Electronic Supplementary Material (ESI) for Molecular BioSystems

Supporting Information

Orientation of GST-tagged lectins *via in situ* surface modification to create an expanded lectin microarray for glycomic analysis

Daniel C. Propheter and Lara K. Mahal*

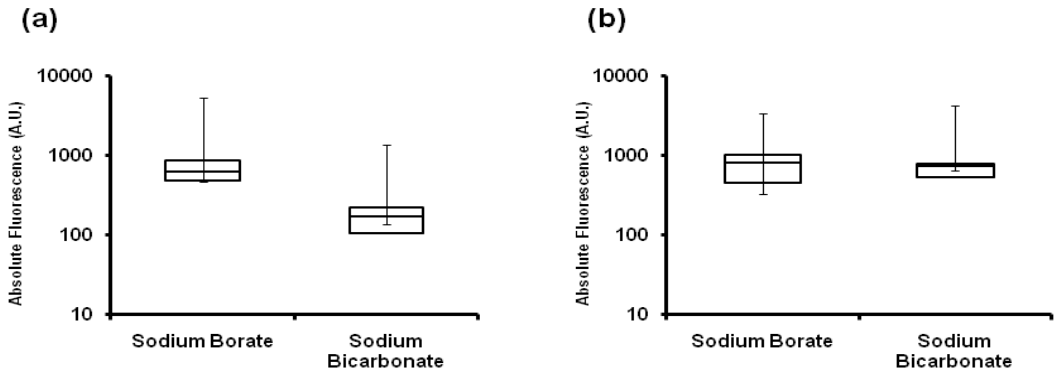

Figure S1: The distribution of binding of GST-GafD and GST-RS-IIL versus two glycoproteins with respect to two buffer systems. (a) A box and whisker plot of the observed binding of GST-GafD (20 µM) and GST-RS-IIL (24 µM) that were printed in either 50 mM sodium borate buffer (pH 8.5) or 100 mM sodium bicarbonate (pH 9.3) supplemented with various concentrations of GSH (100, 50, 25, 10, and 0 mM). Samples were assayed against various concentrations of Cy5-labeled chicken egg ovalbumin (OVA-Cy5, 10, 5, 2.5, 1.25, 0.625, and 0.313 µM). (b) A box and whisker plot of the binding of GST-RS-IIL (24 µM) that was printed in either 50 mM sodium borate buffer (pH 8.5) or 100 mM sodium bicarbonate (pH 9.3) supplemented with various concentrations of GSH (100, 50, 25, 10, and 0 mM). Printed samples were assayed against various concentrations of Cy3-labeled RNase B (RNase B-Cy3, 10, 5, 2.5, 1.25, 0.625, and 0.313 µM). Each box plot represents the distribution of binding (in absolute fluorescence). The interquartile range (boxed) represents the distribution of 50% of the total binding signals and was used to compare between the different printing conditions in both orienting the GST-fusion lectin and maintaining glycan-binding activity.

