Analyst

Microarrays for the screening and identification of carbohydrate-binding peptides

Journal:	*Analyst*
Manuscript ID	AN-ART-08-2019-001465.R1
Article Type:	Paper
Date Submitted by the Author:	12-Oct-2019
Complete List of Authors:	Shastry, Divya; Rensselaer Polytechnic Institute Karande, Pankaj; Rensselaer Polytechnic Institute, Howard P. Isermann Department of Chemical and Biological Engineering

ARTICLE

Microarrays for the screening and identification of carbohydrate-binding peptides

Divya G. Shastry*[a,b] and Pankaj S. Karande†[b,c]

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

The development of carbohydrate-binding ligands is crucial for expanding knowledge on the glycocode and for achieving systematic carbohydrate targeting. Amongst such ligands, carbohydrate-binding peptides (CBPs) are attractive for use in bioanalytical and biomedical systems due to their biochemical and physicochemical properties; moreover, given the biological significance of lectin–carbohydrate interactions, these ligands offer an opportunity to study peptide sequence and binding characteristics to inform on natural target/ligand interactions. Here, a high-throughput microarray screening technique is described for the identification and study of CBPs, with a focus on polysialic acid (PSA), a polysaccharide found on neural stem cells. The chemical and biological uniqueness of PSA suggests that an ability to exclusively target this glycan may promote a number of diagnostic and therapeutic applications. PSA-binding peptides from phage display screening and from epitope mapping of an ScFv for oligosialic acid were screened in an optimized microarray format with three ligand density conditions. Hypothesis-driven mutations were additionally applied to select peptides to modulate peptide affinity and selectivity to PSA. Peptide compositional and positional analyses revealed the significance of various residues for PSA binding and suggested the importance of basic residue positioning for PSA recognition. Furthermore, selectivity studies performed directly on microarrays with chondroitin sulfate A (CS-A) demonstrated the value of screening for both affinity and selectivity in the development of CBPs. Thus, the integrated approach described, with attention to design strategy, screening, and peptide characterization, successfully identified novel PSA-binding ligands and offers a platform for the identification and study of additional polysaccharide-binding peptides.

Introduction

Carbohydrates demonstrate crucial and versatile roles in a number of biological processes, including cellular growth and differentiation, microbial pathogenesis, and immune recognition.[1] For some time, the importance of understanding carbohydrate-encoded information has been acknowledged. Recognition of the significance of harnessing this information for molecular and cellular targeting has followed suit, with examples of natural and engineered carbohydrate targeting spread throughout scientific literature.[2–5] However, with the paradigm of biological coding still evolving,[6] particularly in regards to the glycocode, it is not surprising that systematic methods for targeting have not reached the rigor of protein targeting.

Numerous carbohydrate-binding ligands and platforms have been identified (or developed) and studied, the most common being lectins (i.e., carbohydrate-binding proteins).[7] Artificial lectins, antibodies, aptamers, and peptides have also been developed, with each molecular class offering its own advantages.[2,7–10] The stability of protein ligands has commonly been cited as an impetus to generate artificial ligands, like synthetic lectins (including borono-lectins), nucleic acid-based aptamers, and peptides, against carbohydrates.[2] Additionally, limitations in ligand synthesis (e.g., need for cellular protein synthesis machinery) have also been stated, though biotechnological advancements are rapidly overcoming these drawbacks. Nevertheless, apart from stability and synthesis aspects, smaller ligands offer distinct advantages of their own. Specifically, advantages in application arise from molecular size and structure, especially relating to integration of small molecule ligands in bioanalytical and biomedical platforms.

Both peptides and aptamers are similar in that they present an ease of synthesis over wholly synthetic ligands,[2,8] and carbohydrate-binding peptides (CBPs) and aptamers have been discovered with random and rational methodologies, incorporating structure-based design and non-natural moieties.[2,10,11] However, overcoming repulsive charge interactions is a greater barrier with aptamer development when targeting carbohydrates with high negative charge.[11] Additionally, from solely a design and characterization perspective, peptides offer an excellent biomimicry opportunity in comparing binding characteristics to natural lectins, for which specific recognition capabilities have been extensively studied.

[a] Department of Biological Sciences, Rensselaer Polytechnic Institute, Troy, NY 12180, USA. Email: shastd3@gmail.com
[b] Center for Biotechnology and Interdisciplinary Studies, Rensselaer Polytechnic Institute, Troy, NY 12180, USA.
[c] Howard P. Isermann Department of Chemical and Biological Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180, USA. Email: karanp@rpi.edu
* Corresponding author
† Co-corresponding author
Electronic Supplementary Information (ESI) available. See DOI: 10.1039/x0xx00000x

A number of studies on peptides as artificial lectins have been reported.[12–19] Most commonly, these studies isolate peptides from lectin-binding pockets, for which structural and/or molecular characterization has been accomplished, or approach peptide discovery with screening of large random or combinatorial libraries. Possibly due to constraints arising from low affinity binding,[2,7,17] few reports relating to CBPs utilize multiple strategies to optimize and/or characterize peptide sequences and binding properties (as compared to studies with naturally occurring lectins). This may not be disadvantageous in regard to identification of desired peptides; yet, many methods for peptide identification fail to provide flexibility in lead generation while simultaneously offering a platform for initial ligand characterization. Hence, versatility of these methodologies and their ability to expand upon knowledge of the glycocode is minimized.

The above is especially true for the generation of peptide ligands for polysaccharides. A number of CBPs have been created for monosaccharides and oligosaccharides, but there are few examples of polysaccharide-binding peptides, either naturally occurring or engineered.[2] The higher structural rigidity of smaller carbohydrates, as well as relative ease of structure determination (for these sugars and their natural lectin complexes), may contribute to this disparity. Notably, Svarovsky and colleagues have developed random peptide arrays enabling detection of lipopolysaccharides (LPS) on bacterial cell walls, where glycan portions of the LPS are detected.[20–22] Heparin-binding peptides have also been identified, though established consensus binding motifs from heparin-binding proteins facilitated these studies.[23–27] Heparin-binding motifs were also used for the generation of a peptide library challenged with chondroitin sulfate A (CS-A),[28,29] representing a rational approach driven by lack of knowledge on CS-A binding motifs. In separate work, successful *in vivo* targeting was demonstrated with CS-A peptides from a random phage display library.[30]

The limited examples available for systematic targeting of polysaccharides with peptide ligands necessitates the development of a platform that offers flexibility at multiple stages—in peptide design, screening, and characterization—thus offering an iterative approach for the study of CBPs. Here, a high-throughput peptide microarray screening technique is described for the identification of CBPs, using polysialic acid (PSA) as an exemplary glycan target. Peptides derived from random screening and rational design were immobilized in array format and challenged with the PSA target for assessment of PSA affinity and selectivity.

Unlike in previous studies, an integrated random and rational approach is demonstrated, wherein novel ligands are identified from a combination of phage display screening, lectin mapping, and mutational approach, and ligand characterization informs on target binding properties. Additionally, while peptides binding to the sialic acid monosaccharide have been identified,[31–35] this report is the first to identify peptide binders to PSA. PSA is a highly negatively charged polysaccharide playing important roles in physiology, development, and pathogenesis, though limited information on structure, binding motifs, and ligands is available.[36–38] Thus, the approach herein enabled identification of peptides with appreciable affinity and selectivity to a large glycan target and demonstrated a methodology that can be used for characterization of peptide–polysaccharide interactions.

Experimental

Phage display screening

Phage display screening of colominic acid (CA) was conducted with the Ph.D.™-12 Phage Display Peptide Library (New England BioLabs, Inc., Ipswich, MA). Biopanning of CA was carried out in accordance with standard NEB protocols,[39] with modifications for the polysaccharide target described as follows. CA was immobilized prior to each round of biopanning on wells of a polystyrene 96-well plate (CELLTREAT Scientific Products, Pepperell, MA), pre-coated with poly-L-lysine (PLL) to enhance immobilization.[40] PLL coating was achieved through incubation of 300 μL of 0.1 mg/mL PLL in sterile phosphate-buffered saline (PBS; 0.01% w/v solution) (Cultrex, Trevigen, Inc., Gaithersburg, MD) for 1 hour at room temperature. Following removal of PLL solution, CA was immobilized on PLL-coated wells through overnight incubation of 300 μL of 5 mg/mL CA sodium salt (Nacalai USA, Inc., San Diego, CA) in deionized water at 4 °C in a humidified container. CA solution was removed, and blocking buffer, sterile-filtered (1% bovine serum albumin [BSA] in phosphate-buffered saline [PBS], w/v; 10 mM phosphate, 100 mM NaCl, pH 7.4) was added at room temperature for 1 hour. All incubation steps were performed with mild shaking.

