 (2001).

46. Caldara, M. *et al.* Mucin Biopolymers Prevent Bacterial Aggregation by Retaining Cells in the Free-Swimming State. *Curr. Biol.* **22**, 2325–2330 (2012).

47. Kavanaugh, N. L., Zhang, A. Q., Nobile, C. J., Johnson, A. D. & Ribbeck, K. Mucins Suppress Virulence Traits of Candida albicans. *mBio* **5**, e01911-14 (2014).

48. Frenkel, E. S. & Ribbeck, K. Salivary mucins protect surfaces from colonization by cariogenic bacteria. *Appl. Environ. Microbiol.* AEM.02573-14 (2014). doi:10.1128/AEM.02573-14

49. Lewis, W. G., Robinson, L. S., Gilbert, N. M., Perry, J. C. & Lewis, A. L. Degradation, Foraging, and Depletion of Mucus Sialoglycans by the Vagina-adapted Actinobacterium Gardnerella vaginalis. *J. Biol. Chem.* **288**, 12067–12079 (2013).

50. Cauci, S. & Culhane, J. F. High sialidase levels increase preterm birth risk among women who are bacterial vaginosis–positive in early gestation. *Am. J. Obstet. Gynecol.* **204**, 142.e1-142.e9 (2011).

51. McGregor, J. A. *et al.* Bacterial vaginosis is associated with prematurity and vaginal fluid mucinase and sialidase: results of a controlled trial of topical clindamycin cream. *Am. J. Obstet. Gynecol.* **170**, 1048–1060 (1994).

52. Meis, P. J. *et al.* The preterm prediction study: significance of vaginal infections. National Institute of Child Health and Human Development Maternal-Fetal Medicine Units Network. *Am. J. Obstet. Gynecol.* **173**, 1231–1235 (1995).

53. Silver, H. M., Sperling, R. S., St Clair, P. J. & Gibbs, R. S. Evidence relating bacterial vaginosis to intraamniotic infection. *Am. J. Obstet. Gynecol.* **161**, 808–812 (1989).

54. Gravett, M. G., Hummel, D., Eschenbach, D. A. & Holmes, K. K. Preterm labor associated with subclinical amniotic fluid infection and with bacterial vaginosis. *Obstet. Gynecol.* **67**, 229–237 (1986).

55. Hillier, S. L. *et al.* Association between bacterial vaginosis and preterm delivery of a low-birth-weight infant. The Vaginal Infections and Prematurity Study Group. *N. Engl. J. Med.* **333**, 1737–1742 (1995).

56. Care, A. G. *et al.* Predicting preterm birth in women with previous preterm birth and cervical length ≥ 25 mm. *Ultrasound Obstet. Gynecol. Off. J. Int. Soc. Ultrasound Obstet. Gynecol.* **43**, 681–686 (2014).

57. Hassan, S. S. *et al.* Patients with an ultrasonographic cervical length < or =15 mm have nearly a 50% risk of early spontaneous preterm delivery. *Am. J. Obstet. Gynecol.* **182**, 1458–1467 (2000).

58. Grobman, W. A. *et al.* Prediction of spontaneous preterm birth among nulliparous women with a short cervix. *J. Ultrasound Med. Off. J. Am. Inst. Ultrasound Med.* **35**, 1293–1297 (2016).

59. Hiersch, L. *et al.* Role of Cervical Length Measurement for Preterm Delivery Prediction in Women With Threatened Preterm Labor and Cervical Dilatation. *J. Ultrasound Med. Off. J. Am. Inst. Ultrasound Med.* **35**, 2631–2640 (2016).

60. Melamed, N. *et al.* Predictive value of cervical length in women with threatened preterm labor. *Obstet. Gynecol.* **122**, 1279–1287 (2013).

61. Iams, J. D. *et al.* The preterm prediction study: can low-risk women destined for spontaneous preterm birth be identified? *Am. J. Obstet. Gynecol.* **184**, 652–655 (2001).

62. Goldenberg, R. L., Mayberry, S. K., Copper, R. L., Dubard, M. B. & Hauth, J. C. Pregnancy outcome following a second-trimester loss. *Obstet. Gynecol.* **81**, 444–446 (1993).

