Electronic Supplementary Material (ESI) for Analyst.

Supplementary Information for:

Microarrays for the screening and identification of carbohydrate-binding peptides

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

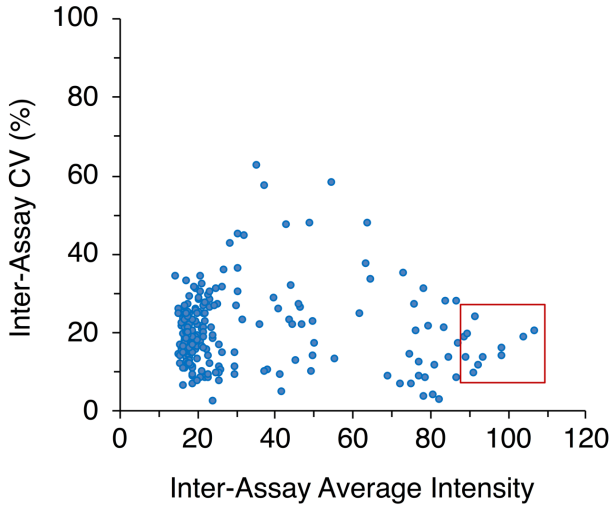
[a.] Department of Biological Sciences, Rensselaer Polytechnic Institute, Troy, NY 12180, USA. Email: shastd3@gmail.com

[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

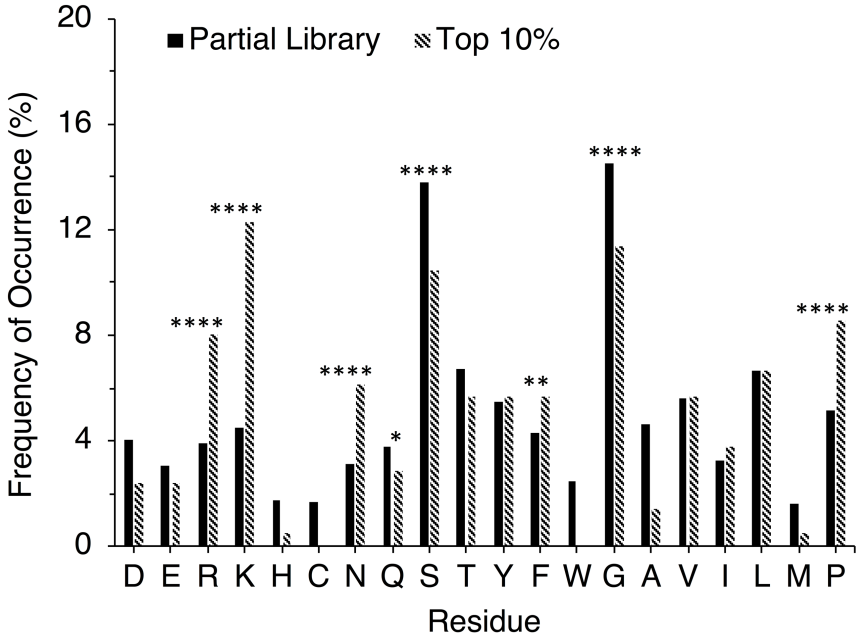
† Co-corresponding author



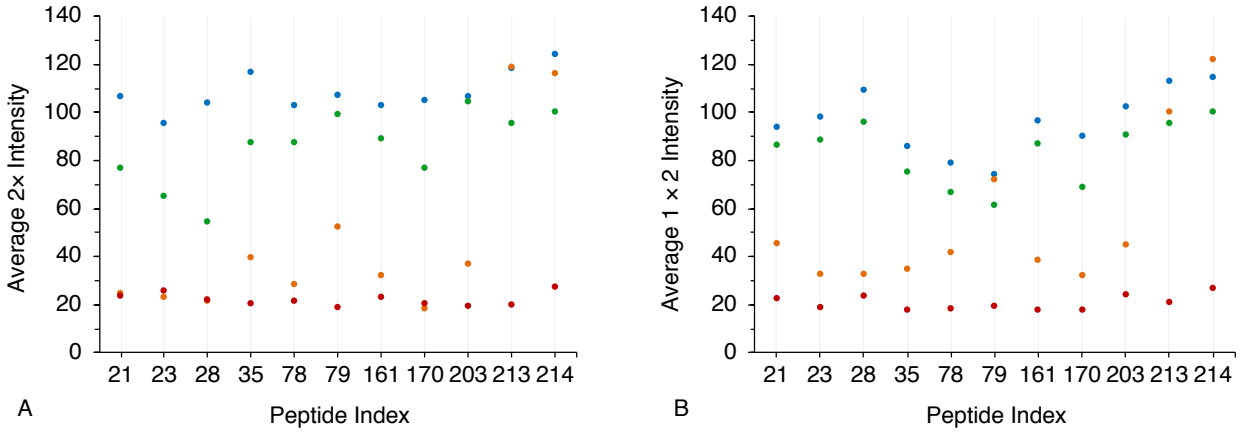
ESI Fig. 1 Comparison of variation in peptide intensity between printing conditions. Percent coefficient of variation (CV) of 1×, 2×, and 1 × 2 assays are displayed for corresponding inter-assay peptide intensities from affinity screening. Peptides within the red box represent the top 5% of peptides, identified as high affinity polysialic acid (PSA)-binding peptides (11 peptides total; Table 1 and ESI Table 1). Intensities are shown in relative intensity units.