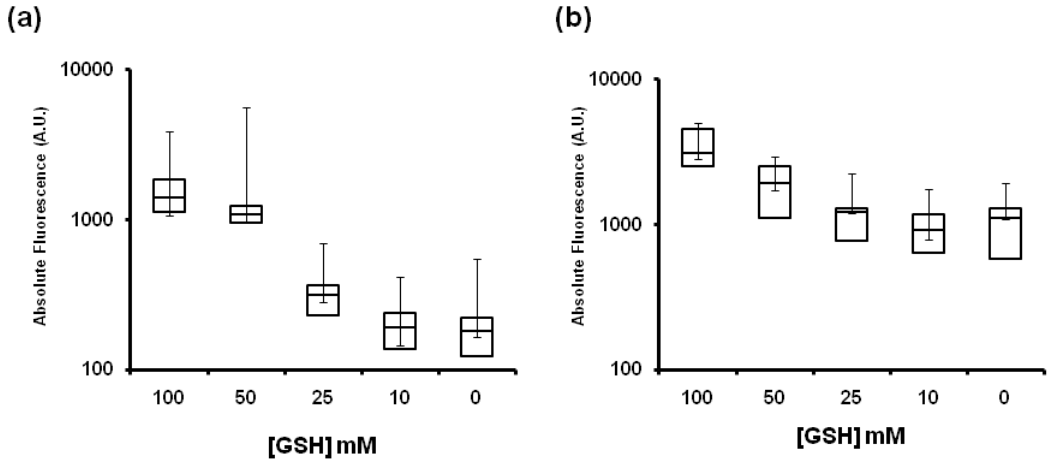

Figure S2: The distribution of binding of GST-GafD and GST-RS-IIL versus two glycoproteins with respect to various GSH concentrations. (a) A box and whisker plot of the observed binding of GST-GafD (20 µM) and GST-RS-IIL (24 µM) that were printed in both 50 mM sodium borate buffer (pH 8.5) or 100 mM sodium bicarbonate (pH 9.3) supplemented with various concentrations of GSH (100, 50, 25, 10, and 0 mM). Samples were assayed against various concentrations of Cy5-labeled chicken egg ovalbumin (OVA-Cy5, 10, 5, 2.5, 1.25, 0.625, and 0.313 µM). (b) A box and whisker plot of the binding of GST-RS-IIL (24 µM) that was printed in both 50 mM sodium borate buffer (pH 8.5) or 100 mM sodium bicarbonate (pH 9.3) supplemented with various concentrations of GSH (100, 50, 25, 10, and 0 mM). Printed samples were assayed against various concentrations of Cy3-labeled RNase B (RNase B-Cy3, 10, 5, 2.5, 1.25, 0.625, and 0.313 µM). Each box plot represents the distribution of binding (in absolute fluorescence). The interquartile range (boxed) represents the distribution of 50% of the total binding signals and was used to compare between the different printing conditions.

Discussion of Figures S1 and S2: In order to optimize conditions for lectin orientation, we focused on two variables: the buffer pH and the concentration of GSH. In our previous work[1], we used three different buffers to test GSH coupling to the NHS-activated surface: 10 mM phosphate buffered saline (PBS, pH 7.4), 50 mM sodium borate (pH 8.5), and 100 mM sodium bicarbonate (pH 9.3). We showed that pH of the GSH-coupling buffer played a key role in the coupling chemistry of the primary amine of the aspartate residue (amino terminus) of the GSH peptide and the polymeric, NHS-activated ester surface on which we printed. Compared to the higher pH buffers, PBS showed very little activity. Sodium borate showed intermediate activity and sodium bicarbonate displayed the best activity. Given that pH is a large factor in the competition between GSH and the lysines of proteins, we decided to test both the sodium borate and sodium bicarbonate buffers. We printed both GST-GafD (20 µM) and GST-RS-IIL (24 µM) in the sodium borate and sodium bicarbonate buffers. In both buffers, a range of concentrations of GSH was tested (100, 50, 25, 10, 0 mM). In Figure S1a, we plot the observed activity of GST-GafD and GST-RS-IIL against various concentrations of OVA-Cy5 with respect to each printing buffer containing all GSH concentrations. In Figure S1b, we probed desposited GST-RS-IIL (24 µM) under the same conditions in (a). In both conditions, the sodium borate buffer was found to be the better buffer and statistically different from the sodium bicarbonate values ($p < 0.001$). We initially believed that 50 mM GSH would be the optimal concentration based on our previous work. When we analyzed same set of data to determine the optimal GSH concentration as described above, (Figure S2a) the activity of GST-GafD and

GST-RS-IIL against OVA-Cy5 at 100 and 50 mM GSH concentrations are very similar ($p = 0.7$), however, as shown in Figure S2b, the differences between the activity of GST-RS-IIL against RNase B differ a significant amount between 100 and 50 mM GSH ($p < 0.001$). However, when the data for OVA-Cy5 binding for GST-GafD and GST-RS-IIL were analyzed separately, the difference between the 100 and 50 mM GSH concentrations for GST-RS-IIL was statistically significant ($p = 0.025$), although for GST-GafD they were not ($p = 0.96$). Thus we concluded that the optimal buffer for GST-lectin activity was 50 mM sodium borate (pH 8.5) with 100 mM GSH.

