



HHS Public Access

Author manuscript

Methods Mol Biol. Author manuscript; available in PMC 2015 July 22.

Published in final edited form as:

Methods Mol Biol. 2015 ; 1273: 203–214. doi:10.1007/978-1-4939-2343-4_15.

Exploring the specificities of glycan-binding proteins using glycan array data and the GlycoSearch software

Doron Kletter¹, Bryan Curnutte², Kevin Maupin², Marshall Bern¹, and Brian B. Haab^{2,*}

¹Palo Alto Research Center (PARC)

²Van Andel Research Institute

Summary

The glycan array is a powerful tool for investigating the specificities of glycan-binding proteins. By incubating a glycan-binding protein on a glycan array, the relative binding to hundreds of different oligosaccharides can be quantified in parallel. Based on these data, much information can be obtained about the preference of a glycan-binding protein for specific subcomponents of oligosaccharides, or motifs. In many cases the analysis and interpretation of glycan array data can be time consuming and imprecise if done manually. Recently we developed software, called GlycoSearch, to facilitate the analysis and interpretation of glycan array data based on the previously developed methods called Motif Segregation and Outlier Motif Analysis. Here we describe the principles behind the software, the use of the software, and an example application. The automated, objective, and precise analysis of glycan array data should enhance the value of the data for a broad range of research applications.

1. Introduction

A valuable tool for characterizing the nature of interactions between glycans and proteins is the glycan microarray (1–7). Glycan arrays are composed of diverse oligosaccharides, either produced synthetically or purified from natural sources, immobilized on a solid support. Glycan-binding proteins such as lectins and glycan-binding antibodies can be incubated on a glycan array to characterize relative binding to each oligosaccharide. From such data, much information can be gathered regarding the specificity of the glycan-binding protein. Glycan arrays have grown in complexity and usefulness to the glycobiology community since their introduction in 2002. Access to this technology has been provided by the Consortium for Functional Glycomics, which offers a service to run glycan microarray experiments using samples provided by individual investigators.

While glycan array data have great potential for characterizing glycan-binding specificity, the interpretation of the data can be challenging. The difficulty is particularly great for proteins that bind multiple structures or that have varying affinity depending on the presentation or overall context of a particular structure. For many proteins, the primary glycan-binding specificity is known, but details about the fine specificity, such as preferred

*Corresponding author: Brian B. Haab, Ph.D., Van Andel Research Institute, 333 Bostwick NE, Grand Rapids, MI 49503, (616) 234-5269, brian.haab@vai.org.

presentations of binding determinants or potentially blocking side chains, are not clear. The manual analysis and interpretation of complex glycan array data can be time consuming and subject to user bias. Objective and automated methods of analyzing glycan array data could address those difficulties.

Previously we introduced the Motif Segregation method as an approach to systematizing and automating the analysis of glycan array data (8) and further developed the method using Outlier Motif Analysis (9). Motif Segregation operates by identifying the motifs (component parts of oligosaccharides) that are selectively present in the glycans on a glycan array that are strongly bound by a particular lectin, and Outlier Motif Analysis is an approach to refining and improving the output from Motif Segregation (described in more detail below). We demonstrated the accuracy of Motif Segregation for extracting the primary binding specificities of a wide variety of glycan-binding proteins and the use of Outlier Motif Analysis for a more detailed definition of binding specificity.

Recently we developed software to facilitate the analysis of glycan array data by these approaches and additional methods. This chapter describes the theory behind the analysis methods, the use of the software, and an example application. The software is available upon request to the authors.

