## PHA-E

# Automated Analysis report for PHA-E evaluated at 1 ug/mL

#### List of Primary PHA-E Motifs



Minimal and complete motif definitions match the same set of glycans defined in the array. Components found in the complete motif but excluded in the minimal motif are not tested by the array. Monosaccharide identities and subsitution intolerance may or may not have been tested by the array, they are retained from the complete motif for readability.

#### **Boxplot of Primary Motifs**



### List of Fine-Specificity PHA-E Motifs

	Nearest Common Name		Relative	Number of	
Motif ID	(Accuracy%**)	Motif Graphic Structure	Binding	Glycans	P-Value***
A0*		$R \xrightarrow{3} \beta^{\beta 4} \overbrace{\beta^{\beta 2}}^{\beta 2} \overbrace{\alpha}^{\alpha} \underset{R}{\beta^{\beta 4}}^{\beta} R$	1.00	8	<0.001
A1		$R \xrightarrow{3} \beta^{4} \beta^{2} \alpha^{\alpha} R$ $R \xrightarrow{4} \beta^{2} \alpha^{\alpha} R$	0.74	4	<0.001
BO		$\mathbf{R} \xrightarrow{4} \mathbf{B}^{\beta 2} \mathbf{O}_{\alpha}$ $\mathbf{R} \xrightarrow{4} \mathbf{B}^{6} \mathbf{B}$ $\mathbf{R}$	0.12	36	0.013
0	Non-Binders (100%)		0.00	24	NA





See Symbol Nomenclature for Glycans (SNFG) for complete key: https://www.ncbi.nlm.nih.gov/glycans/snfg.html (https://www.ncbi.nlm.nih.gov/glycans/snfg.html)

PHA-E

\*Motif indicates the remaining glycans not matched by motifs which are a subset. Motif definition needs to be taken in the context of the model.

\*\*Accuracy describes the consistency between common-name definition of the motif and the formal, text-based definition of the motif, in terms of percent agreement in the glycans containing the two motifs. Common Name label definitions given here (http://carbogrove.org/MotifLabels.php).

\*\*\*P-Value refers to difference from Non-Binders with multiple testing correction (Dunnet's Test)

Motifs with a red motif ID fail to show a logistic response to protein concentration in the range of concentrations analyzed. These motifs may be nonbinding motifs (motifs which define nonbinding exceptions) or simply fail to fit. Nonbinding motifs are determined based on concentration dependent response when available or the average binding of non-motif glycans otherwise.



#### **Boxplot of Fine-Specificity Motifs**

Figure 1. Glycan binding grouped by motif and motif family. Individual glycans are given as points on the plot.

#### Motif Intensity Map

PHA-E



Figure 2. Glycan intensity and motif distribution plot. The top half of the plot presents the observed glycan binding intensity of various glycans used in the array over their rank binding intensity; only the top glycans are shown. The second plot indicates the position of glycans containing the various motifs in the top plot with a yellow tick.

#### Motif Family Membership Map



\*Motif indicates the remaining glycans not matched by motifs which are a subset. Motif definition needs to be taken in the context of the model.

Figure 3. Treemap of glycan binding grouped by motif and family structure. The model structure can be represented as nested boxes where box size is proportional to the number of glycans with the motif and color changes with change in average relative binding of glycans with the motif. Only three layers of data splitting are included here, though further splitting may be possible.

#### **Detailed Model Breakdown**

Motif Glycan Examples:



PHA-E



Figure 4. Boxplots of glycan binding grouped by motif for each dataset in the model. Motifs are listed in ascending average binding intensity (for the selected concentration) and colored by family.

Model Structure:



\*Motif matches the remaining glycans not matched by earlier motifs in the model.

Figure 5. Tree representation of the regression tree model trained on array data. Data flows through the tree (topdown) and is split by the various motifs. The motif used the split the data at each point has the id "family+split number" except when further split. In the case of futher splits the id of the motif used to split the data is denoted with an asterisk.

Curve Fitting:

No curves were fit for model

Motif Text Structures:



PHA-E



# Automated Analysis report for PHA-L evaluated at 10 ug/mL

#### List of Primary PHA-L Motifs

Primary Motif ID	Primary Motif Minimal Structure	Primary Motif Complete Structure	Relative Binding	Number of Glycans
A	$\mathbf{R} = \mathbf{R} = \mathbf{R} = \mathbf{R}$	$\mathbf{R} = \begin{bmatrix} \mathbf{B} & \mathbf{B} & \mathbf{B} \\ \mathbf{B} & \mathbf{B} & \mathbf{B} \\ \mathbf{R} & \mathbf{R} & \mathbf{B} \\ \mathbf{R} & \mathbf{R} & \mathbf{R} $	1.00	18
0			0.02	54

Minimal and complete motif definitions match the same set of glycans defined in the array. Components found in the complete motif but excluded in the minimal motif are not tested by the array. Monosaccharide identities and subsitution intolerance may or may not have been tested by the array, they are retained from the complete motif for readability.