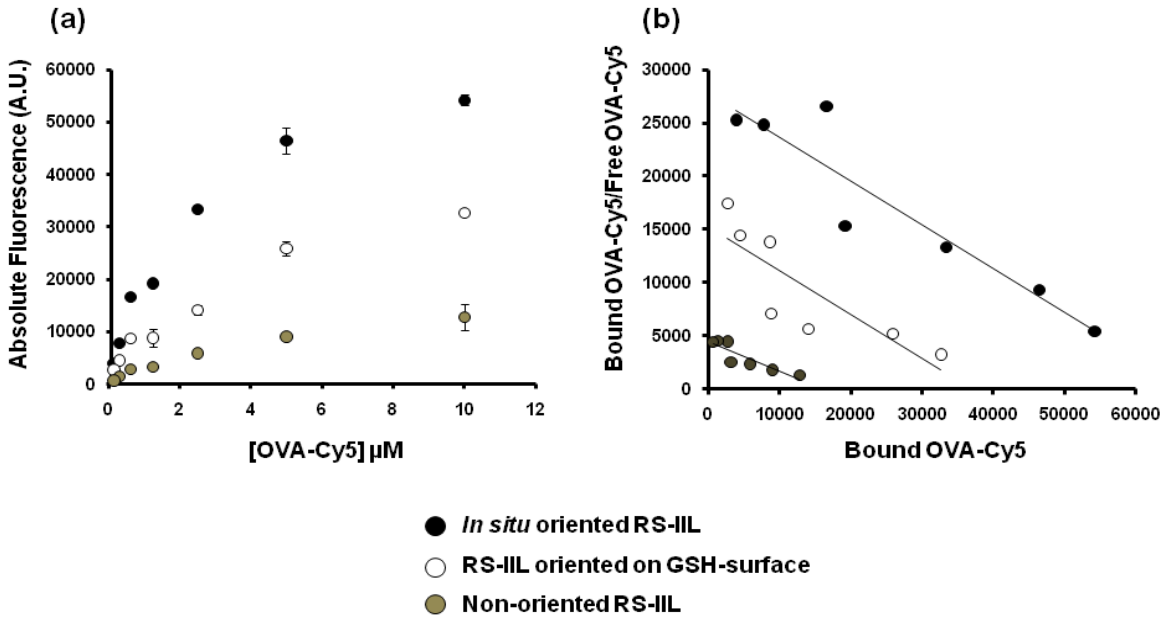

Figure S3: Analysis of GST-RS-IIL (24 µM) binding to varying concentrations of OVA-Cy5 (10, 5, 2.5, 1.25, 0.63, 0.31, and 0.16 µM). GST-RS-IIL was printed in GSH-B on NHS-activated slides (*In situ* oriented RS-IIL, closed circles), in PB on NHS-activated slides (Non-oriented RS-IIL, tan circles), and in PB printed on a pre-modified GSH layer (RS-IIL oriented on GSH-surface, open circles). (a) Activity of GST-RS-IIL printed in GSH-B, PB, or PB on a GSH-modified surface as previously described[1]. (b) Scatchard plot of data represented in (a). Plots were analyzed using linear regression, and apparent dissociation constants (K_d) were determined from the slope of each of the lines (*In situ* oriented $K_d = 2.05 \pm 0.33$ µM, non-oriented $K_d = 3.59 \pm 0.39$ µM, GST-RS-IIL oriented on GSH-surface, $K_d = 2.63 \pm 0.47$ µM). Graphical data is representative of triplicate arrays. Error bars indicate the standard deviation from the mean value.

Discussion of Figure S3: To directly compare this new method of protein orientation with our previously published method, we printed GST-RS-IIL using both conditions and probed for activity against OVA-Cy5. As anticipated, the apparent dissociation constants for RS-IIL were in the micromolar range under all conditions.[1] However, a significant difference in overall activity was observed as seen in the level of binding to the fluorescent protein (Figure S3a). In accordance with our previous data, GST-RS-IIL printed on a GSH-surface displayed ~3-fold increase in activity compared to the non-oriented lectin. Comparison between the two orientation methods revealed that *in situ* orientation resulted in higher levels of activity than deposition on the pre-made surface (~2-fold increase). While the exact reason for the increase in activity is unknown, it may arise from differences in the fabrication process. Based on our optimized procedures for each condition, the *in situ* print buffer contains twice as much GSH as the derivitization

buffer. This may result in overall lower levels of modification. More experiments would be needed to clarify this point.