2. Materials

The GlycoSearch software tool analyzes glycan array data generated on any glycan array platform. Several different versions of glycan arrays have been produced, including arrays that mostly contain variants of sialic acids, and the software will work for each, provided the sequences of the glycans on the array are available. In a typical experiment, a purified glycan-binding protein (GBP, such as a lectin or glycan-binding antibody) is incubated on an array, and the level of binding of the GBP to each glycan on the array is quantified (Fig. 1A). An example method of quantitation is to use a biotinylated GBP, probe the array with dye-conjugated streptavidin, and scan the array for fluorescence. The output data from a glycan array experiment consists of a list of the glycans on the array followed by a numerical quantification of the fluorescence detected at each glycan (Fig. 1A). Other parameters could be associated with each glycan such as the level of background signal around the glycan spot or the standard deviation between pixels within the spot, which may be useful for quality control purposes. The output file, saved either as a tab-delimited text file or as a Microsoft Excel workbook, can be used directly as the input for the GlycoSearch software, which then produces two output files (Fig. 1B). The glycan names should be in a condensed IUPAC notation as defined by the Consortium for Functional Glycomics (CFG, www.functionalglycomics.org) and in the *Essentials of Glycobiology* textbook (10). Some deviations from IUPAC notation can be properly interpreted, and the software is flexible enough to catch and automatically correct common errors in glycan expressions. The software is written in Java and runs on various computer platforms, including Windows, Macintosh, and Linux platforms (Note 1).

¹Be sure the latest version of Java is installed on your computer. Java is freely available on the web by download. The program must be resident on the local disk of the computer.

3. Methods

3.1. Overview of the methods used in GlycoSearch

Motif segregation—The principles of the Motif Segregation approach were described in detail elsewhere (8) and are briefly reviewed here. Motifs, or component substructures of oligosaccharides, are predefined by the software. Examples include the blood group A antigen, terminal β -linked galactose, and internal lactosamine. The initial release of the software starts with over 200 such predefined motifs. For every glycan on an array, the software determines whether each motif is present or absent. Based on that determination, a matrix of the glycans and motifs (in the rows and columns of the matrix, respectively) is populated with 1s and 0s indicating presence or absence, respectively, of each motif in each glycan. Glycan array data is provided to the program which indicates the signal intensity at each glycan (corresponding to the amount of binding of a GBP). For each motif, the program statistically compares the signals of the glycans that contain the motif to those that do not contain the motif. Using the Mann-Whitney test, a p value is generated for each motif indicating the likelihood that the observed pattern of signals could be generated by chance. For ease of comparison between motifs, the software takes the logarithm (base 10) of the p value and adds a +/- sign, with positive indicating that motif-containing glycans have a higher average intensity, and negative indicating the opposite. This signed, logged p value is referred to as the motif score.

Motif refinement through combination motifs and outlier analysis—Because the set of predefined motifs may not include the precise binding determinant of a given glycan-binding protein, the software includes two methods for further refining the motif definitions. The first employs combination motifs, and the second employs outlier analysis.

The combination motifs method works by the iterative combining of motifs to test whether any combinations add to or take away from the strength of the motif score. The top scoring motif is combined with another motif through an 'OR' operator. For example, if a glycan contains either the polylectosamine motif OR the α 2,3 sialic acid motif, it would contain the new combination motif. The program individually calculates the motif score for the top-scoring motif combined with each of the other motifs. The combination with the best score is recorded. The program then looks for motifs to exclude using an AND NOT operator. For example, a glycan could have a new combination motif consisting of either the polylectosamine motif OR the α 2,3 sialic acid motif, AND NOT 6' sulfation. In this way, motifs that have a negative effect on lectin binding could be identified. This process is continued as long as the score is improved by adding new combinations. The program records each best combination and the score improvement resulting from the combination. This process enables a refinement of the motifs and an exploration of unexpected modifiers to binding. An example of the output and its interpretation are given below.

Outlier analysis (9) provides additional information for the refinement of motifs. Glycans that contain motifs with high motif scores would be expected to have a high signal, indicating strong binding from the GBP. Likewise, glycans that do not have any motifs with high motif scores would be expected to have a low signal. However, the array data sometimes show exceptions to this rule. Outlier glycans, which either 1) contain high-

scoring motifs yet have low signal or 2) do not contain high-scoring motifs yet have high signal, can provide more detailed information about the specificity of GBPs. While the primary binding specificity may be accurately revealed by the top-scoring motifs, the fine specificity of a GBP can be probed by comparing the structures of the outlier glycans to the non-outlier glycans. The GlycoSearch software provides a convenient way to do that, described in more detail below.