63. Goldenberg, R. L. *et al.* The preterm prediction study: fetal fibronectin testing and spontaneous preterm birth. NICHD Maternal Fetal Medicine Units Network. *Obstet. Gynecol.* **87**, 643–648 (1996).
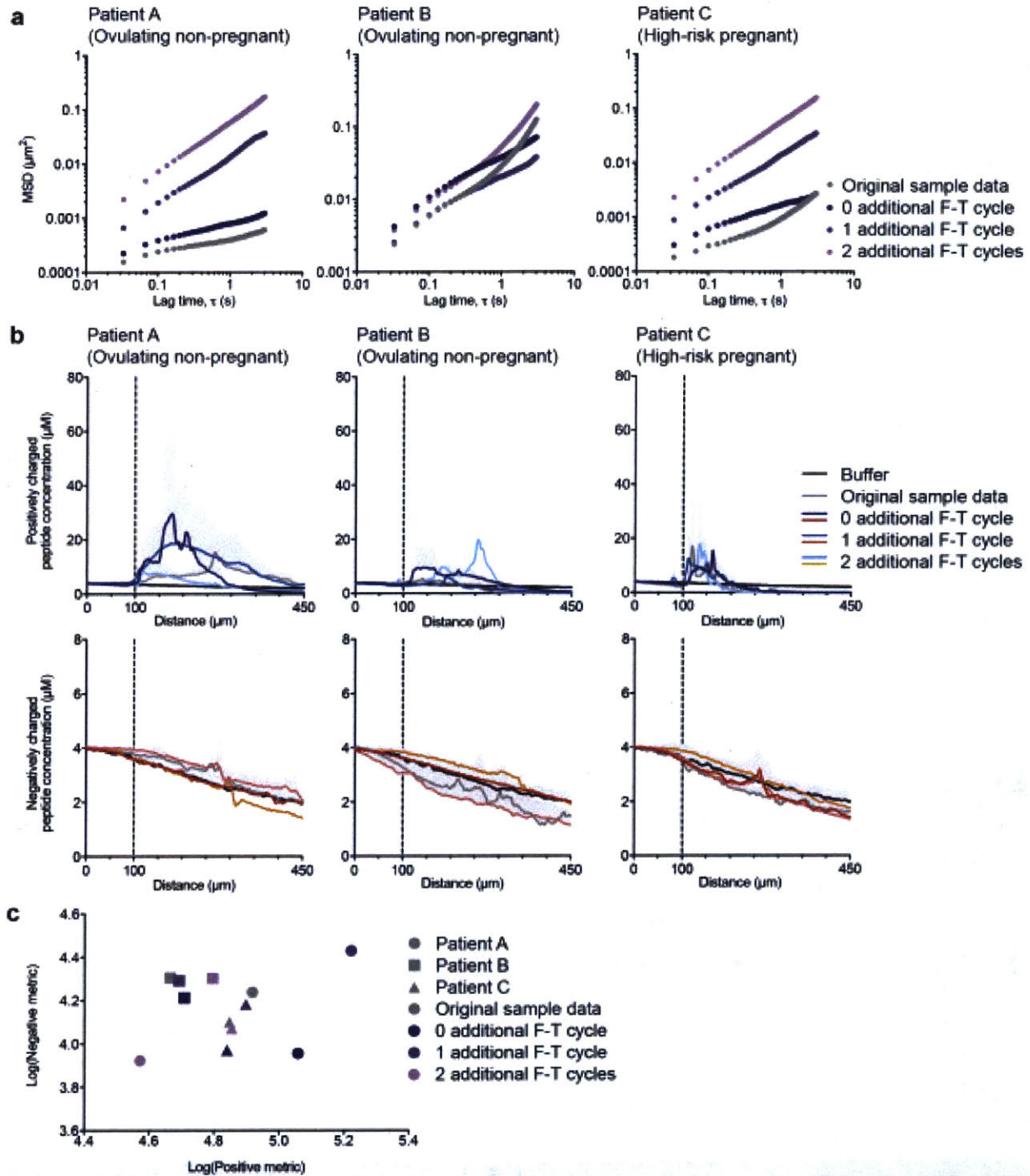
B.6 SUPPORTING INFORMATION

a

Patient A (Ovulating non-pregnant)