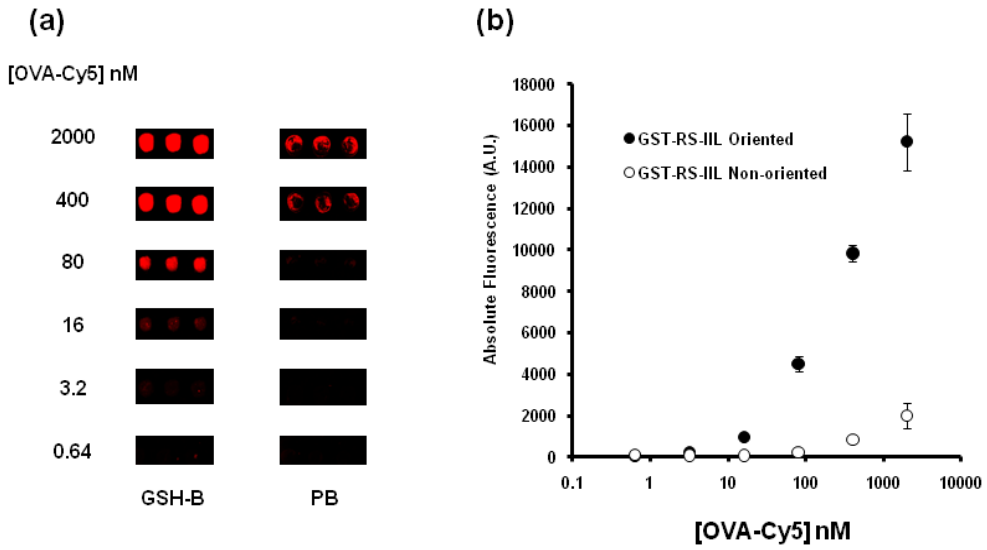

Figure S4: (a) GST-RS-IIL (24 µM) was printed in either GSH-B (*in situ* oriented) or PB (non-oriented) and hybridized with Cy5-labelled ovalbumin (OVA-Cy5). (b) Graphical representation indicates that the largest difference in activity is ~8-fold in favor of the oriented GST-RS-IIL. The enhanced spot morphology of the oriented lectin is also clearly visible. Error bars indicate the standard deviation from the mean value. Data is representative of triplicate arrays.

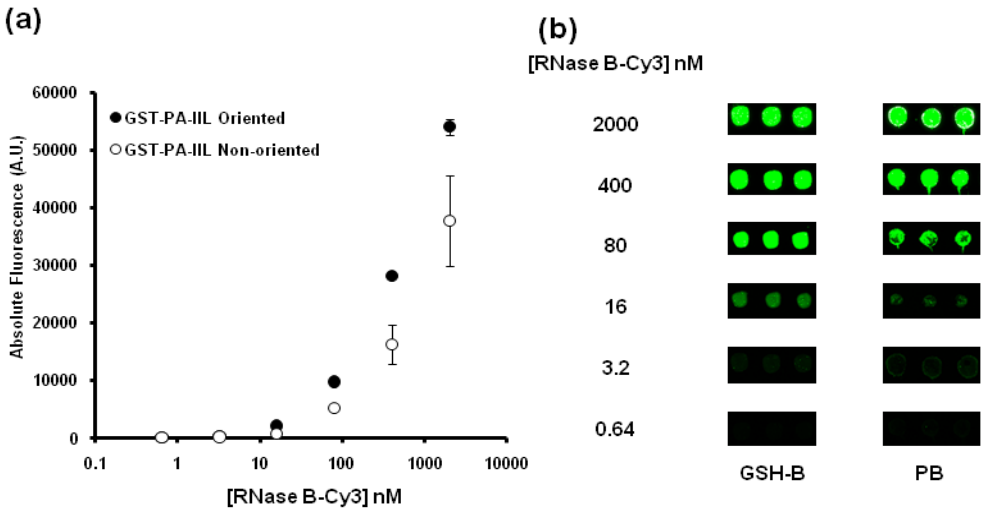

Figure S5: GST-PA-IIL (24 µM) printed in either GSH-B (*in situ* oriented) or PB (non-oriented) and hybridized with varying concentrations of Cy3-labeled RNase B (RNase B-Cy3). (a) Graphical representation of the data in (b) indicates that binding of RNase B by GST-PA-IIL shows modest improvement in activity (~1.5 – 2x) when oriented. Error bars indicate the standard deviation from the mean value. (b) Array data showing GST-PA-IIL binding to RNase B. Data is representative of triplicate arrays.