3.2. Using GlycoSearch

Primary analysis—The GlycoSearch software is simple to use. Upon starting the program, a dialog box asks for the file to be analyzed. As described above, the input file should be an Excel workbook containing columns for the glycans names, the signal associated with each glycan, and a standard deviation associated with each spot. Specific column headers are not required, since the program will recognize the data types of each column. The software processes the data and provides two output files. The first is an Excel workbook that contains the results and that is named the same as the input filename, appended with “_lectin_out.” For example an input file named ConA would produce an output file named ConA_lectin_out. The second is a text file that contains a description of the binding motifs learned by the analysis and that is named the same as the input filename, appended with “_learned_motifs.”

The primary information is in the _lectin_out workbook, in a tab called “Motif Scores.” The first three rows list all the motifs and the score for each, and the next three rows show the motifs sorted by motif score, in decreasing score from left to right. Use this information to obtain the primary binding specificity of the lectin. The highest motif scores are most significant, and a motif score of 1.3 indicates $p = 0.05$ ($\log_{10}(0.05) = 1.3$), which can be a useful threshold for significance. The subsequent rows give the glycans in descending order of intensity along with a list of the motifs that are represented in each glycan and the sum of the motif scores above 3.0. This information is useful for observing which motifs are present or not present in the high-signal glycans.

Motif refinement through combination motifs—The results from the combination motif analysis (described above) are provided at the bottom of the Motif Scores tab. Each iteration listed shows the outcome of searches for combinations that improve the score. The score of the combination is given, followed by the gain over the previous best combination (for the first iteration it is the gain over the highest individual motif score), the total number of glycans included in the new combination, and the motif used in the combination (both the motif ID and motif name are given). If exclusions using the AND NOT operator produce an improvement in the motif score, the excluded motifs are listed as well. The details of each iteration step and more information about the motifs added or excluded at each step are provided in the _learned_motifs text file. The example application section below provides more information on how to interpret and use this information.

Motif refinement through outlier analysis—The second tab of the _lectin_out file, called “Outlier Analysis,” provides a convenient summary of outlier glycans, which can be used to further refine the understanding of the GBP binding specificity. This sheet lists the

motifs, sorted by motif score. Following each motif is the total number of glycans containing that motif and the numbers of positive outlier glycans, negative outlier glycans, and correct glycans. Positive outliers contain the motif but have low signals, negative outliers do not contain the motif but have high signals, and correct glycans both contain the motif and have high signals. One can use this information to assess the correspondence between the presence of a motif and the binding of the GBP. A high level of positive outliers suggests that the GBP does not bind the motif in certain presentations, such as if substituted in a certain way. A high level of negative outliers suggests that the GBP also binds other motifs. To gain more insight into what causes the outliers, the next step is to compare the structures of the outlier glycans to the non-outlier glycans, which is explained more in the example application below.

Exploring user-defined motifs—The study of the motif analyses may suggest that additional motifs not represented in the predefined list can be important for the binding of the GBP. In that case, the user would want to introduce other newly-defined motifs in order to examine the possibility of obtaining further improvements in motif scores. The software allows that type of analysis through the option of including user-defined motifs. The list of new motifs should be placed in a new Excel file with the lectin name followed by “_user_motifs.xlsx.” For example, if the file containing the lectin data were named “BPL.xlsx,” the user motifs file should be named “BPL_user_motifs.xlsx,” and the two files should be in the same directory. The program will add the user-defined motifs to the analysis (Note 2). A supplementary file with detailed guidelines for the motif naming conventions is available separately from the authors. The program can accept any combination of standard IUPAC-like motif expressions (component parts of oligosaccharides) with logical operators such as AND, OR, and NOT. The software can additionally accept any combination of the pre-defined motif terms used by the conventional Consortium for Functional Glycomics workbook for lectin binding analysis.

3.2. Example application

Here we provide an example of the analysis and interpretation of an individual glycan array dataset. The *bauhinia purpurea* lectin (BPL) is isolated from a flowering deciduous tree native to southeast Asia. The primary specificity of BPL is known to be terminal, beta-linked galactose (11), but additional details about its specificity would be useful for the interpretation of data obtained using BPL as an analytical reagent. We used GlycoSearch to analyze data available from the CFG for the incubation of BPL on array version 2.0.