Briefly, biopanning against CA encompassed screening of the Ph.D.™-12 phage library on CA-coated wells, elution of binders, and amplification of binders for the next round of panning. All wash steps were performed with PBS buffer (pH 7.4). Bound phages were eluted at high pH (30 mM triethylamine; neutralized with 1 M Tris-HCl, pH 7.4)[41] and low pH (0.2 M glycine-HCl, pH 2.2 with 1 mg/mL BSA; neutralized with 1 M Tris-HCl, pH 9.1) in separate screens. A total of three rounds of panning were conducted for each screen. Twenty phages were selected at random at the end of the third round of panning for each elution screen, and sequencing was performed (Genscript, Piscataway, NJ).

Peptide library design and synthesis

A peptide library was constructed from linear epitope mapping of mAb735, a monoclonal antibody specific to oligosialic acid,[42] and from lead peptides derived from phage display screening. In contrast to peptides originating from random screening, peptides were designed from mAb735 through linear epitope mapping of the scFv (Peptide IDs 1–102). Peptides 15 amino acids in length were designed from the primary sequence of the mAb with 13 residue overlaps, such that consecutive peptides were shifted by 2 amino acids. For phage display peptides, 81 mutant peptides were included in the library along with 40 parent sequences derived from random screening (Peptide IDs 102–223). All mutations to phage display-derived peptides

encompassed single and double mutations to select parent sequences, except one peptide (peptide 170) into which putative PSA-binding residues from the protein Siglec-11 were spliced.[43] Mutations were rationally introduced into parent peptides, and hence positions and identities of residues selected for mutation (as well as identities of introduced residues) did not encompass all combinatorial possibilities.

Peptides were synthesized in high-throughput with SPOT technology using standard Fmoc (fluorenylmethyloxycarbonyl) chemistry in an automated peptide synthesizer (Multipep RS, INTAVIS Bioanalytical Instruments AG, Germany) as described previously.[44–46] Briefly, peptide synthesis was carried out in a parallel fashion from C- to N-terminus on solid cellulose-based discs. Though number and spacing of functional groups for amino acid conjugation are uniform amongst commercially available discs, the same batches of discs and reagents were used across the peptide library to minimize any sequence-specific variability in peptide density.

To maintain charge of mAb peptides as they occur within the parent protein sequence, these peptides were N-terminally acetylated. Phage display-based peptides (parent and mutant sequences) were synthesized with free N-termini and C-terminal GGGS linkers so as to mimic peptide presentation on phage particles.[39] Two positive and negative control peptides derived from the Siglec-11 ligand were also synthesized with N-terminal acetylation for parallel microarray screening with the peptide library (details on origins of control peptides are provided with ESI Table 3). Control peptides were subsequently assessed with SPR spectroscopy as described in ESI Fig. 7. Following synthesis, peptide–cellulose discs were subjected to the standard work-up procedure outlined by Intavis[47] and reconstituted in 250 µL of dimethyl sulfoxide (DMSO) to create stock solutions of peptide–cellulose conjugates. Peptide stock solutions were diluted 1:2 for spotting on microarray substrates.

Peptide microarray preparation

Peptide solutions prepared as described above were spotted in triplicate in a spatially addressable manner on nitrocellulose-coated glass microscope slides using the Slide Spotting Robot from Intavis Bioanalytical Instruments AG. Printed microarrays were air-dried prior to heating for 2 hours at 65 °C to improve adhesion of peptide–cellulose conjugates to slides. Microarrays encompassing the complete peptide library were prepared with the following three printing conditions: i) 60 nL spot volume (1x condition), ii) 120 nL spot volume (2x condition), and iii) 60 nL spot volume, air-dried, followed by a second printing of 60 nL directly superimposed on the initially printed spots (1x2 condition). For 2x and 1x2 conditions, the peptide library was divided into two microarrays due to space limitations, and differences arising from slide and screening variability were accounted for in microarray data analysis (see below). As per manufacturer provided values on cellulose disc parameters, each 60 nL spot contained approximately 0.1 pmol cellulose fibers and 11.5 pmol peptide (corresponding to approximately 6.4×10^{10} and 7.42×10^{11} molecules, respectively).

Microarray screening and data acquisition

Microarray screening was conducted at room temperature (approximately 20 °C) with shaking. Prepared microarrays were blocked for 3 hours in 5 mL of 5% (w/v) BSA in PBS (10 mM phosphate, 100 mM NaCl, pH 7.4). Following removal of blocking buffer, microarrays were subjected to three 10-minute washes in PBS (5 mL each). The same wash procedure was employed after all subsequent incubation steps. Target incubation of the microarrays was performed with 5 mL of CA sodium salt (Nacalai USA, Inc.) at 10 µM in PBS for 3 hours. For selective-binding studies, chondroitin sulfate A sodium salt from bovine trachea (CS-A) (Sigma-Aldrich, St. Louis, MO) was incorporated in target incubation through addition of 0.1, 1, or 10 µM of CS-A to the CA solution (see ESI Fig. 6 for further information regarding molecular weight estimation of CS-A). CA binding to peptides in the presence or absence of competitor was detected indirectly through antibody-based detection using anti-PSA-NCAM primary antibody, clone 2-2B (MilliporeSigma, Burlington, MA) at 1:5000 in 2.5% (w/v) BSA in PBS and horseradish peroxidase (HRP)-conjugated goat anti-mouse IgG, IgM (H+L) secondary antibody (Thermo Fisher Scientific, Waltham, MA) at 1:500,000 in 2.5% (w/v) BSA in PBS. Each antibody was incubated for 1 hour (5 mL per microarray). Binding of primary and secondary antibodies to the peptide library in the absence of CA were assessed to account for non-target antibody–peptide binding, if any, in binding intensity analysis. Similarly, dot blot assays, in which CA and CS-A were immobilized at various concentrations on nitrocellulose membranes and targeted with screening antibodies, were performed to ensure antibodies did not show cross-reactivity with CS-A. Throughout the microarray screening procedure, hydration of the microarray surface was ensured, as partial or complete drying of the surface at any step prior to imaging severely compromised image quality.

Chemiluminescence imaging was performed using a ChemiDoc™ XRS+ System and Image Lab 4.0 software (Bio-Rad, Hercules, CA) with a protocol optimized for acquisition of high-resolution microarray images. A 5 mL volume of SuperSignal™ West Femto Maximum Sensitivity Substrate (Thermo Fisher Scientific) was added to each slide immediately prior to imaging. Images were acquired in signal accumulation mode (SAM) in high resolution chemiluminescent imaging mode (with 2x2 binning) to maximize signal-to-noise ratio without signal saturation. SAM parameters were set to the first image time at 1 second, the last image time at 30 seconds, and a total of 15 images.

Positive and negative control peptides synthesized with the peptide library were established as such through parallel screening with the complete peptide library on microarrays with CA and CS-A. To validate the capability of microarray screening to distinguish between binding and non-binding

peptides, the control peptides were evaluated for binding to PSA with SPR spectroscopy. Further details of SPR spectroscopy are provided along with ESI Fig. 7.

Microarray data analysis

Images captured at 30 seconds exposure were selected for analysis. In order to enable comparison amongst various microarray printing and screening conditions, all microarray images were normalized to identical image transform values and exported at 1200 dpi. Image transformation in Image Lab did not alter underlying image data; transformation removed variation arising from chemiluminescent imaging conditions (e.g., auto scaling of a slide with low signal to a compressed intensity range) and did not remove variation due to screening conditions of interest (e.g., CS-A concentration). Transform values were selected to maximize dynamic range of spot intensities and signal-to-noise ratio of peptide spot intensity to background microarray intensity in subsequent microarray image analysis in TIGR Spotfinder **(Dana Farber Cancer Institute, Boston, MA)**. Exported images were modified in ImageJ (NIH, Bethesda, MD) to enable analysis in the microarray analysis software**.** ImageJ modifications encompassed image conversion to 8-bit grayscale, background subtraction using the rolling ball radius algorithm (rolling ball radius of 50 pixels), and color inversion. Quantification of peptide spot intensities in modified images was conducted with the Otsu segmentation method utilizing local background subtraction in TIGR Spotfinder with the following gridding and processing parameters: mask size 10, minimum spot size 3, maximum spot size 10, and top background cut-off 5%. Any aberrant spots arising from slide irregularities, printing error, or physical marks during screening were flagged during processing to prevent erroneous calculations of peptide average intensity and standard deviations. Only clearly aberrant spots were excluded in such a manner (0–0.01% of spots per slide).

