Molecular Characterization of Mucus Binding

Jacob Witten†,‡, Tahoura Samad,§ and Katharina Ribbeck*,†

†Department of Biological Engineering and ‡Computational and Systems Biology Initiative, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States

Supporting Information

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 mucin binding properties.

INTRODUCTION

Mucus is a viscoelastic hydrogel layer that coats every nonkeratinized epithelial surface of the body, including the lungs and intestines.1 Its activity has important ramifications for biology and medicine.2−5 The selective binding of small molecules to mucus may affect the activity and diffusivity of any molecule acting in a mucosal environment.6 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.7 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.8,6 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,7 suggesting that mucin binding is correspondingly more important in these pathologies. Synthetically cross-linked, 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.8 Mucins can also nonenzymatically catalyze some organic reactions by binding to relevant molecules.9

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 pathogen10). Both eDNA and bacterial polysaccharides are commonly found in CF lung mucus and may also bind small molecules.3 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.2 In these experiments, increased hydrophobicity typically correlates with slower drug diffusion.11 Lipids, and the hydrophobic domains of mucins, are believed to account for this slowed diffusion by binding to hydrophobic drugs3,5,8,12, although the glycans grafted to mucins may also play a role.9 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).13−18 This binding is believed to be mediated by the polyanionic glycosylated...
domains of mucins and, in CF, eDNA, and bacterial polysaccharides.\textsuperscript{3,5\textendash}7

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 mucin. One primary reason for this knowledge gap is that studies of mucus-small molecule interactions typically use small data sets, with fewer than 20 molecules per study (with one interesting exception that measured 41 molecules\textsuperscript{8,9,20,21}). 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 electrochemical methods,\textsuperscript{20} 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 data set (96 molecules in total), which constitutes a substantial increase over previous studies of drug\textendash 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,\textsuperscript{10} 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\textsuperscript{22} 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.

### EXPERIMENTAL SECTION

**Buffer Preparation and Specifications.** Phosphate-buffered saline (10\texttimes\ PBS) was purchased from AccuGene. At 1X concentration, it consisted of 1 mM KH\textsubscript{2}PO\textsubscript{4}, 2 mM Na\textsubscript{2}HPO\textsubscript{4}, and 150 mM NaCl at pH 7.4. Phosphate-citrate buffer (PCB) at 1X consisted of 10 mM each of phosphate and citrate, and 58 mM sodium, at pH 7.0.

**Mucin Preparation and Fluorescent Labeling.** MUC5AC was purified from fresh pig stomach scrapings as described in Celli et al.,\textsuperscript{23} 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, mucin was scraped from fresh pig stomachs (Research 87, Inc., Boylston, MA, U.S.A.) and mucins were purified using size exclusion chromatography.

To fluorescently label the mucin, 0.6 mg was solubilized in 500 \mu\text{L} of pH 9.0 0.1 M sodium bicarbonate buffer overnight at 4 °C. Alexa Fluor 647 succinimidyl ester (ThermoFisher Scientific) dissolved in dimethyl sulfoxide (DMSO) at 10 mg/\mu\text{L} was added to the mucin solution to a concentration of 0.02 mg/\mu\text{L} (1:100 mass ratio with respect to the 2 mg/\mu\text{L} mucin), and the mixture was incubated at room temperature for 1 h with shaking. An equal volume of 50 mM tris(hydroxymethyl)aminomethane was added to quench the reaction and the mixture was left at room temperature for 5 min with shaking. The volume was diluted to 20 \mu\text{L} with PBST (PBS with Tween 20) in a Corning Spin-X 100 kDa molecular weight cutoff ultrafiltration concentrator (Fisher) and centrifuged at 3000g for 25 min at 4 °C; the flow-through was discarded. These steps (dilute to 20 \mu\text{L} with PBST, centrifuge, discard flow-through) were repeated twice more; the volume remaining after each centrifugation was approximately 5 \mu\text{L}. The labeled mucin was then diluted to the appropriate concentration.