Patient B (Ovulating non-pregnant)

Patient C (High-risk pregnant)

- Original sample data
- 0 additional F-T cycle
- 1 additional F-T cycle
- 2 additional F-T cycles

b

Patient A (Ovulating non-pregnant)

Patient B (Ovulating non-pregnant)

Patient C (High-risk pregnant)

- Buffer
- Original sample data
- 0 additional F-T cycle
- 1 additional F-T cycle
- 2 additional F-T cycles

c

- Patient A
- Patient B
- Patient C
- Original sample data
- 0 additional F-T cycle
- 1 additional F-T cycle
- 2 additional F-T cycles

Figure B.S1 Repeated freeze-thaw (F-T) cycles affect the microstructure of cervical mucus and possibly its adhesiveness. To determine the impact of F-T cycles on mucus integrity, permeability was measured in three cervical mucus samples across three cycles. F-T cycles were performed by snap-freezing the sample in liquid nitrogen, and then thawing it at 4 °C for as many cycles as indicated. (a) MSD of particles in mucus from three individuals after additional F-T cycles reveals an increase in the mobility of the micron-sized tracer particles as a function of number of F-T cycles in mucus from two patients (A and C), but not the third (B). Data presented in the main body of the manuscript for each individual patient are shown in grey. (b) Average concentration profiles of positively (top, blue) and negatively (bottom, red) charged peptides through mucus from three individuals show no drastic change in the permeability of either peptide with number of F-T cycles for patients B and C, while data gathered from the mucus of patient A were more variable between cycles. Peptide diffusion profile in buffer is shown in black, and data presented in the main body of the manuscript for each individual patient are shown in grey. In all concentration profiles, the vertical, dotted lines indicate the buffer-mucus interface. (c) Scatter plot of metric values confirms consistent peptide diffusion behavior across F-T cycles in mucus from patients B and C, but more heterogeneity is observed in mucus from patient A. Metric values represent the difference in peptide transport between mucus and buffer, integrated over the channel length and experiment duration.
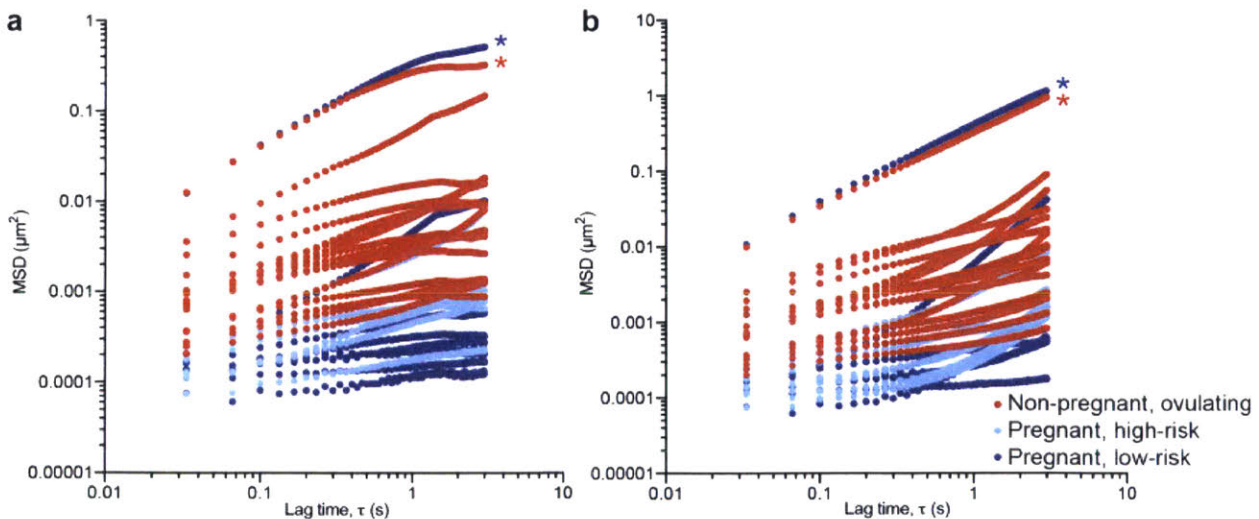