Statistical analyses on intensity data thus obtained were performed in Microsoft Excel 2016. Average intensities and standard deviations for triplicate peptide spots were determined from spot intensities using Visual Basic macros, and these data were used in the calculation of inter- and intra-assay coefficients of variation (CVs), in a similar manner as reported by Sachse *et al*.[46] For comparison of microarray printing conditions, bivariate scatterplots for pairwise conditions were created, and Pearson correlation coefficients (r) were calculated. Values of $0 < r < 0.30$ were noted as very weak, $0.30 \leq r < 0.50$ as weak, $0.50 \leq r < 0.70$ as moderate, $0.70 \leq r < 0.90$ as strong, and $0.90 \leq r < 1.00$ as very strong correlations, with $r = 0$ and $r = 1.00$ considered no correlation and perfect correlation, respectively. Based on the low intra-assay CV for high intensity peptides obtained in 1x, 2x, and 1x2 conditions with screening for peptide–CA binding in the absence of competitor (i.e., screening with a target solution of only 10 μM CA in PBS, from henceforth referred to as "affinity" screening‡, in contrast to screening with addition of CS-A, termed "selectivity" screening), the top 5% of binding peptides from averaged results from the three screens were identified. **Selectivity of these peptides in 1x, 2x, and 1x2 screening with 10 μM CA was assessed through determination of percent selectivity, where percent selectivity was calculated as percent of average binding intensity in 10 μM CA screening remaining with the addition of 1 μM CS-A (10:1 molar ratio CA:CS-A). Additionally, percent selectivity of all PSA-binders in the complete peptide library was determined to identify selective binders regardless of affinity rank. Peptides displaying binding above background or non-binding levels in affinity screening were identified as PSA-binders, where background was defined as intensities falling in the bottom 20% of the intensity range of a screen (approximately 70–80% of the library.) Binding of all high affinity and high selectivity peptides was visually verified in microarray images.**

Peptide residue compositional and positional analyses were conducted to assess variation in frequency of residue occurrence between the unmodified peptide library and peptides with top 10% affinity binding in this set; frequency variation was assessed in the peptide sequences (compositional analysis) and at each position within the sequences (positional analysis).[28] Binding data from the 1x printing condition was used for these analyses. Only unmodified sequences from mAb epitope mapping and phage display screening were used for compositional and positional analyses to avoid sampling bias arising from mutant phage display-derived peptides (comprising >10% of the library), as select parent phage peptides were chosen for modification and overrepresentation of these sequences would otherwise bias statistical outcomes. For compositional analysis, statistical significance of changes in residue propensity was determined using the test statistic *z*, calculated as in eqn (1):

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \quad (1)$$

where $\hat{p}$ is the percentage of a residue, or type of residue, occurring in the top 10% affinity binders; p_0 is the percentage of the same residue or residue type occurring in the library; and n is the total number of residues in the library.

Results and discussion

In this study, a peptide library composed of random and rationally-designed peptides was screened for binding to polysialic acid (PSA) in an optimized high-throughput microarray format. The modification of existing peptide microarray technology allowed for rapid detection of PSA-binding peptides. PSA was chosen as a polysaccharide target for structural and functional considerations. From a structural perspective, PSA is a linear homopolymer with a degree of polymerization estimated to vary between 8 and 200 residues,[36]

and thus presents an opportunity to target repeating monomeric units. While PSAs of differing polymeric composition exist, the α-(2,8)-linked form of PSA used in this study occurs naturally in mammalian neural stem and progenitor cells and in the polysaccharide capsules of *Escherichia coli* K1 and *Neisseria meningitidis* serogroup B.[36] In these occurrences, changes in polysialylation on proteins or cell surfaces is tied to functional states. The identification of peptides for targeting or detection of PSA can thus serve a range of diagnostic and therapeutic purposes relating to such changes in healthy or disease contexts.

The high-throughput microarray screening employed allowed for the simultaneous screening of up to 1200 peptide spots per microarray for polysaccharide binding. The lower equilibrium dissociation constants (micromolar to millimolar; K_D's) generally associated with carbohydrate–protein binding translates to carbohydrate–peptide interactions, with the additional challenge of higher entropic penalties for small molecule binding.[2,7] In order to identify such "weak" (as compared with nanomolar binding) affinity carbohydrate interactions, two conditions exist: i) a technique must either be of sufficiently high sensitivity to detect low affinity interactions or ii) the strength of the interaction must be enhanced. These conditions are not mutually exclusive; for example, increasing the strength through avidity effects can promote detection of the interaction. Here, successful screening of a peptide library and identification of PSA-binding peptides addressed both detection (through assay development‡) and interaction strength (through peptide density and peptide design).

Peptide microarray screening

In contrast to previous studies where glycan targets were modified for detection,[20–22,28] PSA-binding peptides were indirectly detected with chemiluminescence-based detection of colominic acid (CA) (Fig. 1B–C; analyses from microarray data are shown in Fig. 1A). CA is the bacterial homolog of PSA derived from *E. coli* K1 and is structurally identical to the PSA found in humans, specifically poly-alpha 2,8-(N-acetylneuraminic acid) (the CA used has an average molecular weight of 30 kD with a normal distribution of chain lengths; ESI Fig. 6).[40] Direct labeling of CA, as with a fluorescent label, may influence peptide binding either through allosterically affecting PSA–peptide interaction or through non-target binding to peptides or the cellulose support. This is especially true for interactions of μM–mM K_D's, where the strength of "weak" non-specific interactions may be on par with that of target interactions. This was supported by a high false positive rate arising from assays where CA was conjugated with hydrophobic fluorescent dyes as well as indirect detection assays using fluorescently labeled antibodies (data not shown). In comparison, antibody binding controls incorporated with indirect chemiluminescent detection showed no significant non-specific antibody binding to peptides, and hence control antibody binding subtraction was not performed in subsequent analyses.

While direct labeling with a chemiluminescent agent may not have resulted in false positives as well, versatility of the microarray assay would be reduced; a glycan target would require chemical ligation with an agent, such as HRP, in contrast to indirect detection, where only availability of target antibodies is limiting. While commercial availability of detection agents may be a drawback for other glycans of interest (and is itself an incentive for development of CBPs), the advantage of the ELISA-like microarray methodology described is flexibility in format and detection techniques. For instance, assay format may be optimized based on commercially available reagents and/or target properties (e.g., biochemical properties linked to detrimental effects from a specific labeling technique). Thus, for PSA, indirect chemiluminescence-based detection of PSA–peptide interaction minimized the likelihood of false positives while increasing assay versatility.

Previous studies have approached the problem of carbohydrate affinity primarily with multivalent presentation of either the peptide or carbohydrate ligands.[5,49–52] Here, a similar approach is undertaken with microarray presentation of peptides. The three-dimensional presentation of peptides in high density on immobilized cellulose fibers[45–47] provides a platform for selecting interactions of highest affinity and selectivity,[53] which may then be studied in low-throughput experimental or computational platforms (as for thermodynamic, kinetic, or structural analysis). Two synthesis and screening control peptides incorporated with the peptide library exemplify this approach (ESI Table 3 and ESI Fig. 7), wherein microarray binding of a positive control and non-binding of a negative control are corroborated by low-throughput analysis with SPR spectroscopy.

In order to determine whether printing-based density effects significantly altered peptide binding outcomes, three printing conditions were examined (Fig. 1). As can be seen in Fig. 1A, intra-assay comparisons of peptide intensities show strong or very strong correlations, suggesting that a peptide binder with one printing condition is likely to be selected as a binder in another. Even with uniform occurrence of reactive functional groups across cellulose discs on which peptides are synthesized, it is possible that peptide-to-peptide differences exist in peptide density on cellulose fibers due to differences in amino acid coupling yields. However, variation of peptide density through altered printing conditions allows for preliminary assessment of density effects on binding for individual peptides. Furthermore, the intra-assay comparisons demonstrate that variability in microarray printing (common for physical adsorption-based methods) does not significantly affect discrimination of binding peptides from non-binding peptides. Thus, the microarray assay provides flexibility of printing conditions in lead peptide candidate selections.

Nevertheless, it is evident that some variation exists amongst these conditions, primarily in higher intensity binding with 2x and 1x2 conditions; inter-assay correlations for the sub-set of peptides with above background binding are moderate or

strong. For the top 30% of intensities, in comparison to 1x condition, spot area increased by 170% for the 2x condition, while increasing 140% for the 1x2 condition. At the same time, mean spot intensity per unit area increased by 110% for the 2x condition and by 130% for the 1x2 condition. This suggests that the higher intensities for either 2x or 1x2 conditions are indeed due to higher peptide densities. The nature of this increase in density is likely complex, relating to peptide drying or deposition on the slide surface and/or possible multivalent binding effects with density differences; further investigations on molecular binding mechanisms are outside the scope of this study.

Inter-assay variability, in part arising from the above density effects, was considered in the identification of high "affinity" (see subsequent discussion) PSA-binding peptides (ESI Fig. 1 and ESI Table 1). The highest binding peptides in screening against CA with no competitor all displayed low coefficients of variation (CVs; <25%). In comparison, in all screens, at least 90% of peptides with intensities above background (i.e., in the top 80% of experimental intensity range) had intra-assay CVs below 20%. Also, the intra-assay CVs for 86–100% of peptides with top 10% binding intensities in different screens were <10%, demonstrating good reproducibility of peptide binding within and between different microarray assays.