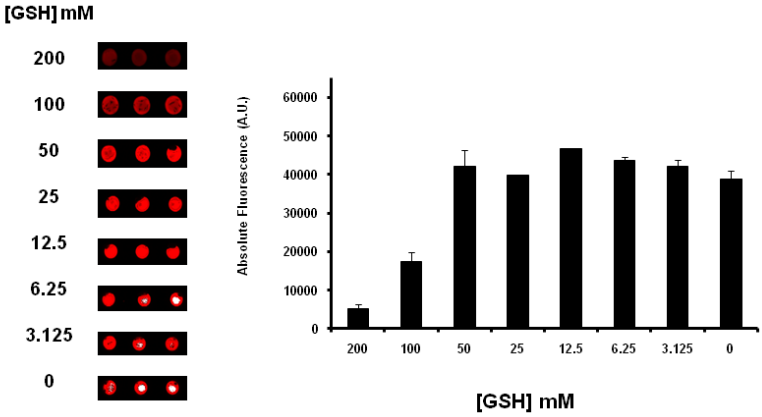

Figure S6: Printing of Alexafluor 647-conjugated BSA (BSA-AF, 3.8 µM) in 50 mM sodium borate buffer (pH 8.5) with varying amounts of GSH (200, 100, 50, 25, 12.5, 6.25 , 3.125, and 0 mM). At high concentrations of GSH, the deposition of BSA is out-competed by the small molecule, which at such concentrations, effectively reacts faster with the NHS-activated esters than the side-chains lysines from the labeled BSA.

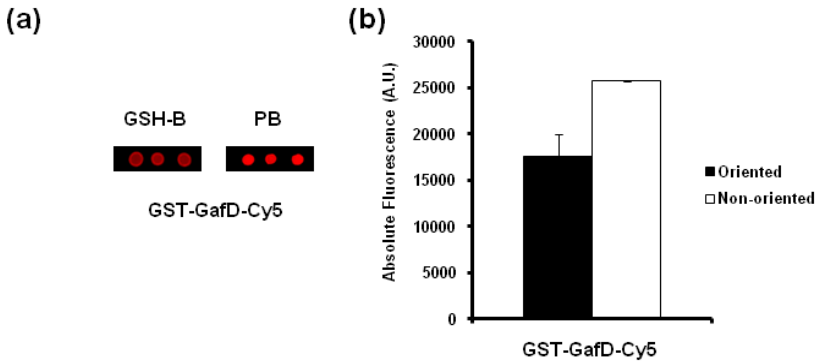

Figure S7: Comparison of Cy5 labeled GST-GafD (GST-GafD-Cy5[2]) printed in orienting buffer, GSH-B, and non-orienting buffer, PB. (a) GST-GafD-Cy5 (10 µM) was printed GSH-B or PB. Array data is shown. (b) Graphical representation of data depicted in (a). Error bars indicate the standard deviation from the mean value. Data is representative of duplicate arrays.

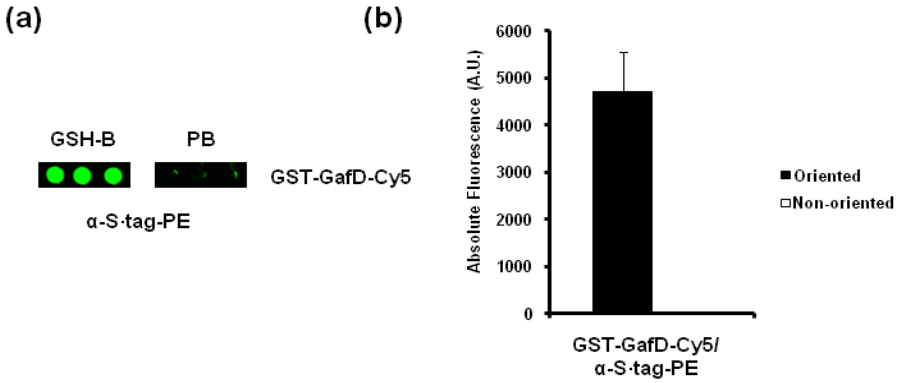

Figure S8: Comparison of the indirect detection of GST-GafD-Cy5[1] printed in the orienting buffer, GSH-B, and non-orienting buffer, PB. (a) Array shown in Figure S8 was probed with 1 µg/mL of PE-conjugated α-S·tag antibody (α-S·tag-PE). (b) Graphical representation of data depicted in (a). Detection of the S·tag domain of the oriented GST-GafD-Cy5 is far greater than that compared to the unoriented, randomly deposited lectin. This is not an issue of protein deposition as the deposition of these spots is shown in Figure S8. As shown in previous work, the S·tag epitope is more accessible due to orientation, hence GST-GafD-Cy5 in GSH-B is more easily detected. Error bars indicate the standard deviation from the mean value. Data is representative of duplicate arrays.

Discussion of Figure S7 and S8: GST-GafD-Cy5 was printed in GSH-B or PB to determine protein deposition and orientation (see paper for discussion). As previously described, labeling of GST-GafD with Cy5-NHS destroys the activity of the lectin, but not GST. This is due to the labeling of the protein while bound to GSH-beads. This precludes direct comparison of protein concentration to activity.