Primary specificities—The first step was to examine the top-scoring individual motifs. Table 1 lists the motifs with a score above 2.0, taken from the _lectin_out.xlsx output file. The top-scoring motif (score = 5.99) is “Terminal Galactose AnyLinkage,” which is consistent with the known specificity of BPL. (The AnyLinkage modifier indicates that the motif includes both beta-linked and alpha-linked terminal galactose.) The next two highest

²GlycoSearch has several switches that allow more detailed analyses or customization. The batch file that starts the program can be edited using a standard text editor to add the switches. For example, to obtain detailed information on how each motif was interpreted, add the line “-CHECK_ARRAY_SPREADSHEET = 1” to the file. This option is useful to check whether a user-defined motif was properly interpreted and to see the glycans in which the motif was found. A list of available switches is provided with the software.

motifs also are terminal galactose but specific to a beta linkage. The next motifs have weaker but still significant scores and suggest that BPL has broader specificity than just the top motif. For example, motifs 32, 188, 127, 164, and 30 all suggest that BPL can bind alpha-linked galactose or N-acetylgalactosamine (GalNAc) under certain circumstances. The simple motifs of Terminal Gal α 1,2; Terminal Gal α 1,3; and Terminal Gal α 1,4 scored 1.04, 1.98, and 0.30, respectively, indicating that terminal, alpha-linked galactose is bound by BPL only when attached to the 3' carbon of the next monosaccharide.

Combination motifs—To gain additional insights into BPL specificity, the combination motifs table (found at the bottom of the Motif Scores tab in the `_lectin_out.xlsx` workbook) can be consulted (Table 2). The first line of the table gives the top-scoring individual motif, with which other motifs are combined in subsequent iterations. In each iteration, the program tests the combination of the top motif with every other individual motif using an OR operator and the exclusion of individual motifs using the AND NOT operator. In the first iteration of this example, the combination of motif 19 with motif 23 (Terminal GalNAc β) and exclusion of several motifs resulted in 68 glycans with the combined motif. The combined motif had a score of 10.39, which is a major gain over the starting score of 5.99. This result indicates that the binding specificity of BPL primarily encompasses both terminal, beta-linked Gal and GalNAc, but with exclusions. Those exclusions are mainly sulfation and sialylation, which are known to prevent BPL binding. The individual motif score of Terminal GalNAc β was 2.06 (Table 1), far less than the score for Terminal Gal β (score = 5.31), indicating that BPL prefers Gal over GalNAc. However, the binding to GalNAc β is significant on its own.

The next two iterations show that BPL could also bind both alpha-linked GalNAc and GlcNAc under certain circumstances (Table 2). The individual motif scores for motifs 107 (Tn antigen, GalNAc α -Serine/Threonine) and 59 (Terminal GlcNAc α 1,6) were 1.64 and 1.53, respectively, which are weakly significant but suggestive of real binding. The following iterations in Table 2 reinforce the above interpretations and provide information on specific circumstances in which BPL binds. The overall picture that emerges is that BPL binding is broader than simply beta-linked, terminal Gal, encompassing beta-linked GalNAc and alpha-linked GalNAc and GlcNAc in certain configurations, although with weaker preference.

Outlier analysis and user-defined motifs—The outlier analysis and the addition of user-defined motifs could be used in conjunction with the combined motif analysis to arrive at a more complete description of the specificity of a GBP. To identify outlier glycans, the correlation between the signal intensity of each glycan and the summed motif scores (of all scores 3.0) (both parameters are given in the Motif Scores tab of the `_lectin_out.xlsx` workbook) can be examined (Note 3) (Fig. 2A). Intensity thresholds are calculated by the program that identify low-signal glycans (below the bottom threshold), high-signal glycans (above the top threshold), and moderate-signal glycans (between the thresholds). An

³In the initial version of the software, the user must manually make the plots and perform the comparisons. The future versions of the software are planned to include automated generation of outlier plots, comparisons of motifs between outlier and non-outlier glycans, and generation of candidate new motifs.

examination of the BPL outlier plot (Fig. 2A) reveals several positive outliers (high summed motif score but low signal) and negative outliers (low summed motif score but high signal), as presented previously (9). The positive outliers were mostly sulfated glycans (our definition of “Terminal Gal β ” encompassed sulfated Gal β , so sulfated glycans had a high summed-motif score even though they did not bind BPL), while the negative outliers were mostly terminal GalNAc α or terminal/penultimate Gal α sequences for which we previously did not have specific motifs (these motifs and many others are now included in GlycoSearch).