ESI Table 1 Variation in peptide intensity of high affinity PSA-binding peptides between 1×, 2×, and 1 × 2 assays. High affinity PSA-binding peptides are defined as peptides with the top 5% of inter-assay average intensities (shown in red box in ESI Fig. 1). Percent coefficients of variation (CVs) are displayed with corresponding inter-assay peptide intensities (in relative intensity units). CVs of all 11 peptides are <25%.

Peptide index	Peptide sequence	Inter-assay average intensity	Inter-assay coefficient of variation (%)
21	YLQKPGQSPKPLIYR	94	14
23	PGQSPKPLIYRVSNR	91	10
28	RVSNRFSGVPDRFSG	99	14
35	GSGSGTDFTLKISRV	99	16
78	SGNTKYNEKFKGKAT	89	14
79	NTKYNEKFKGKATLT	89	19
161	TLPYILQSSGTRGGGS	90	20
170	TLERGSRVRQSSGTRG	93	12
203	KISSPLLWNPFRGGGS	92	24
213	AISSPLLRNPFRGGGS	104	19
214	AISSPLLKNPFRGGGS	107	20



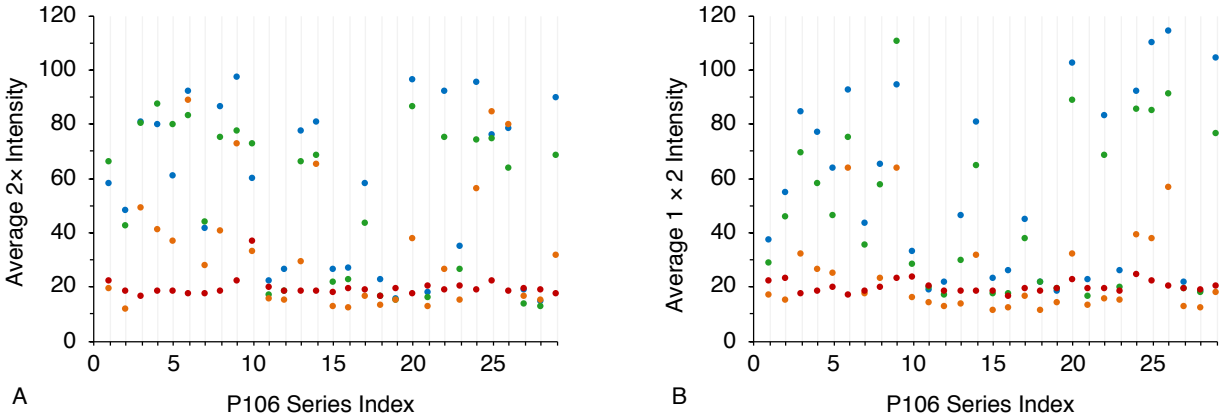
ESI Fig. 2 Peptide residue compositional analysis. Residue occurrence in the partial peptide library (consisting of all mAb epitope mapping-derived peptides and unmodified phage display screening-derived peptides) is compared to occurrence in the top 10% of this peptide set, where top 10% refers to peptides with highest affinity binding in the partial library (14 peptides). Statistical significance of an increase or decrease in occurrence of the following residues could not be determined as low residue occurrence in the sample population precluded the assumption of normal distribution: D, E, H, C, W, A, I, and M. In contrast, decrease in T and L and increase in Y and V were deemed not significant. (Two-tailed *z* test for population proportions; *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$). The greatest relative change in occurrence is evident for basic amino acids R and K (positively charged at the screening pH of 7.4), followed closely by N. It is possible that the changes seen for S and G are influenced by biases from phage display screening, as well as simultaneous evaluation of phage-based and other peptides with the microarray platform. Specifically, the high initial frequencies of G and S arise from C-terminal linkers included to mimic phage display conditions, and decreases in these frequencies are expected if high binders are not assumed to arise solely from phage display-derived peptides.