Calculation of relative of percent activity of non-oriented GST-GafD: To estimate the relative percentage of GST-GafD that is active when it is not oriented, we made the following assumptions: 1. All oriented GST-GafD is active; 2. The relative activities of oriented versus non-oriented protein can be estimated from the differences in fluorescence from OVA-Cy5 binding, and; 3. The concentration of unlabeled oriented GST-GafD is also 30% lower than non-oriented GST-GafD. We calibrated the activity to protein amount to compensate for the difference.

Percent activity of GST-GafD non-oriented , X = (relative protein deposition of oriented/ relative activity of oriented protein)*100

X = (0.7/9)*100 = ~8%

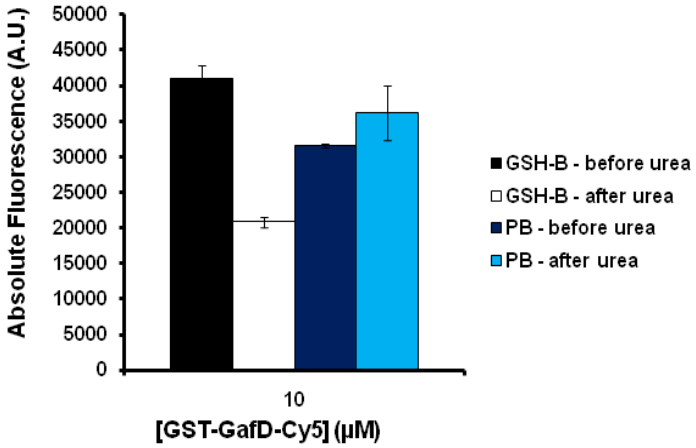

Figure S9: The affect of treatment of *in situ* oriented GST-GafD-Cy5 with 6 M urea. GST-GafD-Cy5 (10 µM) was printed in GSH-B and scanned in the Cy5 channel (before urea). The array was then treated with 6 M urea in PBS, and the same slide was scanned in the Cy5 channel again (after urea). Error bars indicate the standard deviation from the mean value. Data is representative of duplicate arrays.

Discussion of Figure S9: In order to determine if the *in situ* GSH-modified surface forms a non-covalent interaction with our GST-tagged protein, we printed GST-GafD-Cy5 in GSH-B. After 1 hr incubation the slide was scanned (Figure S8, before urea), then treated with 6 M urea to denature the proteins, which we expect to disrupt the protein folding required for GST-GSH interaction. The array was then washed, and re-scanned under the same scanning conditions as before (Figure S8, after urea). As a result of urea treatment, we observed ~50% decrease in fluorescence from our *in situ* oriented protein but no decline in signal from our non-oriented protein. This observation indicates that at least 50% of our protein is non-covalently bound to our GSH surface. However, half of all protein was still present after the denaturing conditions. There are several possible reasons for this result. First, it is plausible that the GSH binding domain of GST is buried within the Nexterion H surface, which may protect it from denaturation by urea. When in a dry state, the NHS-PEG surface of the slide is ~10 nm thick. When hydrated, the surface expands to between 50 – 100 nm, thus allowing adsorption of the sample into the slide.[3] Considering that proteins, in general, are well within a ~100 nm diameter, the GSH binding domain could be buried within the Nexterion slide, protecting the protein and also making the binding interaction difficult to compete out.

GST-lectin[4]	Binding Specificity
GST-GafD	β-GlcNAc
GST-PA-IL	Galactose
GST-PA-IIL	Fucose/Mannose
GST-PapGII	GbO4
GST-PapGIII	GbO5
GST-RS-IIL	Mannose/Fucose

Table S1: List of GST-tagged recombinant lectins used in this study.