Affinity and selectivity of PSA-binding peptides

The peptide library was subjected to a screening assay against CA at a single concentration (10 μM) to identify high "affinity" PSA-binding peptides (Table 1). As noted,§ affinity in this context refers to screening in the absence of competing macromolecules (in contrast to selectivity screening) and does not indicate specific K_D's. Binding analysis for thermodynamic characterization of peptide–carbohydrate binding, through screening at multiple CA concentrations (for construction of a ligand binding curve), was not possible; use of higher CA concentrations required for determination of μM–mM K_D's was limited due to the solubility limit of CA in water-based buffers (approximately 1.67 mM), increased viscosity of CA solution at higher concentrations, and poor slide background at very high CA concentrations. Additionally, any determination of binding affinity constants through microarrays would likely be high as compared to methods where 1:1 stoichiometry can be ensured. In fact, enhancement of multivalent interactions through clustered presentation of immobilized peptides has been shown to increase apparent K_D.[51,54,55]

As evident from Table 1, peptides with the highest binding to CA in the absence of a competing glycan originate from both mAb epitope mapping and phage display screening. The incidence of phage display peptides with mutations to positively charged residues is expected given repeating carboxyl groups on the sialic acid polymer. Similarly, all high binding peptides contained positively charged residues. To determine whether occurrence of certain residues in higher binding peptides was likely even without bias introduced with rational mutations, compositional and positional analyses were undertaken on unmodified sequences in the peptide library (Fig. 2) in a method adapted from Butterfield *et al*. and others.[26–28] As expected, Fig. 2A shows a statistically significant increase in basic residues in PSA-binding sequences and a statistically significant decrease in acidic residues. Interestingly, both polar and nonpolar residues display significant decreases as well, though propensities of residues within categories vary (ESI Fig. 2). For example, there is an increase in phenylalanine propensity, which may be due to CH–pi interactions with sialic acid.[56] Polar, hydrophobic, and aromatic residues have all been implicated in carbohydrate–protein binding, and studies on characterization of binding of carbohydrates by peptides,[7] though more limited, have arrived at similar conclusions.[2,21,28] Definition of such binding generalities for PSA-binding peptides, beginning with compositional residue propensities, serves as the first step in establishing a molecular understanding of PSA binding, which can be used to improve upon peptide design.

Positional analysis (Fig. 2B) of PSA-binding peptides can similarly aid in peptide characterization. Analysis of changes in basic residues (R and K), which display the greatest increase, shows the largest increases in lysines in the middle of sequences and arginines in the middle and C-terminal end (the greater increases for lysine may be due to higher positive charge density). It is possible that N-terminal presentation of peptides on cellulose fibers on the microarrays influences peptide binding, but projection of peptides from the matrix may minimize this effect. Increases in R and K frequencies at nearly every two positions suggests the existence of a binding motif for PSA, where peptide primary sequence and subsequent conformation (i.e., final geometric presentation) play an important role. This agrees with previous studies where electrostatic interactions, especially from positively charged groups, have been shown to serve as "anchors" that guide binding affinity, with other residues contributing to affinity and selectivity.[7,9,28,57,58] It is possible that PSA-non-binding peptides of acidic or highly hydrophobic nature present unfavorable enthalpic contributions to binding, while non-binding peptides containing positive charge lack the necessary binding motif (or flexibility required to orient and mimic such a motif). As suggested in prior work,[27,28] it is possible that the residue occurrence and positioning data uncovered with peptide binding studies serve as preliminary evidence, though tangential, for the understanding of naturally occurring interactions (here, for biological lectin–PSA interactions).

The increased incidence of basic residues at certain positions within high binding peptides suggests that peptides do not bind the negatively charged target indiscriminately, but demonstrate specific target binding. Importantly, qualitative (Fig. 1C and Fig. 3) and quantitative (Tables 1–2) descriptions of selectivity show clear evidence for specific and selective target binding by some, if not all, high "affinity" PSA-binding peptides (here, "specific" denotes binding to the target over "non-specific" interactions with background components such as the

microarray surface, with no other biomolecules present, and "selective" denotes the ability to bind the target in the presence of non-target, non-buffer species).

The selective binding of high affinity peptides to PSA in the presence of the competing glycan chondroitin sulfate A (CS-A) varies dramatically (Table 1 and Fig. 3; selectivities of high affinity peptides with 2x and 1x2 screening shown in ESI Fig. 3). CS-A serves as a negatively charged polymer with both similar and distinct functional groups to PSA, and thus is a good exemplary competitor for study of peptide selective binding. Like PSA, CS-A is a polydisperse polysaccharide playing important roles in the mammalian nervous system; in contrast to PSA, CS-A prevents plasticity through inhibition of axonal growth and remyelination.[59,60] The applicability of PSA-binding peptides may be envisioned with selective peptide–PSA binding in a therapeutic system for nervous system regeneration. The differing selectivity of high-binding peptides to CS-A emphasizes the importance of selectivity assessment in early stages of peptide development and why selectivity cannot be assumed based on high affinity, library size,[20] or peptide origin.

The quantification of peptide selectivity in the context of a fraction of competitor (10%), as opposed to equimolar amounts, served two purposes: i) relative assessment of selectivity of high affinity peptides, none of which showed significant binding to CA with equimolar CS-A and ii) evaluation of a larger range of peptides (few peptides displayed binding with equimolar CS-A). Of high affinity PSA-binding peptides, only peptides 79 and 214 demonstrate >80% selectivity by this definition regardless of assay, thus serving as binders displaying both affinity and an appreciable degree of selectivity to the target. Comparatively, PSA-binding peptides that demonstrated >80% selectivity regardless of affinity (i.e., binding intensity with CA alone) are given in Table 2 and ESI Table 2. Peptides 36 and 214 are noteworthy in that they display consistently higher selectivities; in fact, peptide 36, despite its higher variability between affinity screens and generally lower affinity, was one of the few peptides displaying above background binding with equimolar CS-A. Such peptides, displaying binding characteristics that may not be identified solely through assessment of high affinity, may be evaluated for direct binding to one or more competitors of interest through low-throughput platforms following lead selection. Considering that a peptide's K_D and selectivity may display a different relationship than affinity intensity and selectivity, selectivity evaluation thus ensures that peptides with only low or moderate affinity binding intensity to CA are not discounted for study.

Peptide design strategies

Peptides of higher affinity and selectivity originated from both phage display screening and mAb epitope mapping. Here, a smaller library with peptides from exemplary sources demonstrated: i) use of a previously elucidated carbohydrate–protein interaction as a reference for peptide–target binding, ii) the ability of the screening methodology to rapidly verify random screening leads, and iii) potential advantages in combining random and rational screening, such as initial large library screening combined with a hypothesis-driven approach. Furthermore, multiple sources and strategies may have offered a higher probability of peptide identification for a target for which *a priori* binding knowledge is minimal.

Epitope mapping is a strategy commonly employed for the generation of peptide ligands against both protein and carbohydrate targets. mAb735 is an scFv that interacts discontinuously with oligosialic acid. Linear epitope mapping for this interaction is thus not intended to "map" the true epitope, as with discontinuous epitope mapping, but may provide peptides from small linear and conformational epitopes from continuous portions of the protein interaction (ESI Fig. 5). For example, despite the critical contribution of a number of tyrosine residues for octasialic acid-binding of mAb735,[42] PSA-binding peptides display neither a higher propensity of tyrosine (ESI Fig. 2) nor do the majority of peptides containing the corresponding "mAb-critical" tyrosine residues demonstrate high PSA affinity. In contrast, peptide 79 and similar sequences demonstrate higher PSA binding. Though many corresponding residues from these peptides also do not display direct ligand binding in the mAb, the corresponding mAb sequence occurs with local continuity in close proximity to the binding site and potentially promotes ligand binding through enhancing local basicity. Thus, in peptide form, a sequence serves as a novel entity in entropic and conformational aspects, but the residue composition in lectins (carbohydrate-binding proteins) may inherently provide peptides with a higher probability of carbohydrate interaction (both affinity and selectivity),[61] even with regard to antibody sequences, for which sequence diversity may be higher.[62]

While phage display screening has been applied for diverse targets, including biomolecules, synthetic materials, and metals,[63] this report is the first to use this technique in the targeting of PSA. A number of phage display peptides demonstrated lack of binding to CA on microarrays, which was not unexpected given the false positive rate (from target unrelated peptides) with this screening technique.[64] Unlike with phage ELISAs, which verify peptide–target binding when peptides are presented on a phage particle, microarrays are able to verify peptide binding with alternative molecular context. The microarray platform also provides an opportunity to rapidly create and screen modified sequences from parent phage peptides while remaining *in vitro*. Fig. 4 demonstrates variation in affinity and selectivity of mutants from one phage display-derived peptide (sequences shown in Table 3; see ESI Fig. 4 for binding in 2x and 1x2 conditions). Evaluation of mutational results supports library-level conclusions on peptide properties conferring higher affinity and selectivity. For example, peptide 170, where a high density of positive charge was introduced, displays higher affinity binding but very poor selectivity. This supports residue occurrence data, which demonstrates importance of charge placement, and selectivity

data, in which highest affinity was not necessarily correlated to high selectivity.