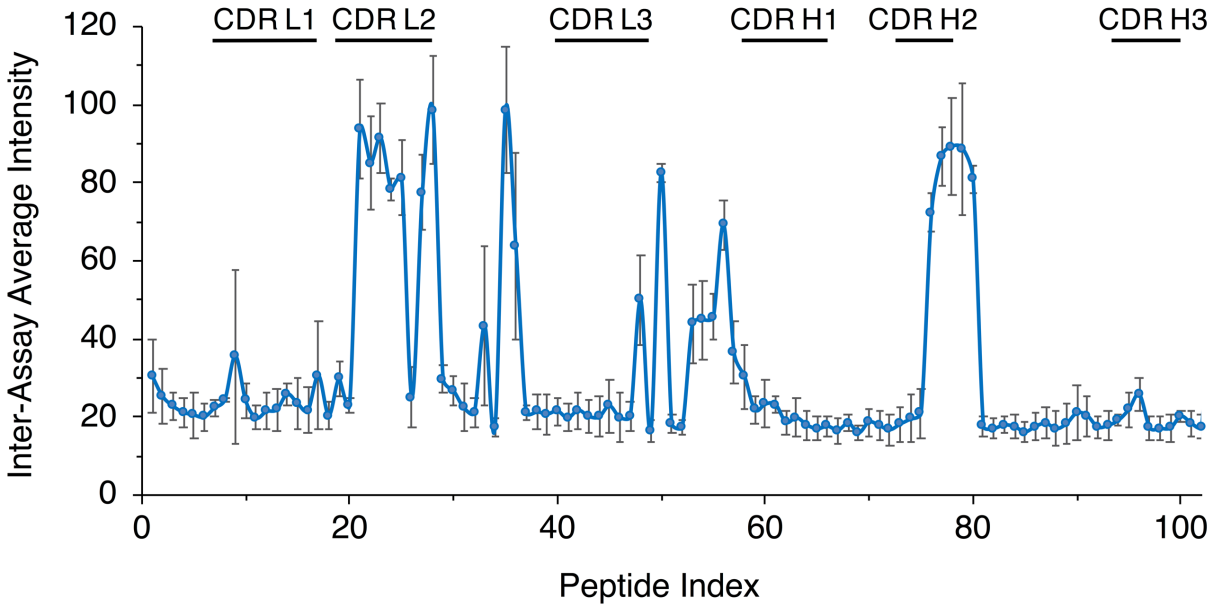
ESI Fig. 3 Selectivity of high affinity peptides (Table 1) with (A) 2× and (B) 1 × 2 printing (intra-assay average intensities from triplicate measurements shown in relative intensity units). Blue circles = 10 μM colominic acid (CA) with 0 μM chondroitin sulfate A (CS-A), green circles = 10 μM CA with 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. As with the 1× condition (Fig. 3), peptides of similar binding intensities in affinity screening display different levels of binding ability in the presence of 10% competitor, with peptides 79 and 214 demonstrating relatively higher selectivities. No high affinity peptides show significant (i.e., above background) levels of binding to PSA with equimolar concentrations of CS-A. It is possible that the generally lower selectivity displayed by these 11 high affinity peptides in the 2× and 1 × 2 conditions as compared to 1× (Fig. 3) are due to higher peptide densities (ligand density has been shown to affect both affinity and selectivity of carbohydrate interactions[1]).

ESI Table 2 Selective peptides from 2× and 1 × 2 conditions. Peptides with % selectivity greater than 80% are shown with corresponding intra-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 70–80% of the library) were considered in selectivity assessment (intensity range in 2× affinity screening = 13–114 [non-binding < 33] and intensity range in 1 × 2 affinity screening = 18–124 [non-binding < 39]). As in Tables 1–2, percent selectivities at or above 100% generally fall within error, with higher error for peptides with lower binding intensities. However, it is possible that in these cases, the presence of another glycan promotes peptide binding to CA, especially with higher peptide densities. While the assay enables classification of selective peptides from non-selective peptides, it does not support discrimination of peptide binding ability to this extent. * Represented in Table 1 in list of high affinity PSA-binding peptides.