The insights from the comparisons of the outliers to the non-outliers can be used to form additional motifs that could be added into the analysis. In addition, the combination motif analysis described above also could suggest new motifs. For example, both analyses show that sulfation should be excluded and that GalNAc α 1,3Gal without fucose on the Gal should be added. New motifs can be added to the analysis by creating a new Excel spreadsheet containing the motif names (using the guidelines given above) in individual rows.

Based on an analysis of the original output, we defined two new motifs: Terminal GalNAc α 1-3Gal Inclusive No Subterminal Fucose (motif 217), and Terminal GlcNAc β 1-6Gal Inclusive (motif 218). Running GlycoSearch again with the new motifs showed that motif 217 was moderately significant (motif score = 1.72) and motif 218 was not significant (motif score = 1.13) on their own. However, both motifs were used in the combination motifs, with motif 217 included in the third and motif 218 in the seventh iterations, resulting in an overall gain of 0.4 in the motif score. The success of new motifs in describing binding specificity can be visually checked by creating an outlier plot after the reanalysis (Fig. 2B). Fewer outlier glycans should be present if the new motifs have properly accounted for the previous outliers. Work is underway on the automated generation of new motifs, which could further streamline the analysis to rapidly arrive at the best description of binding specificity.

Acknowledgments

The authors gratefully acknowledge funding of this work by a bridging grant from the Consortium for Functional Glycomics and by the Van Andel Research Institute.

References

1. Bathe OF, Shaykhtudinov R, Kopciuk K, Weljie AM, McKay A, Sutherland FR, Dixon E, Dunse N, Sotiropoulos D, Vogel HJ. Feasibility of identifying pancreatic cancer based on serum metabolomics. *Cancer Epidemiol Biomarkers Prev.* 2010; 20:140–147. [PubMed: 21098649]
2. Lo JF, Yu CC, Chiou SH, Huang CY, Jan CI, Lin SC, Liu CJ, Hu WY, Yu YH. The epithelial-mesenchymal transition mediator s100a4 maintains cancer initiating cells in head and neck cancers. *Cancer Res.* 2010; 71:1912–1923. [PubMed: 21169409]
3. Drickamer K, Taylor ME. Glycan arrays for functional glycomics. *Genome Biol.* 2002; 3:REVIEWS1034. [PubMed: 12537579]
4. Yue T, Haab BB. Microarrays in glycoproteomics research. *Clin Lab Med.* 2009; 29:15–29. [PubMed: 19389548]
5. Manimala JC, Roach TA, Li Z, Gildersleeve JC. High-throughput carbohydrate microarray analysis of 24 lectins. *Angew Chem Int Ed Engl.* 2006; 45:3607–3610. [PubMed: 16639753]

6. Blixt O, Head S, Mondala T, Scanlan C, Huflejt ME, Alvarez R, Bryan MC, Fazio F, Calarese D, Stevens J, Razi N, Stevens DJ, Skehel JJ, van Die I, Burton DR, Wilson IA, Cummings R, Bovin N, Wong CH, Paulson JC. Printed covalent glycan array for ligand profiling of diverse glycan binding proteins. *Proc Natl Acad Sci U S A*. 2004; 101:17033–17038. [PubMed: 15563589]
7. Lee SM, Chan RW, Gardy JL, Lo CK, Sihoe AD, Kang SS, Cheung TK, Guan YI, Chan MC, Hancock RE, Peiris MJ. Systems-level comparison of host responses induced by pandemic and seasonal influenza A H1N1 viruses in primary human type I-like alveolar epithelial cells in vitro. *Respir Res*. 2010; 11:147. [PubMed: 21029402]
8. Porter A, Yue T, Heeringa L, Day S, Suh E, Haab BB. A motif-based analysis of glycan array data to determine the specificities of glycan-binding proteins. *Glycobiology*. 2010; 20:369–380. [PubMed: 19946132]
9. Maupin KA, Liden D, Haab BB. The fine specificity of mannose-binding and galactose-binding lectins revealed using outlier-motif analysis of glycan array data. *Glycobiology*. 2011; 22:160–169. [PubMed: 21875884]
10. Varki, A.; Cummings, R.; Esko, J.; Freeze, H.; Stanley, P.; Bertozzi, CR.; Hart, G.; Etzler, ME. *Essentials of glycobiology*. 2. Cold Spring Harbor Laboratory Press; Cold Spring Harbor, NY: 2009.
11. Wu AM, Wu JH, Liu JH, Singh T. Recognition profile of bauhinia purpurea agglutinin (bpa). *Life Sci*. 2004; 74:1763–1779. [PubMed: 14741734]