Figure B.S2 Supplemental Figure S2. MSD of particles de-drifted with the (a) individual particle algorithm introduces anomalies at large delay times in comparison to the (b) center-of-mass algorithm.
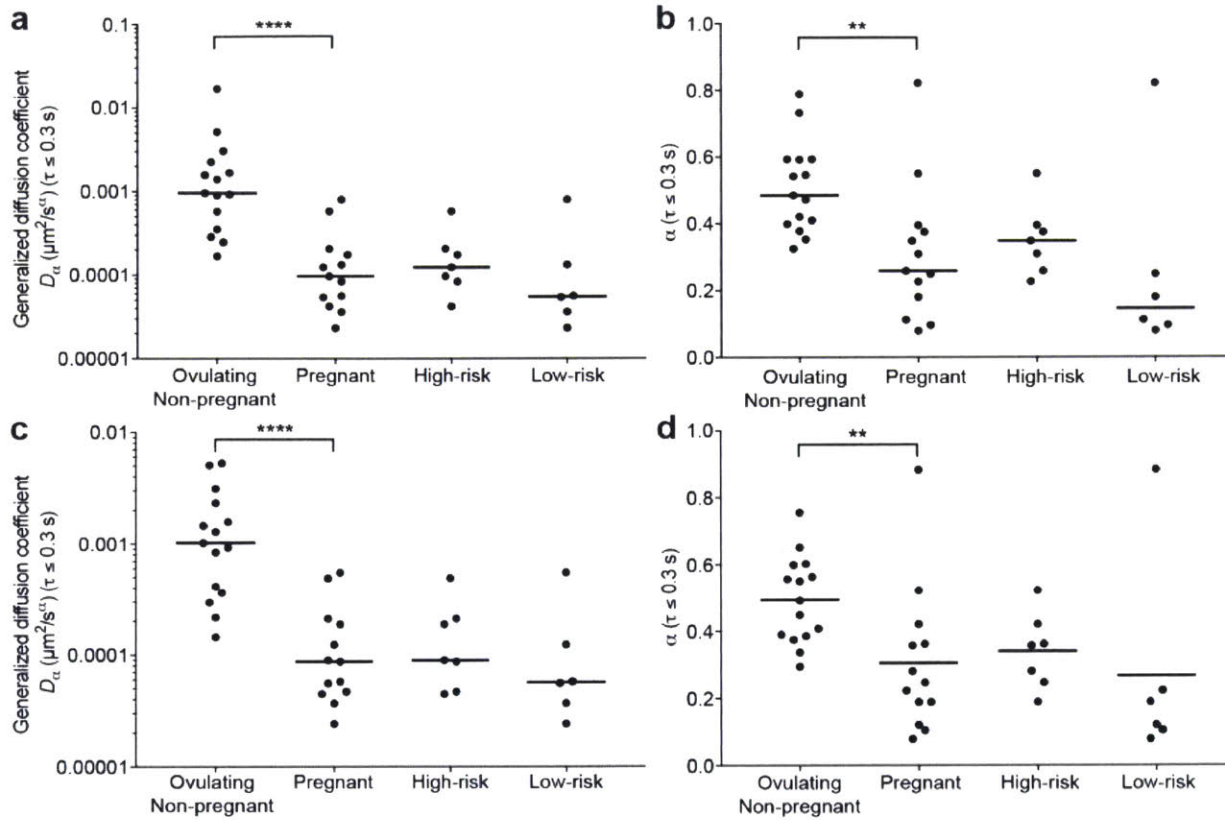
Figure B.S3 Significance of restricted particle mobility statistics (generalized diffusion coefficient, $D\alpha$, and (subdiffusion exponent, α) in mucus from pregnant patients is unaffected by dedrifting algorithm (a and b: individual particle, c and d: center-of-mass). (a) Diffusivity ($D\alpha$) and (b) subdiffusion exponent α values extracted from MSD results at early lag times ($\leq$ 0.3 s). Each point represents a single patient sample from non-pregnant ovulating (n=16) or pregnant (n=14) patients. Pregnant patient data are further subdivided into high (n=7)- and low (n=7) -risk groups. Bars, median values for each patient group. Significance was determined with the Mann-Whitney test (****P<.0001; **P<.0021
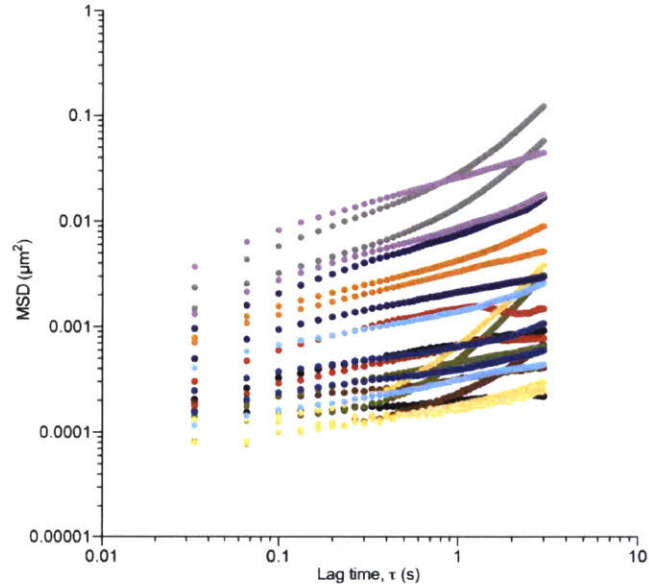
Figure B.S4 MSD curve from particle tracking in separate aliquots of cervical mucus from an individual patient is reproducible. 2 or 3 separate aliquots of cervical mucus from 11 patients were measured. Measurements from the same patient are paired by matching colors.
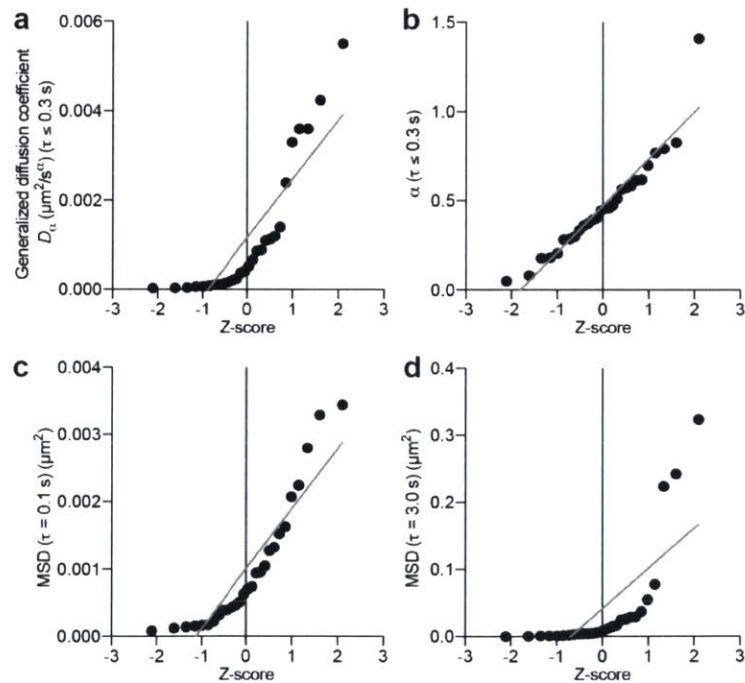