Conclusions

In this study, PSA-binding peptides were identified in an integrated approach emphasizing flexibility of assay design when limited biochemical information is available on the target and its natural ligands. Peptide ligand characterization, as through amino acid compositional and positional analysis, suggested biochemical characteristics necessary for PSA-binding by peptide sequences, including the importance of positive charge positioning for PSA interactions. The discovery of sequence characteristics serves as the first step for enhancement of affinity and selectivity of novel ligands and also provides a resource for understanding the molecular basis of biological interactions involving PSA. Though rational strategies are common in ligand design, successful application of such strategies for polysaccharide-binding peptides is limited. Here, rational design utilizing well-known mutational approaches was shown to enhance hypothesis-driven ligand discovery for PSA-binding peptides.

The high-throughput assay developed directly incorporates a selectivity test. This goes in hand with potential applications for CBPs, where peptide selectivity to a glycan amongst many structurally complex, yet similar, glycans would prove advantageous and where preliminary assessment of selectivity along with affinity is strategic for ligand development. Given the amenability of the method to modifications, including to ligand density, peptide design, and selectivity screening, this approach can be adapted to other carbohydrates, particularly those for which rational ligand design is challenging, but for which random methodologies alone do not offer sufficient analytical capacity for fundamental understanding and iterative design.

Conflicts of interest

Rensselaer Polytechnic Institute has filed a patent application on peptide sequences described in this paper (U.S. Patent Application No. 62/879,176).

Acknowledgements

This work was supported by the National Science Foundation Graduate Research Fellowship (to DGS) under Grant No. DGE-1247271. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. The authors thank Dr. Patrick Maxwell and Dr. Jennifer Hurley (RPI) for use of the ChemiDoc™ XRS+ System for microarray and gel imaging. DGS thanks Dr. John P. Trasatti, Dr. Robert Linhardt (RPI), Dr. Steven Cramer (RPI), and Dr. Linda McGown (RPI) for critical input during conception and design of the work. DGS also thanks Rebecca Jackson, Sagar Shankar, Samuel Mayo, and Priyanka Raju for assistance with experiment preparation. Parts of this research were undertaken in the Cell and Molecular Biology Research Core and Microbiology Research Core in the Center for Biotechnology and Interdisciplinary Studies (RPI).

Notes and references

‡ In addition to the methodology discussed, assay development to enhance detection of carbohydrate–peptide interactions included standard screening optimization (involving optimization of antibody concentrations, reduction of non-specific antibody binding with BSA, optimization of incubation volumes, optimization of CA concentration, and improvement in signal detection).

§ The term affinity in this usage is not meant to indicate the derivation of thermodynamic parameters such as equilibrium binding constants from this type of screening, but refers to an assay to distinguish peptide binding capability to the target at a selected concentration in the absence of non-target species (excluding buffer components).

1 A. Varki, *Glycobiology*, 2017, **27**, 3.
2 J. Arnaud, A. Audfrey and A. Imberty, *Chem Soc. Rev.*, 2013, **42**, 4798.
3 T. Johannssen and B. Lepenies, *Trends Biotechnol.*, 2017, **35,** 334.
4 S. A. Krumm and K. J. Doores, *Curr. Top. Microbiol. Immunol.*, 2018, DOI: 10.1007/82_2018_103.
5 N. F. Reuel, B. Mu, J. Zhang, A. Hinckley and M. S. Strano, *Chem. Soc. Rev.*, 2011, **41**, 5744.
6 H.-J. Gabius, *Biosystems*, 2018, **164**, 102.
7 M. Ambrosi, N. R. Cameron and B. G. Davis, *Org. Biomol. Chem.*, 2005, **3**, 1593.
8 A. P. Davis, *Org. Biomol. Chem.*, 2009, **7**, 3629.
9 P. Kosma and S. Müller-Loennies, *Anticarbohydrate Antibodies: From Molecular Basis to Clinical Application*, Springer-Verlag/Wien, New York, NY, 2012.
10 W. Sun, L. Du and M. Li, *Curr. Pharm. Des.*, 2010, **16**, 2269.
11 Q. Zhang and R. Landgraf, *Pharmaceuticals*, 2012, **5**, 493.
12 S. A. Svarovsky and L. Joshi, *Curr. Drug Discovery Technol.*, 2008, **5**, 20.
13 M. Rauschenberg, S. Bandaru, M. P. Waller and B. J. Ravoo, 2014, **20**, 2770.
14 S. André, C. J. Arnusch, I. Kuwabara, R. Russwurm,H. Kaltner, H. J. Gabius and R. J. Pieters, *Bioorg. Med. Chem.*, 2005, **13**, 563.
15 T. J. Molenaar, C. C. Appeldoorn, S. A. de Haas, I. N. Michon, A. Bonnefoy, M. F. Hoylaerts, H. Pannekoek, T. J. van Berkel, J. Kuiper and E. A. Biessan, *Blood*, 2002, **100**, 3570.
16 K. Yamamoto, Y. Konami, T. Osawa and T. Irimura, J*. Biochem.*, 1992, **111**, 436.
17 L. A. Landon, J. Zou and S. L. Deutscher, *Mol. Divers.*, 2004, **8**, 35.
18 N. Röckendorf, M. Borschbach and A. Frey, *PLoS One*, 2012, **8**, DOI: 10.1371/journal.pcbi.1002800.
19 C. LI, X. Chen, F. Zhang, X. He, G. Feng, J. Liu and S. Wang, Anal. Chem., 2017, **89**, 10431.
20 K. W. Boltz, M. J. Gonzalez-Moa, P. Stafford, S. A. Johnston and S. A. Svarovsky, *Analyst*, 2009, **134**, 650.
21 C. Morales Betanzos, M. J. Gonzalez-Moa, K. W. Boltz, B. D. Vander Werf, S. A. Johnston and S. A. Svarovsky, *ChemBioChem*, 2009, **10**, 877.
22 S. A. Svarovsky and M. J. Gonzalez-Moa, *ACS Comb. Sci*., 2011, **13**, 634.
23 A. Verrecchio, M. W. Germann, B. P. Schick, B. Kung, T.