Number	Row	Lectin	[Lectin] μM	[Lectin] μg/mL	Monosaccharide
1	1	AAA	14	1000	Fucose
2	1	AIA, Jacalin	10	500	Galactose
3	1	BPA	3	500	Galactose
4	1	ConA	10	500	Mannose
5	1	DBA	4	500	Galactose
6	2	DSA	6	500	Lactose
7	2	ECA	19	500	Galactose
8	2	GNA	10	500	Mannose
9	2	GS-I	4	500	Galactose
10	2	GS-II	4	500	GlcNAc
11	3	HPA	6	500	Galactose
12	3	LcH	20	1000	Mannose
13	3	Lotus	9	500	Fucose
14	3	MAA	4	500	Lactose
15	3	PAA	N/A	500	GlcNAc
16	4	PNA	5	500	Galactose
17	4	PHA-E	4	500	Lactose
18	4	PHA-L	5	500	Galactose
19	4	SBA	4	500	Galactose
20	4	SNA	3	500	Lactose
21	5	STA	5	500	GlcNAc
22	5	TJA-I	8	500	Lactose
23	5	TJA-II	8	500	Lactose
24	5	UEA-I	8	500	Fucose
25	5	WGA	28	1000	GlcNAc
26	**6**	**GST-GafD**	**20**	**1000**	**GlcNAc**
27	**6**	**GST-GafD**	**10**	**500**	**GlcNAc**
28	**6**	**GST-GafD**	**5**	**250**	**GlcNAc**
29	**6**	**GST-GafD**	**2.5**	**125**	**GlcNAc**
30	**6**	**GST-GafD**	**1.25**	**62.5**	**GlcNAc**
31	**7**	**GST-PA-IL**	**23**	**1000**	**Galactose**
32	**7**	**GST-PA-IL**	**11**	**500**	**Galactose**

33	**7**	**GST-PA-IL**	**6**	**250**	**Galactose**
34	**7**	**GST-PA-IL**	**3**	**125**	**Galactose**
35	**7**	**GST-PA-IL**	**1.4**	**62.5**	**Galactose**
36	**8**	**GST-PA-IIL**	**23**	**1000**	**Fucose**
37	**8**	**GST-PA-IIL**	**11**	**500**	**Fucose**
38	**8**	**GST-PA-IIL**	**6**	**250**	**Fucose**
39	**8**	**GST-PA-IIL**	**3**	**125**	**Fucose**
40	**8**	**GST-PA-IIL**	**1.4**	**62.5**	**Fucose**
41	**9**	**GST-PapGII**	**19**	**1000**	**Galactose**
42	**9**	**GST-PapGII**	**9**	**500**	**Galactose**
43	**9**	**GST-PapGII**	**4.5**	**250**	**Galactose**
44	**9**	**GST-PapGII**	**2.3**	**125**	**Galactose**
45	**9**	**GST-PapGII**	**1.2**	**62.5**	**Galactose**
46	**10**	**GST-PapGIII**	**19**	**1000**	**Galactose**
47	**10**	**GST-PapGIII**	**9**	**500**	**Galactose**
48	**10**	**GST-PapGIII**	**5**	**250**	**Galactose**
49	**10**	**GST-PapGIII**	**2.5**	**125**	**Galactose**
50	**10**	**GST-PapGIII**	**1.2**	**62.5**	**Galactose**
51	**11**	**GST-RS-IIL**	**23**	**1000**	**Mannose**
52	**11**	**GST-RS-IIL**	**11**	**500**	**Mannose**
53	**11**	**GST-RS-IIL**	**6**	**250**	**Mannose**
54	**11**	**GST-RS-IIL**	**3**	**125**	**Mannose**
55	**11**	**GST-RS-IIL**	**1.4**	**62.5**	**Mannose**

Table S2: Print list for lectin array shown in Figure 4. Three spots were printed for each lectin/concentration and 15 spots/5 lectins were printed per row. Bolded samples were printed in GSH-B (100 mM GSH, 50 mM sodium borate buffer, pH 8.5 with 0.5 mg/mL BSA). All others printed in PB (10 mM sodium phosphate, 15 mM sodium chloride, pH = 7.4 with 0.5 mg/mL BSA)

Supplemental Methods and Materials:

Microarray Fabrication:

Recombinant lectins were expressed and purified as described previously.[3] GST-GafD-Cy5 was expressed, purified, and labeled as previously described.[1] Unless otherwise noted, all microarrays were printed *via* the following protocol: Lectins were diluted in either GSH-B (50 mM sodium borate buffer, 100 mM GSH, pH 8.5) or PB (phosphate buffered saline (PBS), 10 mM sodium phosphate, 15 mM sodium chloride). For lectin printing, both print buffers contain 1 mM $CaCl_2$, 1 mM $MgCl_2$, and 1 mM of appropriate monosaccharide (see Table S2), and 0.5 mg/mL BSA. Prepared samples were loaded into a 384-well microplate (Whatman, Piscataway, NJ), and loaded into the SpotBot2 Personal Microarrayer (ArrayIt, Sunnyvale, CA). Printing programs were created with the MMF Spocle Program. Samples were