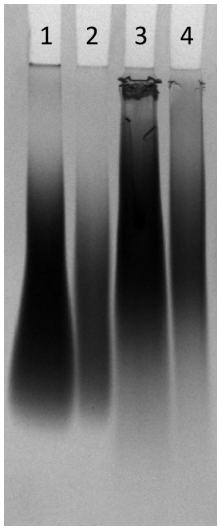
2×				
Peptide index	Peptide sequence	Peptide origin	Average 2× affinity intensity	% Selectivity
36	GSGTDFTLKISRVEA	mAb735	90 ± 2	90 ± 2
79*	NTKYNEKFKGKATLT	mAb735	74 ± 4	97 ± 5
104	HLSLKNPLRMDLGGGS	Phage display screening	99 ± 5	86 ± 4
147	TLPAILQAAGTRGGGS	Phage display screening; S8A and S9A mutations	92 ± 5	96 ± 6
155	TLPKILQSSGTRGGGS	Phage display screening; A4K mutation	81 ± 4	80 ± 5
166	TLPAILRSSGTRGGGS	Phage display screening; Q7R mutation	76 ± 5	111 ± 8
167	TLPAILKSSGTRGGGS	Phage display screening; Q7K mutation	78 ± 3	102 ± 3
213*	AISSPLLRNPFRGGGS	Phage display screening; W8R mutation	113 ± 4	89 ± 5
214*	AISSPLLKNPFRGGGS	Phage display screening; W8K mutation	114 ± 1	106 ± 4
1 × 2				
Peptide index	Peptide sequence	Peptide origin	Average 1 × 2 affinity intensity	% Selectivity
36	GSGTDFTLKISRVEA	mAb735	43 ± 7	109 ± 20
104	HLSLKNPLRMDLGGGS	Phage display screening	69 ± 3	119 ± 5
213*	AISSPLLRNPFRGGGS	Phage display screening; W8R mutation	118 ± 1	100 ± 3
214*	AISSPLLKNPFRGGGS	Phage display screening; W8K mutation	124 ± 2	93.6 ± 4.0



ESI Fig. 4 Selectivity of peptides derived from peptide 106 (P106; phage display screening peptide) with (A) 2× and (B) 1 × 2 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 in the above plots (library index 170) displays high affinity as in the 1× assay (Fig. 4), likely due to the replacement of neutral residues with three arginines (among other substitutions). However, this peptide shows poorer selectivity in the presence of 10% competitor than peptides of similar affinity, suggesting that designing selective peptide ligands to target a negatively charged polysaccharide requires considerations beyond a general increase in peptide basicity (such as positioning of charged residues, peptide conformation, etc.).



ESI Fig. 5 PSA-binding of mAb735-derived peptides. Average binding intensities from affinity screening in 1×, 2×, and 1 × 2 assays (with triplicate intra-assay readings) are displayed in relative intensity units. Though differences in assay conditions (specifically peptide density) may influence variation seen in peptide affinity to PSA, inter-assay intensities are plotted to display general trends in binding. Error bars represent standard deviations of inter-assay values. Similar binding of contiguous sequences is in keeping with derivation of sequences from epitope mapping, with adjacent sequences containing 13 residue overlaps. Higher binding regions do not necessarily correspond to CDRs of mAb735 (CDRs from which sequences originate shown above corresponding points), as mAb735 has a discontinuous epitope in the binding of oligosialic acid.[2] However, linear epitope mapping of the discontinuous epitope may have provided peptide fragments encompassing residue composition and positioning with a higher likelihood of target interaction,[3] similar to a rational partners-based design approach in the engineering of peptide ligands to protein targets.[4–6] Additionally, future work with linear epitope mapping of sialic acid-binding proteins with continuous binding pockets may help elucidate lectin–PSA interactions (an advantage not afforded by random screening).