Figure B.S5 Parameters extracted from Single Particle Tracking (SPT) microrheology experiments in cervical mucus are not normally distributed. Quartile-Quartile (QQ) plots of Diffusivity,D, and anomalous diffusion exponent, α, extracted at early lag times (a and b). (c and d) QQ plots ofMSD values at 0.1 s and 3.0 s lag times, respectively.
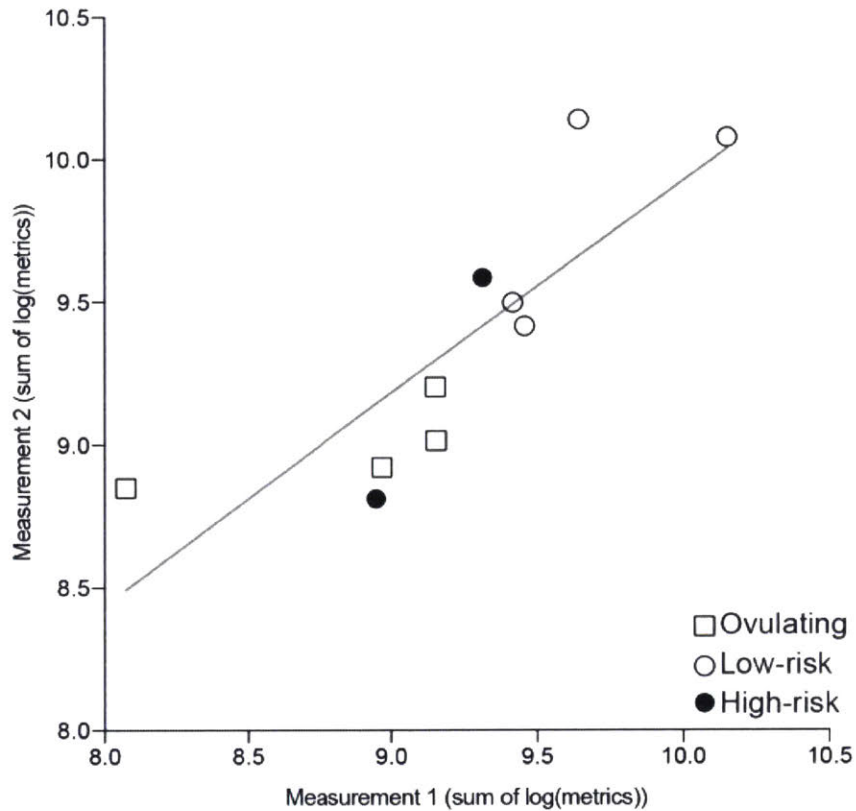
228