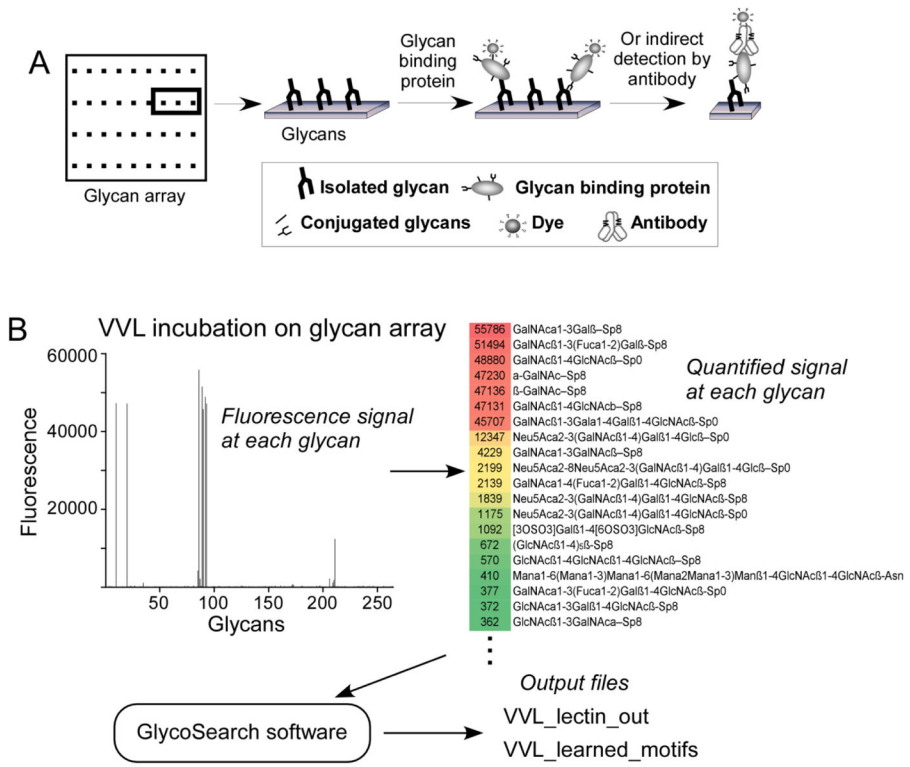


Figure 1. Glycan array data processing

(A) Glycan arrays consist of purified glycans immobilized in discrete locations on slides. A glycan-binding protein is incubated on the slide to allow binding to any glycans. The amount of bound protein at each glycan spot is determined through measuring fluorescence from a tag on the glycan-binding protein or on a secondary detection reagent. (B) The fluorescence signals from the binding of the lectin *vicia villosa* (VVL) to 215 different glycans were quantified (left panel) and placed in an Excel file containing the names of the glycans and the corresponding signals (right panel), which served as the input for the GlycoSearch software. The software outputs two files: an Excel file containing the main analysis, and a text file providing details of the motif refinement by combination motifs.