printed onto Nexterion H slides (Schott North America, Elmsford, NY) with an SMP3 pin (ArrayIt, #SMP3). During the print, the slides were kept at 8 °C with internal humidity maintained at ~50% throughout the print process. After printing, the slides were allowed to warm to room temperature for 2 hrs, while maintaining humidity at ~50%. After 2 hrs, the slides were then placed in a coplin jar and blocked with 50 mM ethanolamine in 50 mM sodium borate buffer (pH 8.5) for 1 hr, at room temperature with mild shaking. After one hour, the slides were washed with PBS with 0.05% Tween (0.05% PBS-T, 3 x 3 min) and once with PBS. The slides were dried using a slide spinner (Labnet Intl., Edison, NJ), and then fastened in a 24-well hybridization chamber (ArrayIt). Fluorescently labeled samples were diluted into 0.005% PBS-T, and 100 µL were added to each subarray and samples were incubated for 2 hrs at room temperature with gentle shaking. For visualization with α-S·tag-PE antibody, the buffer was changed to 0.005% PBS-T with 1% BSA. After 2 hrs, samples were aspirated and washed with 0.005% PBS-T (0.005% Tween 20 in PBS, 3 x 3 min) and once with PBS. The slides were dried as before and loaded into the Genepix 4100A slide scanner (Molecular Devices, Union City, CA). Data was extracted with GenePixPro 5.0 (Molecular Devices) and analyzed and graphed using Microsoft Office Excel 2007.

Labeling, printing, and hybridization of BSA-Alexafluor:

Alexa Fluor® 647 C_2-maleimide (667 µM, dissolved in PBS, Invitrogen #A20347,) was added to BSA (10 mg/mL, ~150 µM) in PBS and incubated at room temperature for 1 hr with gentle shaking. After 1 hr, the sample was dialyzed against PBS for 12 hrs, and the final concentration was determined by DC Assay (Bio Rad #500-0112). BSA-AF (3.8 µM) was dissolved in 50 mM sodium borate buffer (pH 8.5) with varying amounts of GSH (200, 100, 50, 10, 5, 1, and 0 mM) and printed as described above. After blocking, washing, and drying the slide, 100 µL of 250 µM Maleimide-PEG_2-Biotin (Thermo Fisher Scientific, #21901), in 0.005% PBS-T was added to the array and incubated at room temperature for 1 hr. After 1 hr, the slide was washed with 0.005% PBS-T (3 x 3min) and once with PBS, and dried. 100 µL of 50 µg/mL of Cy3-labeled streptavidin (Invitrogen, #43-4315) was then added to each well and incubated for 1 hr at room temperature. After 1 hr, slides were then washed, dried, and scanned as previously described.

Urea treatment immobilized lectins

GST-GafD was printed in 50 mM sodium borate (pH 8.5) containing 100 mM GSH (GSH-B) on Nexterion H slides as described above. After printing, the slides were allowed to incubate for two hours, equilibrating to room temperature. The slides were then blocked in 50 mM ethanolamine in 50 mM sodium borate buffer (pH 8.5) for 1 hr, then washed with 0.05% PBS-T (3 x 3 min) and once with PBS. The slides were dried and fastened to a 24-well hybridization chamber (ArrayIt). Arrays were then either treated to 100 µL of 6 M urea in PBS or 0.005% PBS-T, and were incubated for 1 hr at room temp. After 1 hr, the slides were washed with 200 µL of 0.005% PBS-T (3 x 3 min), then were treated with either 10 µM Cy5-labeled ovalbumin (OVA-Cy5) or α-S·tag antibody PE-conjugated in 0.005% PBS-T or 0.005% PBS-T with 1% BSA, respectively. The slides were then washed with 0.005% PBS-T (3 x 3 min) and once with PBS. The slides were dried, scanned, and analyzed as described above.

References:

(1) J. Adam, M. Pokorná, C. Sabin, E. P. Mitchell, A. Imberty, M. Wimmerová, *BMC Struct. Biol.*, **2007**, *7*, 36.
(2) D.C. Propheter, K.L. Hsu, L.K. Mahal, *Chembiochem*, **2010**, *11*, 1203-1207.
(3) G.M. Harbers, K. Emoto, C. Greef, S.W. Metzger, H.N. Woodward, J.J. Mascali, D.W. Grainger, M.J. Lochhead, *Chem. Mater.*,**2007**, *19*, 4405-4414.
(4) K.L. Hsu, J.C. Gildersleeve, L.K. Mahal, *Mol. Biosys*, **2008**, *4*, 654-662.