Figure B.S6 Permeability metric calculation from separate aliquots of mucus from an individual patient is reproducible. The sums of the positive and negative metrics taken from two separate measurements of mucus from separate aliquots form 10 patients are plotted as a function of one another. The grey curve represents the best-fit line, and confirms positive correlation between the measurements (P = 8.5E-3).
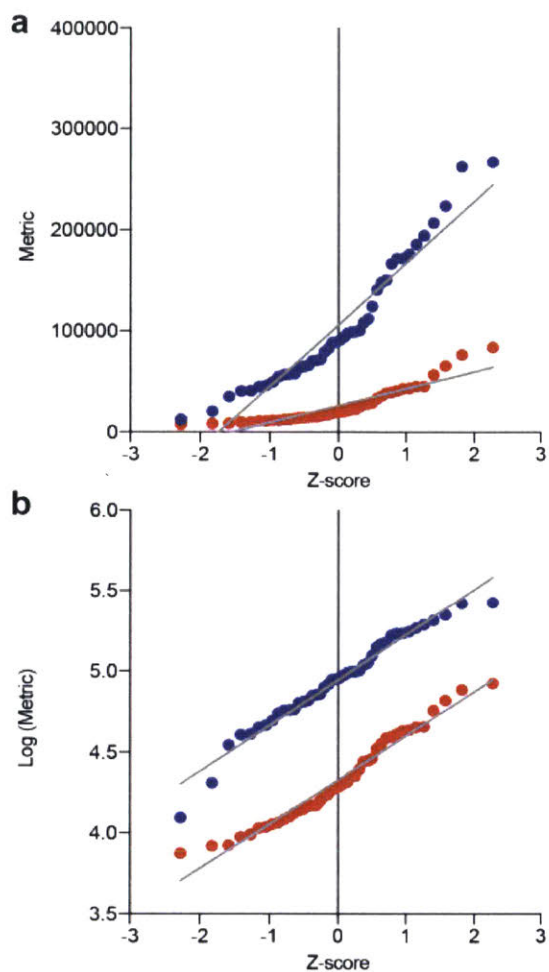
Figure B.S7 Permeability metric values are only normally distributed upon log transformation. Quartile-Quartile (QQ) plots of raw (a) and log-transformed (b) permeability metric values. Metric values corresponding to positive and negative peptides are represented as blue and red spheres, respectively. Grey curves represent best-fit lines to the data.