**Equilibrium Dialysis.** ED was performed with a 12 kDa cutoff Rapid Equilibrium Dialysis device (Thermo Scientific) with 100 \mu\text{L} of 5 mg/mL biopolymer solution or gel, or buffer control, in the sample chamber and 300 \mu\text{L} of matching buffer in the assay chamber. Each experiment began with equal concentrations of the relevant compound cocktail loaded in both chambers (‘1X concentration’). A 1X concentration was 200 nM for every molecule except colistin (10 \mu\text{M}). 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 (3 \mu\text{m} pores, 2.1 \times 100 mm, VWR). The HPLC method is given in Table S1. The first 5 min 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 4000 V, with molecular ions and fragments for all experiments given in Tables S2\textendash 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\textendash 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 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 \( R_C \) of a compound was calculated by (see text in the Supporting Information for derivation):

\[
R_C = 1 + \frac{1}{c_E} - \frac{1}{c_C}
\]

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\textendash 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 three wells for one full biological replicate of \( R_C \), which was then converted to \( \Delta G_C \), so \( n = 3 \) represents nine measured wells.

We used a biopolymer concentration of 0.5% w/v (5 mg/mL). A total of 5 mg/mL is consistent with typical mucin concentrations in lung mucus (0.1\text\textendash 20 mg/mL, depending on disease state\textsuperscript{24,25,26,27}) and DNA concentrations in CF sputum (0.5\text\textendash 5 mg/mL\textsuperscript{24,26}). It is higher than the bulk concentration of alginate in P. aeruginosa-infected CF sputum (0.004\text\textendash 0.1 mg/mL\textsuperscript{28}), but the local concentration in a P. aeruginosa biofilm may approach 5 mg/mL. Alginate (MW 120\text\textendash 190 kDa) and calf thymus genomic DNA were purchased from Sigma.
Small Molecule Microarray. SMM slides were prepared as previously described in Bradner et al.26 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 rinses) were used instead of TBST and TBS, slides were incubated at room temperature rather than 4 °C, 5 mL of protein and wash solution was used rather than 3 mL, 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 eight conditions. Six conditions were experimental, 0.1, 0.5, and 1 μg/mL mucin with 0.05% and 0.2% Tween 20, and two 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 four 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).

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)==N’). 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).

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 “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, U.S.A.).

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 the final concentration of approximately 3 × 10^4 CFU/mL into fresh MHB. A total of 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 1 mg/mL and 0.4 μg/mL. After incubation at 37 °C for 24 h, 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 × 10^6 CFU/mL into fresh MHB. A total of 45 μL of this inoculum was added to each well of a 96-well plate. A total of 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 (2 × MIC). After incubation at 37 °C for 24 h, 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).

Statistical Analysis and Plots. Except for the Fisher’s exact test, statistical analysis and plotting was performed in R. [R Core Team (2017). R is a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/]. Tests of significance 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:

\[ \text{lm}(\text{Mucin,ΔG}) \sim \text{DNA,ΔG} \sim \text{Charge, data = all data}) \]

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 three times, as described in “Equilibrium Dialysis” of the Experimental Section.

Data Availability. The data sets generated during the current study are available from the corresponding author on request.

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

RESULTS AND DISCUSSION

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 (Figure 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 to 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 (Figure 1a, Experimental Section). From these measurements, we calculated the equilibrium uptake ratio \( R_d \) given by

\[ R_d = \frac{[\text{compound}]_{\text{sample}}}{[\text{compound}]_{\text{assay}}} \]

(1)

Log-transforming \( R_d \) gives a value akin to an energy:

\[ \Delta G_{\text{uptake}}/k_BT = -\ln R_d \]  

(2)

For simplicity, we consider the unitless quantity \( \Delta G \) scaled by thermal energy \( k_BT \), meaning \( \Delta G = \Delta G_{\text{uptake}}/k_BT \). This quantity is more likely than \( R_d \) to scale linearly with molecular properties that affect binding energies, so it is a better
shows that mucin does not increase growth in the absence of activity is inhibited by mucin, but not by MC. No colistin control biopolymer-free Mueller-Hinton broth (MHB), in mucin (0.5% w/v colony forming units) after 24 h of growth with colistin in Colistin binds mucin but not MC. Error bars: SEM 0.85, df = 2). (c) Surviving incubated predictions can identify instances where mucus inhibits a drug by degree, while MC had a negligible effect of colistin to a substantial effect on colistin bioactivity differences 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.