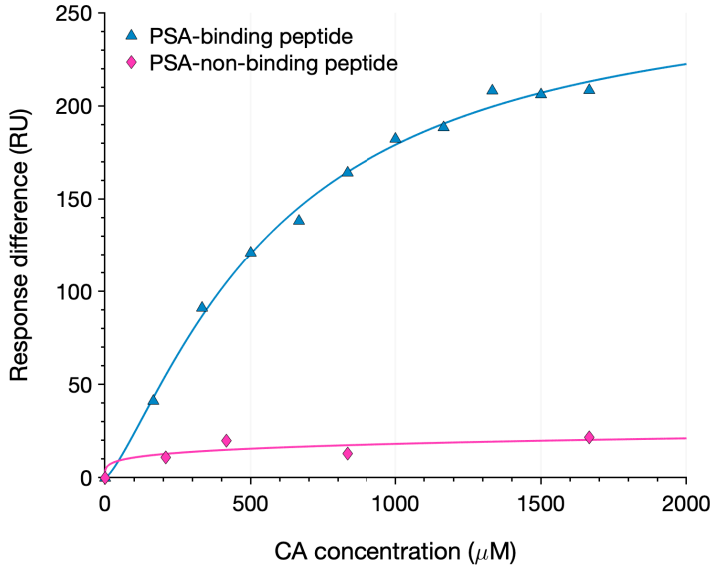


ESI Fig. 6 TBE PAGE for molecular weight assessment of glycans. 4–20% Novex™ TBE gel (Invitrogen, Carlsbad, CA) with lanes 1–2 CA (sodium salt, Nacalai USA, Inc.) and lanes 3–4 CS-A (from bovine trachea, Sigma-Aldrich); 210 μg loaded in lanes 1 and 3 and 52.5 μg loaded in lanes 2 and 4. Gel stained with alcian blue (image acquisition in grayscale). Given polydispersity of samples, rough molecular weight of CS-A estimated as 45 kD based on the following: i) manufacturer provided average molecular weight of CA = 30 kD and observed molecular weight range of CA, ii) expected migration of oligonucleotide fragments and proteins of different molecular weights (Invitrogen gel migration chart for Novex™ pre-cast gels),[7,8] and iii) estimates of bovine trachea CS-A molecular weight from literature.[9,10] Each of the previous likely provides a very broad estimation of molecular weight, and significantly more accurate determination may be performed chromatographically. However, for the purposes of this work, a general estimate suffices for selectivity studies, where selectivity conditions chosen (0, 0.1, 10, and 100% competitor with CA) represent a broad range of low to high (with logarithmic intervals), and relative amounts of competitor between conditions are accurate. Additionally, conclusions on selective binding of peptides are relative (between assay or selectivity conditions or between peptides) and are not based on constants (such as K_D) dependent on glycan concentrations. The term "equimolar," when used to describe CS-A and CA amounts, is thus meant to indicate a general equivalency in amounts of glycans that encompasses error from molecular weight estimation, including from sample polydispersity.

ESI Table 3 Microarray affinity and selectivity of control peptides. Since no peptide ligands have previously been discovered or designed for PSA, positive and negative controls incorporated with microarray screening were based on the minimal information available on known PSA protein ligands, specifically the ligand Siglec-11,[11] for which the putative binding residues were incorporated within the designed peptide ligand 170 (see Table 1 and ESI Table 1). The PSA-binding peptide (sequence EAWYFFKVERGSRVR) is similar to residues from the putative binding region with an R7K mutation (modified to increase basicity but not significantly alter sequence characteristics). In contrast, the non-binding peptide (sequence EHGGGLGLGAALGAG) was obtained from a non-binding region of the same protein (occurring within a hydrophobic, transmembrane domain). Binding and non-binding of these two peptide sequences corresponds to observed peptide compositional data for the peptide library (e.g., basicity and arginine prevalence of the positive control compared to hydrophobicity and N-terminal negative charge presentation of the negative control).

The two peptides were screened alongside mAb and phage display-derived peptides in identical assay conditions. Microarray intensity values provided are for screening in 10 μM CA with specified concentrations of competitor (CS-A). Errors represent standard deviations of triplicate intra-assay measurements.