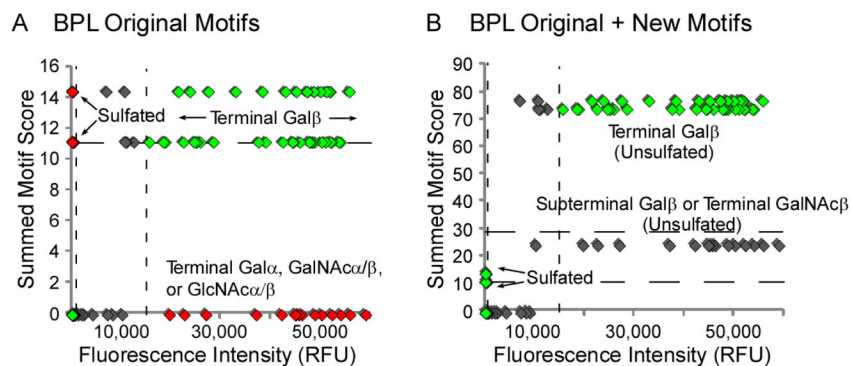


Figure 2. Outlier-motif analysis of the galactose binder BPL

The summed motif scores for each glycan, after analysis with either the original or updated motifs, are plotted with respect to fluorescence intensity after detection with BPL. (A) BPL using the original motifs. (B) BPL using the original plus new motifs. The dashed lines represent thresholds for defining outliers, based on the distributions from all the glycans. Green symbols indicate glycans in expected regions, red symbols indicate outliers, and grey symbols indicate bound glycans that may not represent the primary specificity of the lectin.

Table 1

Individual motif scores for BPL.

Motif ID	Motif Name	Score
19	Terminal Galactose AnyLinkage	5.99
5	Terminal Galb	5.31
2	Terminal Galb1,3	5.27
32	GalNAca 6' Substituted	3.08
113	Core 2 O-glycan	2.90
188	Gala1-3Gal Not Blood Group B	2.85
127	O Glycan (GalNAca-Sp)	2.81
164	Linear B	2.76
30	GalNAca 3' Substituted	2.72
173	Lewis x	2.50
169	Lewis a	2.46
139	Type 3 Chain (Galb1-3GalNAca) Poly LacNAc OR Neo Poly LacNAc	2.45
148	Terminal	2.40
112	Terminal Core 1 O-glycan	2.18
23	Terminal GalNAcb	2.06

Table 2

Combination motif scores

Each iteration shows the motif that was added and the exclusions applied to that glycan. The Exclusions column gives the ID of the excluded motifs and the number of excluded glycans in square brackets. The names of the excluded motifs are given at the bottom of the table.

Itr	Score	Gain	Glycans	ID	Motif	Exclusions
0	5.99	-	109	19	Terminal Galactose AnyLinkage	180 [21]; 85 [28]; 40 OR 13 [1]; 185 [4]
1	10.39	4.40	68	23	Terminal GalNAcb	168 OR 94 [8]
2	11.03	0.65	61	107	Tn Antigen	
3	11.21	0.18	62	59	Terminal GlcNAc1-6	
4	11.37	0.16	63	25	GalNAcb 4' Substituted	
5	11.49	0.11	64	121	Core 6 O-glycan	
6	11.55	0.07	65	189	Forssman Antigen (GalNAc1-3GalNAcb) Not Blood Group A	
7	11.61	0.06	66	188	Gala1-3Gal Not Blood Group B	
8	11.67	0.06	67	161	Blood Group H Type 3	
9	11.72	0.05	68	207	MDPLys	
10	11.76	0.04	69	181	Phosphorylated Glycans	
11	11.79	0.03	70	155	Blood Group B Type 2	
12	11.82	0.03	70	31	GalNAca 4' Substituted	
13	11.87	0.05	71	207	MDPLys	51 [2]
14	11.90	0.04	72	166	Blood Group I GicNAcb1-3(GlcNAcb1-6)Galb	
15	11.92	0.01	73	91	Rhamnose Anywhere	
16	11.93	0.01	74	186	Internal Gala1-4Galb	94 OR 95 [1]

Excluded Motifs

180	Sulfated Glycans
85	Fuca1-2
40	Glc6' Substituted
13	Terminal Gala1-6
185	Terminal Gala1-4Galb
168	Blood Group Sd(a), CAD (GalNAcb1-4(Siaa2-3)Gal)
94	Neu5Aca2-3 Anywhere
51	Terminal GlcNAcb1-4

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

ID	Excluded Motifs
94	Neu5Aca2-3 Anywhere
95	Terminal Neu5Aca2-3