Twardowski and J. D. San Antonio, *J. Biol. Chem.*, 2000, **275**, 7701.
24 D. L. Rabenstein and J. Wang, *Biochemistry*, 2006, **45**, 15740.
25 J. Bae, U. R. Desai, A. Pervin, E. E. Caldwell, J. M. Weiler and R. J. Linhardt, *Biochem. J.*, 1994, **301**, 121.
26 E. E. Caldwell, V. D. Nadkarni, J. R. Fromm, R. J. Linhardt and J. M. Weiler, *Int. J. Biochem. Cell Biol.*, 1996, **28**, 203.
27 J. R. Fromm, R. E. Hileman, E. E. Caldwell, J. M. Weiler and R. J. Linhardt, *Arch. Biochem. Biophys.*, 1997, **343**, 92.
28 K. C. Butterfield, M. Caplan and A. Panitch, *Biochemistry*, 2010, **49**, 1549.
29 K. C. Butterfield, A. Conovaloff, M. Caplan and A. Panitch, *Neurosci. Lett.*, 2010, **478**, 82.
30 G. Loers, Y. Liao Y, C. Hu, W. Xue, H. Shen, W. Zhao and M Schachner, *Sci. Rep.*, 2019, **9**, DOI: 10.1038/s41598-018-37685-2.
31 H.-C. Siebert, S.-Y. Lu, R. Wechselberger, K. Born, T. Eckert, S. Liang, C.-W. von der Lieth, J. Jiménez-Barbero, R. Schauer, J. Vliegenthart, T. Lütteke, S. André, H. Kaltner, H.-J. Gabius and T. Kozár, *Carbohydr. Res.*, 2009, **344**, 1515
32 P. M. Chaudhary, R. V. Murthy, R. Yadav and R. A. Kikkeri, *Chem. Commun.*, 2015, **51**, 8112
33 L. D. Heerze, R. H. Smith, N. Wang and G. D. Armstrong, *Glycobiology*, 1995, **5**, 427.
34 Y. Xiong, M. Li, Q. Lu, G. Qing and T. Sun, *Polymers*, 2017, **9**, DOI: 10.3390/polym9070249.
35 Q. Lu, M. Zhan, L. Deng, G. Qing and T. Sun, *Analyst*, 2017, **142**, 3564.
36 T. Janus and T. Janus, *Biochim. Biophys. Acta*, 2011, **1808**, 2923.
37 C. Sato and K. Kitajima, *J. Biochem.*, 2013, **154**, 115.
38 H. F. Azurmendi, J. Vionnet, L. Wrightson, L. B. Trinh, J. Shiloach and D. I. Freedberg, *Proc. Natl. Acad. Sci. U.S.A.*, 2007, **104**, 11557.
39 New England BioLabs, Inc., Ph.D.™ Phage Display Libraries Instruction Manual, 2016.
40 Y. Haile, K. Haastert, K. Cesnulevicius, K. Stummeyer, M. Timmer, S. Berski, G. Dräger, R. Gerardy-Schahn and C. Grothe, *Biomaterials*, 2007, **28**, 1163.
41 F. Liners, W. Helbert and P. Van Cutsem P, *Glycobiology*, 2005, **15**, 849.
42 M. Nagae, A. Ikeda, M. Hane, S. Hanashima, K. Kitajim, C. Sato and Y. Yamaguchi, *J. Biol. Chem.*, 2013, **288**, 33784.
43 T. Angata, S. C. Kerr, D. R. Greaves, N. M. Varki, P. R. Crocker and A. Varki, *J. Biol. Chem.*, 2002, **277**, 24466.
44 J. P. Trasatti, J. Woo, A. Ladiwala, S. Cramer and P. Karande, *Biotechnol. Prog.*, 2018, **34**, 987.
45 K. Hilpert, D. F. H. Winkler and R. E. W. Hancock, *Nat. Protoc.*, 2007, **2**, 1333.
46 C. Katz, L. Levy-Beladev, S. Rotem-Bamberger, T. Rito, S. G. D. Rüdiger and A. Friedler, *Chem. Soc. Rev.*, 2011, **40**, 2131.
47 INTAVIS Bioanalytical Instruments AG, CelluSpots™ Peptide Arrays, Synthesis and Work-Up – A Short Manual.
48 K. Sachse, K. S. Rahman, C. Schnee, E. Müller, M. Peisker, T. Schumacher, E. Schubert, A. Ruettger, B. Kaltenboeck and R. Ehricht, *Sci. Rep.*, 2018, **8**, DOI: http://10.1038/s41598-018-23118-7.
49 S. Hyun, E. H. Lee, J. Park and J. Yu, *Bioorg. Med. Chem. Lett.*, 2008, **18**, 4011.
50 S. Hyun, J. Kim, M. Kwon and J. Yu, *Bioorg. Med. Chem.*, 2007, **15**, 511.
51 M. Kwon, S. Jeong, K. H. Lee, Y.-K. Park, and J. Yu, *J. Am. Chem. Soc.*, 2002, **124**, 13996.
52 H. Connaris, P. R. Crocker and G. L. Taylor, *J. Biol. Chem.*, 2009, **284**, 7339.
53 N. Röckendorf, S. Bade, T.-R. Hirst, H. H. Gorris and A. Frey, *Bioconj. Chem.*, 2007, **18**, 573.
54 J. J. Lundquist and E. J. Toone, *Chem. Rev.*, 2002, **102**, 555.
55 M. Mammen, S.-K. Choi and G. M. Whitesides, *Angew. Chem. Int. Ed.*, 1998, **37**, 2754.
56 K .L. Hudson, G. J. Bartlett, R. C. Diehl, J. Agirre, T. Gallagher, L. L. Kiessling and D. N. Woolfson, J. *Am. Chem. Soc.*, 2015, **137**, 15152.
57 H. J. Gabius, S. André, J. Jiménez-Barbero, A. Romero and D. Solís, *Trends Biochem. Sci.*, 2011, **36**, 298.
58 J. M. Pröpster, F. Yang, S. Rabbani, B. Ernst, F. H. Allain and M. Schubert, Proc. Natl. Acad. Sci. U. S. A., 2016, **113**, E4170.
59 C. M. Galtrey and J. W. Fawcett, *Brain Res. Rev.*, 2007, **54**, 1.
60 M. B. Keough, J. A. Rogers, P. Zhang, S. K. Jensen, E. L. Stephenson, T. Chen, M. G. Hurlbert, L. W. Lau, K. S. Rawji, J. R. Plemel, M. Koch, C. Ling and V. W. Yong, *Nature Commun.*, 2016, **7**, 1.
61 D. J. Brooks, J. R. Fresco, A. M. Lesk and M. Singh, *Mol. Biol. Evol.*, 2002, **19**, 1645.
62 M. Guharoy and P. Chakrabarti, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 15447.
63 F. C. Dudak, I. H. Boyaci and B. P. Orner, *Molecules*, 2011, **16**, 774.
64 C.-H. Wu, I-J. Liu, R.-M. Lu and H.-C. Wu, *J. Biomed. Sci.*, 2016, **23**, DOI: 10.1186/s12929-016-0223-x.
65 B. Bakhshinejad, H. M. Zade, H. S. Shekarabi and S. Neman, *Amino Acids*, 2016, **48**, 2699.

Table 1 High affinity PSA-binding peptides, that is, peptides with the top 5% of binding intensities in 1x, 2x, and 1x2 screening against CA. Percent selectivities of peptides in different assays provided with standard deviations of triplicate inter-assay measurements.

			% Selectivity		
Peptide index	Peptide sequence	Peptide origin	1x	2x	1x2
21	YLQKPGQSPKPLIYR	mAb735	52 ± 6	48 ± 3	23 ± 1
23	PGQSPKPLIYRVSNR	mAb735	27 ± 6	33 ± 4	24 ± 6
28	RVSNRFSGVPDRFSG	mAb735	45 ± 8	30 ± 1	21 ± 3
35	GSGSGTDFTLKISRV	mAb735	37 ± 5	40 ± 9	34 ± 2
78	SGNTKYNEKFKGKAT	mAb735	63 ± 23	53 ± 11	27 ± 2
79	NTKYNEKFKGKATLT	mAb735	91 ± 5	97 ± 5	49 ± 6
161	TLPYILQSSGTRGGGS	Phage display screening; A4Y mutation	60 ± 4	39 ± 2	31 ± 3
170	TLERGSRVRQSSGTRG	Phage display screening; PAIL to ERGSRVR mutation and C-terminal GGS deletion	30 ± 6	35 ± 2	17 ± 2
203	KISSPLLWNPFRGGGS	Phage display screening; A1K mutation	67 ± 14	44 ± 4	34 ± 5
213	AISSPLLRNPFRGGGS	Phage display screening; W8R mutation	55 ± 19	89 ± 5	100 ± 3
214	AISSPLLKNPFRGGGS	Phage display screening; W8K mutation	86 ± 4	106 ± 4	93.6 ± 4.0

Table 2 Selective peptides from 1x condition. Peptides with % selectivity greater than 80% are shown with corresponding inter-assay average intensities from affinity screening (intensities in relative intensity units). Only peptides with intensities above background levels (bottom 20% of intensity range of a screen; approximately 70% of the library) were considered in selectivity assessment (intensity range = 11-94 [non-binding < 27]). * Represented in Table 1 in list of high affinity PSA-binding peptides.

Peptide index	Peptide sequence	Peptide origin	Average 1x affinity intensity	% Selectivity
36	GSGTDFTLKISRVEA	mAb735	59 ± 5	86 ± 10
50	VPYTFGGGTRLEIKG	mAb735	85 ± 5	89 ± 7
79*	NTKYNEKFKGKATLT	mAb735	85 ± 5	91 ± 5
147	TLPAILQAAGTRGGGS	Phage display screening; S8A and S9A mutations	50 ± 6	109 ± 18
150	RLPAILQSSGTRGGGS	Phage display screening; T1R mutation	70 ± 3	80 ± 9
166	TLPAILRSSGTRGGGS	Phage display screening; Q7R mutation	66 ± 4	84 ± 9
167	TLPAILKSSGTRGGGS	Phage display screening; Q7K mutation	68 ± 7	86 ± 9
214*	AISSPLLKNPFRGGGS	Phage display screening; W8K mutation	82 ± 2	86 ± 4

Table 3 Mutations to peptide 106 (P106; from phage display screening). * Represented in Table 1 in list of high affinity PSA-binding peptides. ** Represented in Table 2 in list of high selectivity PSA-binding peptides.

Peptide index	P106 series index	Peptide sequence	Mutation(s)
106	1	TLPAILQSSGTRGGGS	N/A
143	2	TLPAILQ**T**SGTRGGGS	S8T
144	3	TLPAILQS**T**GTRGGGS	S9T
145	4	TLPAILQ**A**SGTRGGGS	S8A
146	5	TLPAILQS**A**GTRGGGS	S9A
147**	6	TLPAILQ**AA**GTRGGGS	S8A & S9A
148	7	TLPAILQSSGT**K**GGGS	R12K
149	8	**S**LPAILQSSGTRGGGS	T1S
150**	9	**R**LPAILQSSGTRGGGS	T1R
151	10	**K**LPAILQSSGTRGGGS	T1K
152	11	**D**LPAILQSSGTRGGGS	T1D
153	12	**E**LPAILQSSGTRGGGS	T1E
154	13	TLP**R**ILQSSGTRGGGS	A4R
155	14	TLP**K**ILQSSGTRGGGS	A4K
156	15	TLP**D**ILQSSGTRGGGS	A4D
157	16	TLP**E**ILQSSGTRGGGS	A4E
158	17	TLP**H**ILQSSGTRGGGS	A4H
159	18	TLP**S**ILQSSGTRGGGS	A4S
160	19	TLP**T**ILQSSGTRGGGS	A4T
161*	20	TLP**Y**ILQSSGTRGGGS	A4Y
162	21	TLP**W**ILQSSGTRGGGS	A4W
163	22	TLPAIL**N**SSGTRGGGS	Q7N
164	23	TLPAIL**A**SSGTRGGGS	Q7A
165	24	TLPAIL**H**SSGTRGGGS	Q7H
166**	25	TLPAIL**R**SSGTRGGGS	Q7R
167**	26	TLPAIL**K**SSGTRGGGS	Q7K
168	27	TLPAIL**D**SSGTRGGGS	Q7D
169	28	TLPAIL**E**SSGTRGGGS	Q7E
170*	29	TL**ERGSRVR**QSSGTRG	PAIL to ERGSRVR; ΔGGS