Peptide	Printing condition	0 μM CS-A	0.1 μM CS-A	1 μM CS-A	1 μM CS-A
PSA-binding	1×	67 ± 1	43 ± 12	12 ± 2	15 ± 2
	2×	65 ± 1	49 ± 2	19 ± 1	21 ± 2
	1 × 2	92 ± 5	65 ± 5	17 ± 2	17 ± 3
Peptide	Printing condition	0 μM CS-A	0.1 μM CS-A	1 μM CS-A	1 μM CS-A
PSA-non-binding	1×	15 ± 3	20 ± 5	13 ± 4	17 ± 4
	2×	22 ± 3	17 ± 2	21 ± 1	20 ± 1
	1 × 2	19 ± 2	19 ± 3	18 ± 3	17 ± 1



ESI Fig. 7 SPR spectroscopy of positive and negative control peptides (see ESI Table 3) with the Biacore 3000 system (GE Healthcare, Chicago, IL). Peptides were synthesized at >95% purity (confirmed with HPLC) by Biomatik with N-terminal acetylation and C-terminal cysteine addition. Peptides were immobilized by thiol coupling of cysteines to CM5 sensor ships in accordance with standard protocols (GE Healthcare). CA solutions in PBS, pH 7.4 from 0–1.67 mM (0–50 mg/mL) were injected for 180 seconds at a flow rate of 10 µL/min, and the surface was regenerated with 2 M NaCl injection for 90 seconds at the same flow rate. Sensorgrams were obtained following subtraction of reference flow cell response. Equilibrium binding responses at each analyte concentration were determined as the relative response difference from baseline at which binding reached saturation, and equilibrium binding curves were subsequently generated in MATLAB using the four-parameter logistic binding model.[12] A K_D value of 575 ± 74 µM was thus obtained (error represents standard error of the fit with 95% confidence bounds).

Given the unknown absolute ligand density (as for microarrays) and the effect of ligand immobilization on binding constants, though a stoichiometry of 1.32 ± 0.15 was obtained from the fit, it is possible that the affinity constant is an overestimation of 1:1 in-solution binding strength. Additionally, the high viscosity of CA contributed to artifacts within sensorgrams (such as spiking) due to high refractive index changes with analyte injection; extraction of binding values was possible with artifact removal, but such effects may have contributed to error. (Artifacts were minimized with immobilization of CA and injection of peptide; however, responses fell below the detection limit due to the relatively small size of peptides compared to CA.) Nevertheless, SPR responses confirm the ability of the peptide microarray assay to discriminate between binding and non-binding peptides to the polysaccharide target and furthermore, demonstrate the high-to-low throughput approach that may be pursued for future analysis of PSA-binding peptides.

References for supplementary information

1 N. Horan, L. Yan, H. Isobe, G. M. Whitesides and D. Kahne, *Proc. Natl. Acad. Sci. U. S. A.*, 1999, **96**, 11782.

2 M. Nagae, A. Ikeda, M. Hane, S. Hanashima, K. Kitajima, C. Sato and Y. Yamaguchi, *J. Biol. Chem.*, 2013, **288**, 33784.

3 D. J. Brooks, J. R. Fresco, A. M. Lesk and M. Singh, *Mol. Biol. Evol.*, 2002, **19**, 1645.

4 D. Chandra, S. Timmick, C. Goodwine, N. Vecchiarello, D. G. Shastry, A. Mullerpatan, J. P. Trasatti, S. Cramer and P. Karande, *J. Chem. Technol. Biotechnol.*, 2019, **94**, 2345.

5 S. Sachdeva, H. Joo, J. Tsai, B. Jasti and X. Li, *Sci. Rep.*, 2019, **5**, DOI: 10.1038/s41598-018-37201-6.

6 D. J. Huggins, W. Sherman and B. Tidor, *J. Med. Chem.*, 2012, **55**, 1424.

7 Invitrogen, Life Technologies Corporation, TBE Gels, Pub. Part No. IM-6003, 2011.

8 Invitrogen, Life Technologies Corporation, Novex Pre-Cast Gel Electrophoresis Guide, Manual Part No. IM-1002, 2010.

9 N. Igarashi, A. Takeguchi, S. Sakai, H. Akiyama, K. Higashi and T. Toida, *Int. J. Carbohydr. Chem.*, 2013, **2013**, DOI: 10.1155/2013/856142.

10 M. B. Mathews, *Arch. Biochem. Biophys.*, 1956, **61**, 367.

11 T. Angata, S. C. Kerr, D. R. Greaves, N. M. Varki, P. R. Crocker and A. Varki, *J. Biol. Chem.*, 2002, **277**, 24466.

12 J. W. Findlay and R. F. Dillard, *AAPS J.*, 2007, **9**, E260, DOI: 10.1208/aapsj0902029.