Identification of a Mucin Binding Motif by Small Molecule Microarray. Toward this end, we used a small molecule microarray to screen for mucin-binding small molecules (Figure 2a, Experimental Section). 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 nonhits (risk ratio 21.9, p ~ 10−28, Fisher’s exact test without multiple hypothesis testing correction). We termed the motif 2,4-diaminopyrimidine (DAP; Figure 2c,d; Table 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 ~ 10−25).

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 nonseparable 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 ~ 160 mM, pH 7.4) and a lower salt phosphate–citrate buffer (PCB; 60 mM 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,24,30 while pH of lung mucus is neutral to slightly basic for healthy people and slightly acidic in CF patients.31 For binding to mucin in PCB, the average ΔG was lower for molecules with DAP than for molecules without (p = 0.0043, F1,50 = 8.9), meaning that the presence of DAP was positively correlated with increased mucin binding, though this relationship did not hold in PBS (Figures S1 and 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 (1 mM) is unlikely to have a significant impact.

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 representation of binding data for the regression analyses that follow. A negative ΔG implies binding, with more negative ΔG meaning tighter binding and 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.28 Our binding measurements revealed that colistin bound to mucin but not to MC (a neutral biopolymer) at 0.5% w/v (Figure 1b; we retained this biopolymer concentration throughout this entire work, see Experimental Section). 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 (Figure 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.

Figure 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, two 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 h 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.

Biomacromolecules 2019, 20, 1505−1513
DOI: 10.1021/acs.biomac.8b00747
Biomacromolecules 2019, 20, 1505−1513
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 nonpolyanionic 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).

The results of these experiments and analyses are presented in Figure 3, which depicts the slopes of best-fit lines for binding to each biopolymer vs the presence of DAP, net charge, and ClogP (Figure 3; full plots for each individual biopolymer–property pair given in Figures S1 and S2). In PCB, the presence of DAP was significantly correlated not only with binding to mucin but also with binding to DNA and alginate (Figure 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 (Figure 3d,e). These data suggest that (DNA-rich) CF lung mucus, and 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. Increased ClogP was also correlated with stronger binding to each of the biopolymers in PCB (Figure 3f), which is consistent with previous observations that increased hydrophobicity leads to greater mucus binding.

Some caution is required in interpreting these single-variable plots, as there is substantial collinearity between the presence of DAP and the other descriptors (Figure S3a–c). However, these results all hold true when restricting our data set to only molecules with DAP (Figure 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.

**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 Figure 3 held for this more diverse set of molecules (Figure 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 (Figure 4c), which may reflect mucin binding scaffolds beyond DAP that were unidentifed or untested in the SMM. The lack of clear trends observed in the NCC data set suggests that correlations between charge or hydrophobicity and binding to mucus may hold locally, but not globally, in chemical space (Figure 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 Figure 4c, binding sites and conformations would be more varied,
reducing the expectation of correlation between binding and coarse summaries of molecular properties such as charge and ClogP.

**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 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 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 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 mucus and likely abundant in the lungs of CF and COPD patients.6 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),\textsuperscript{5,32,33} mucin-based drug delivery systems,\textsuperscript{8} and studies of catalysis by mucin,\textsuperscript{9} and possibly to mucus-free biopolymer systems such as biofilms.\textsuperscript{34,35} In this vein, we have recently applied the equilibrium dialysis method to examine the binding of antibiotics to MUC5AC, MUC5B, and MUC2 and its relationship to antibiotic efficacy.\textsuperscript{36} Our results with alginate are relevant for the activities of small molecules within biofilms, and the application of our techniques to study small molecule binding to other biopolymer-associated polysaccharides such as the P. aeruginosa-synthesized Pel and PsI\textsuperscript{37} 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.\textsuperscript{38,39} Overall, 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.

**Table 1. Summarized Results Showing Associations between Molecular Properties and Binding to Mucin, Alginate, DNA, and BSA**

<table>
<thead>
<tr>
<th>biopolymer</th>
<th>correlation w/DAP</th>
<th>correlation w/charge</th>
<th>correlation w/hydrophobicity</th>
</tr>
</thead>
<tbody>
<tr>
<td>mucin</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>alginate</td>
<td>+</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>DNA</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>BSA</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

\(a\) 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.


29) Vardakas, K. Z.; Voulgaris, G. L.; Samonis, G.; Falagas, M. E. Inhaled Colistin Monotherapy for Respiratory Tract Infections