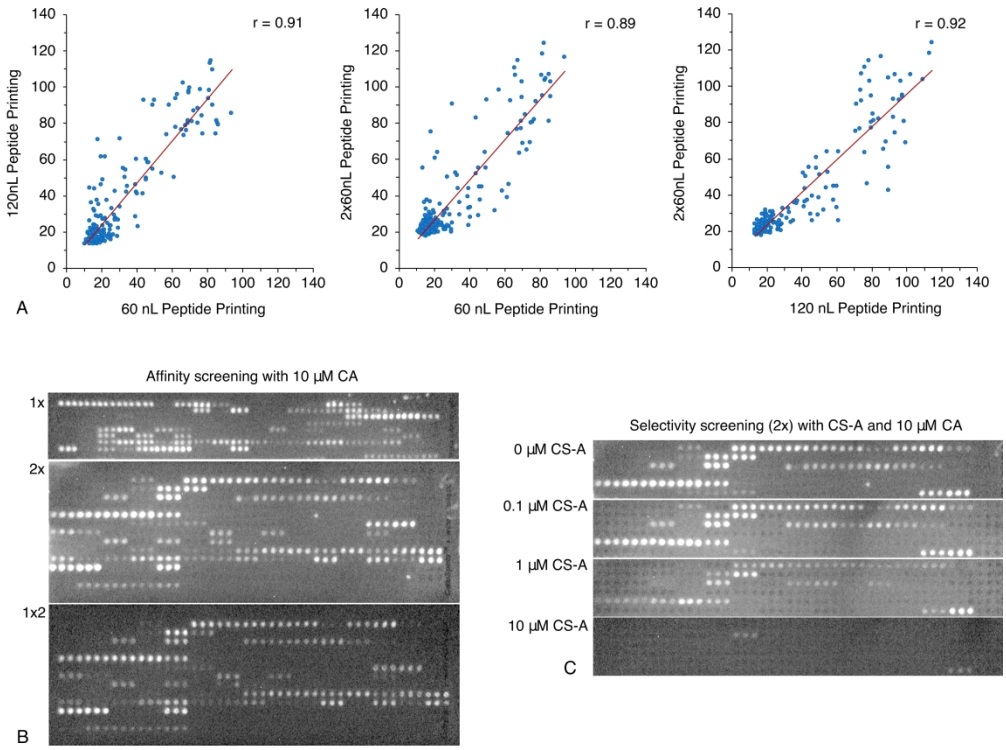


Fig. 1. Comparison of three printing conditions: 1x (60 nL), 2x (120 nL) and 1x2 (2x60 nL). (A) Correlation of average microarray spot intensities from triplicate intra-assay readings with different printing conditions. Raw intensity values gathered from microarray analysis software represent mean spot intensities per unit area in relative intensity units. For the sub-set of peptides above background (i.e., intensities in the top 80% of experimental intensity range), r = 0.69, 0.75, and 0.70, for left, middle, and right correlations, respectively. (B–C) Chemiluminescent images of representative peptide microarrays from: (B) affinity screening with 1x (top), 2x (middle), and 1x2 (bottom) printing and (C) selectivity screening for 2x printing with 0, 1, 10, and 100% (top to bottom) molar concentration of chondroitin sulfate A (CS-A) with 10 µM colominic acid (CA). Polysialic acid (PSA)-binding peptides appear white in affinity screening, and PSA-binding peptides with at least some degree of selective binding to CA over CS-A appear white with increasing amounts of CS-A in selectivity screening (each unique peptide appears in triplicate). Microarray images prior to any background modification for analysis are displayed. Since most peptides shown do not show selective binding to CA in equimolar concentration of competitor, percent selectivity at 10% competitor concentration was determined to enable assessment of selectivity of a larger pool of PSA-binding peptides. However, some peptides did retain the ability to detect CA with equimolar CS-A (C, bottom image).

430x324mm (300 x 300 DPI)

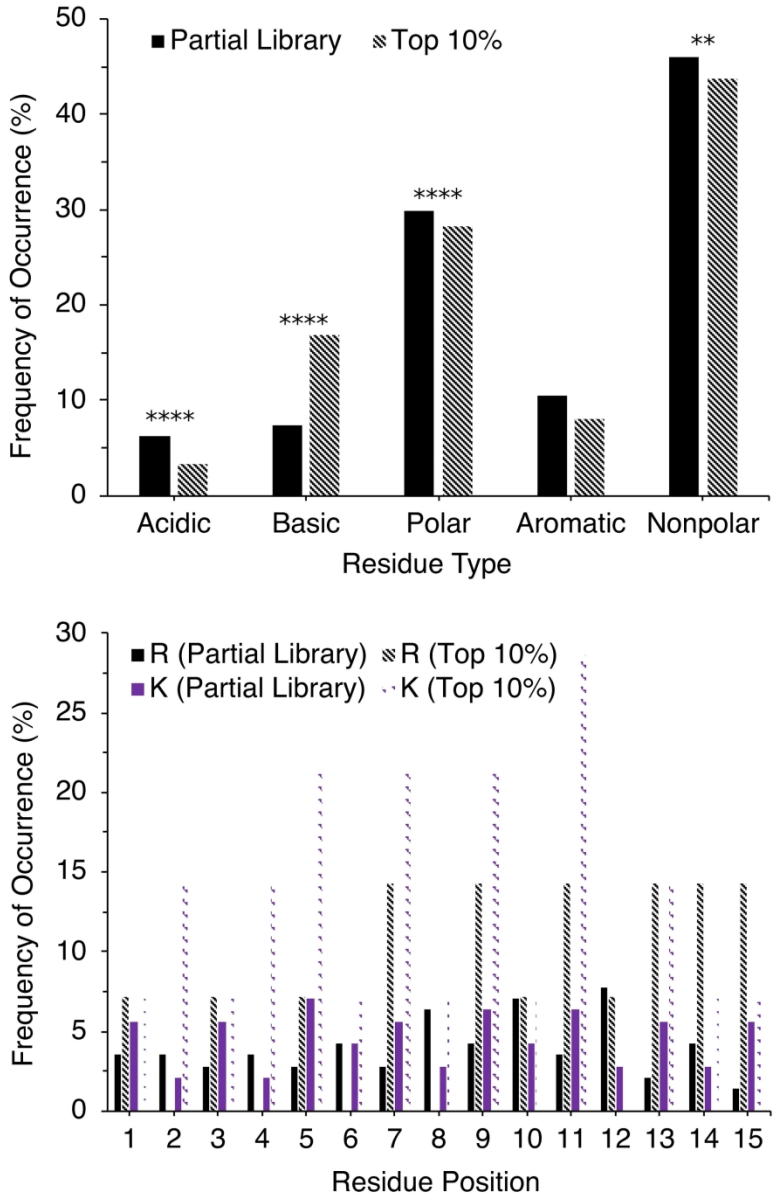


Fig. 2. Peptide residue (A) compositional and (B) positional analysis. For compositional analysis, occurrence of residue types in the partial peptide library (consisting of all mAb epitope mapping-derived peptides and unmodified phage display screening-derived peptides) is compared to occurrence in the top 10% of this peptide set, where top 10% refers to peptides with highest affinity binding in the partial library (14 peptides). Acidic = D and E; basic = R and K; polar = H, C, N, Q, and S; aromatic = Y, F, and W; and nonpolar = G, A, V, I, L, M, and P. (Two-tailed *z* test for population proportions; *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$). For positional analysis, changes in R and K occurrences (i.e., positional changes for basic residues, which display greatest change in compositional occurrence) at positions 1–15 are displayed for the same partial library and top 10% sub-set; position 16 excluded from plot as no library peptides contain these residues at position 16. Statistical significance of an increase or decrease in positional occurrences was not similarly determined as low residue occurrence in the sample population precluded the assumption of normal distribution.

178x272mm (300 x 300 DPI)

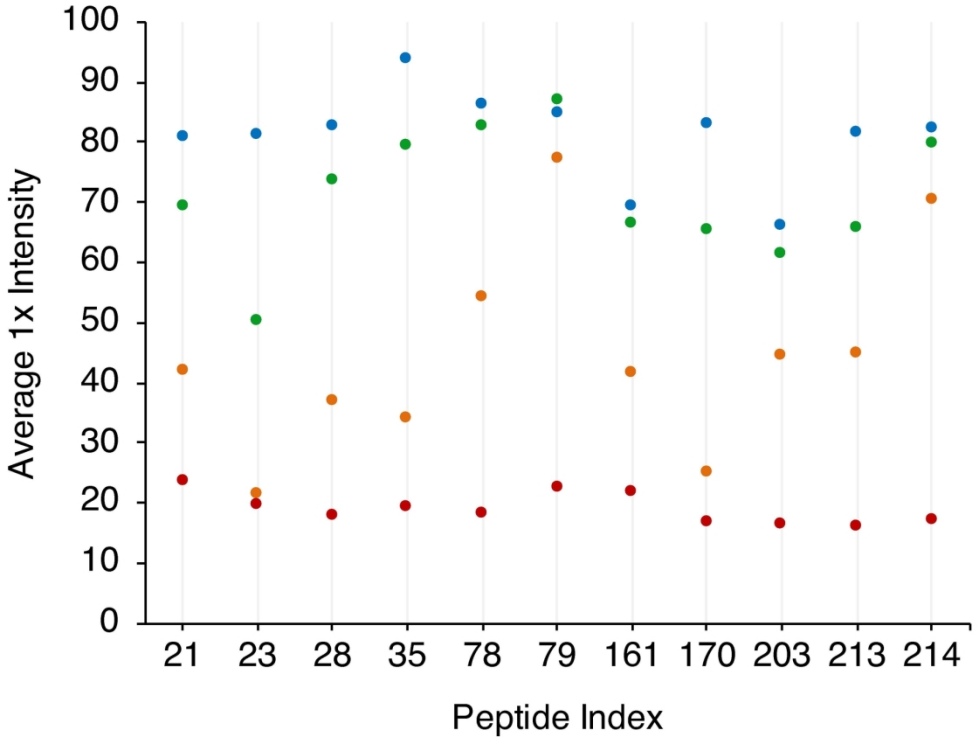


Fig. 3. Selectivity of high affinity peptides (Table 1) with 1x printing (intra-assay average intensities from triplicate measurements shown in relative intensity units). Blue circles = 10 μM CA with 0 μM CS-A, green circles = 10 μM CA with 0.1 μM CS-A, orange circles = 10 μM CA with 1 μM CS-A, and red circles = 10 μM CA with 10 μM CS-A. Peptides of similar binding intensities in affinity screening display different selectivities, that is, levels of binding ability in the presence of 10% competitor. Peptides 79 and 214 display relatively higher selective binding to CA in the presence of CS-A. No high affinity peptides show significant (i.e., above background) levels of binding to PSA with equimolar concentrations of CS-A.

162x122mm (300 x 300 DPI)

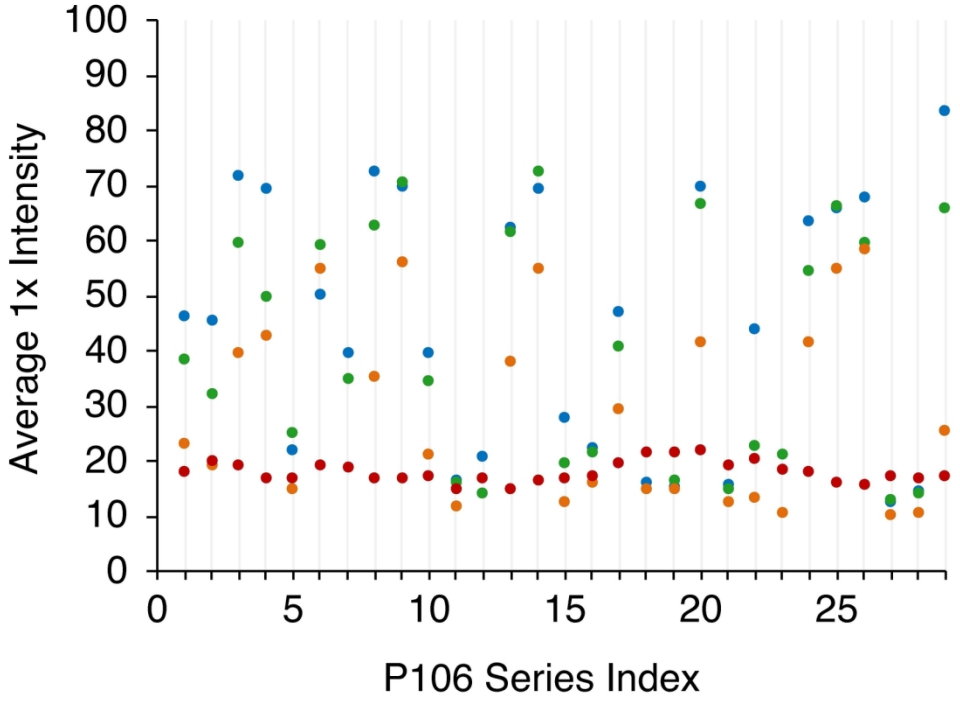
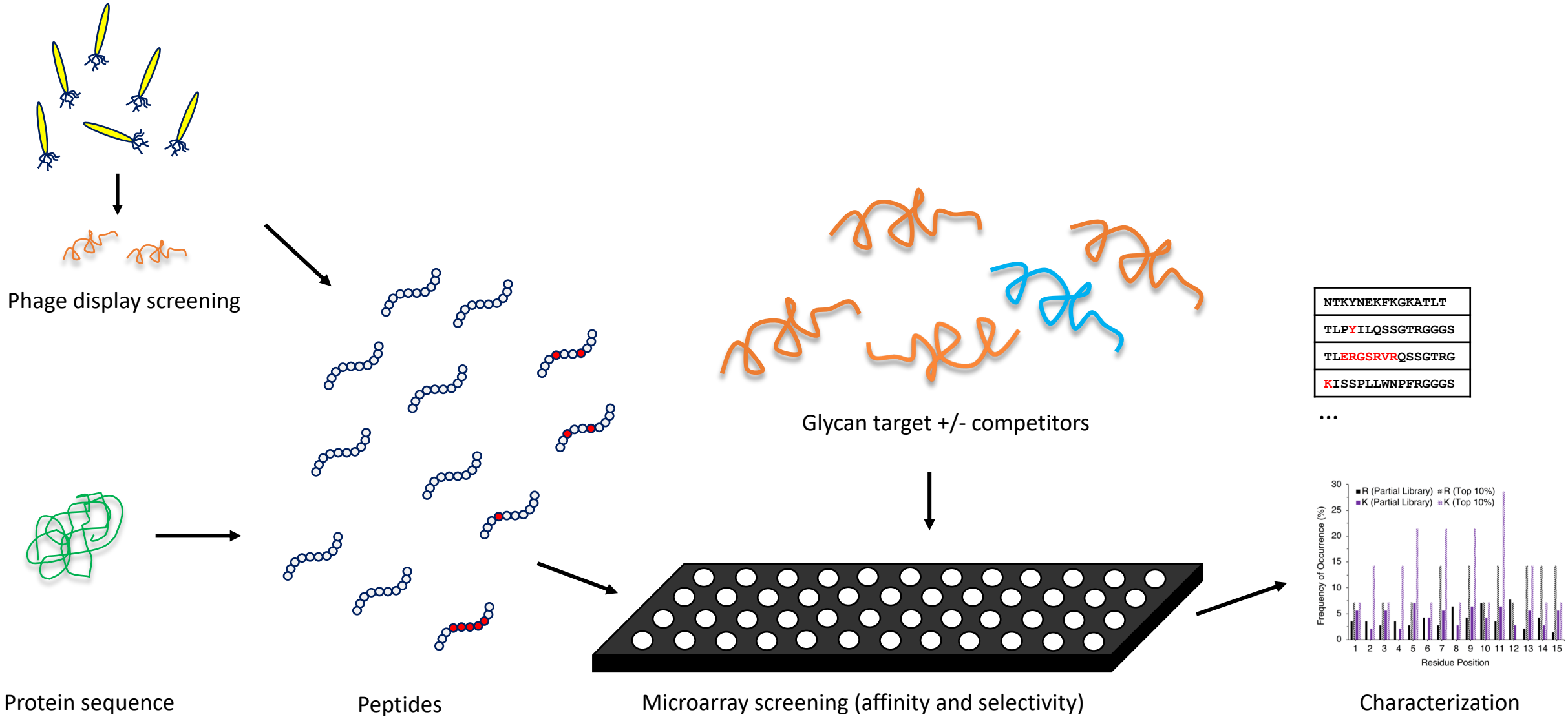


Fig. 4. Selectivity of peptides derived from peptide 106 (P106; phage display screening peptide) with 1x printing (intra-assay average intensities from triplicate measurements shown in relative intensity units). Blue circles = 10 μM CA with 0 μM CS-A, green circles = 10 μM CA with 0.1 μM CS-A, orange circles = 10 μM CA with 1 μM CS-A, and red circles = 10 μM CA with 10 μM CS-A. P106 series indices corresponding to primary peptide library indices are shown in Table 3. Peptide 29 (library index 170) displays high affinity to CA in the absence of competitor, possibly due to substitution of neutral residues with arginines, but has poorer selectivity as compared to other peptides in the series.

163x116mm (300 x 300 DPI)



An integrated approach for the identification of carbohydrate-binding peptides is described, with a focus on the unique glycan polysialic acid.