#### **Boxplot of Primary Motifs**





### List of Fine-Specificity PHA-L Motifs

Motif ID	Nearest Common Name (Accuracy%**)	Motif Graphic Structure	Relative Binding	Number of Glycans	P-Value***
A0*		$\mathbf{R} = \begin{bmatrix} \mathbf{R} & \mathbf{B} & \mathbf{B} & \mathbf{B} \\ \mathbf{R} & \mathbf{B} & \mathbf{B} & \mathbf{B} \\ \mathbf{R} & \mathbf{B} & \mathbf{B} & \mathbf{B} \\ \mathbf{R} & \mathbf{R} & \mathbf{R} \end{bmatrix}$	1.00	9	<0.001
A1		$R \xrightarrow{6}{3} \xrightarrow{\beta4} \xrightarrow{\beta2} \alpha$ $R \xrightarrow{6}{3} \xrightarrow{\beta4} \xrightarrow{\beta2} \alpha$ $R \xrightarrow{6}{3} \xrightarrow{\beta4} \xrightarrow{6} \xrightarrow{\beta} \xrightarrow{6} \xrightarrow{6} \xrightarrow{6} \xrightarrow{6} \xrightarrow{6} \xrightarrow{6} \xrightarrow{6} 6$	0.36	9	0.002
0	Non-Binders (100%)		0.00	54	NA
Key:					
Hexose () Glc () Man () Gal ()	HexNAc Kdn A GlcNAc Neu5Ac Mon GalNAc Neu5Gc Mon Fuc Sia	ny Single osaccharide vny # of osaccharide ~SKIP~ drogen or R- osaccharide			

See Symbol Nomenclature for Glycans (SNFG) for complete key: https://www.ncbi.nlm.nih.gov/glycans/snfg.html (https://www.ncbi.nlm.nih.gov/glycans/snfg.html)

\*Motif indicates the remaining glycans not matched by motifs which are a subset. Motif definition needs to be taken in the context of the model.

\*\*Accuracy describes the consistency between common-name definition of the motif and the formal, text-based definition of the motif, in terms of percent agreement in the glycans containing the two motifs. Common Name label definitions given here (http://carbogrove.org/MotifLabels.php).

\*\*\*P-Value refers to difference from Non-Binders with multiple testing correction (Dunnet's Test)

Motifs with a red motif ID fail to show a logistic response to protein concentration in the range of concentrations analyzed. These motifs may be nonbinding motifs (motifs which define nonbinding exceptions) or simply fail to fit. Nonbinding motifs are determined based on concentration dependent response when available or the average binding of non-motif glycans otherwise.



#### **Boxplot of Fine-Specificity Motifs**

Figure 1. Glycan binding grouped by motif and motif family. Individual glycans are given as points on the plot.

#### Motif Intensity Map



Figure 2. Glycan intensity and motif distribution plot. The top half of the plot presents the observed glycan binding intensity of various glycans used in the array over their rank binding intensity; only the top glycans are shown. The second plot indicates the position of glycans containing the various motifs in the top plot with a yellow tick.

#### Motif Family Membership Map



\*Motif indicates the remaining glycans not matched by motifs which are a subset. Motif definition needs to be taken in the context of the model.

Figure 3. Treemap of glycan binding grouped by motif and family structure. The model structure can be represented as nested boxes where box size is proportional to the number of glycans with the motif and color changes with change in average relative binding of glycans with the motif. Only three layers of data splitting are included here, though further splitting may be possible.

#### **Detailed Model Breakdown**

Motif Glycan Examples:







Figure 4. Boxplots of glycan binding grouped by motif for each dataset in the model. Motifs are listed in ascending average binding intensity (for the selected concentration) and colored by family.

Model Structure:



\*Motif matches the remaining glycans not matched by earlier motifs in the model.

Figure 5. Tree representation of the regression tree model trained on array data. Data flows through the tree (topdown) and is split by the various motifs. The motif used the split the data at each point has the id "family+split number" except when further split. In the case of futher splits the id of the motif used to split the data is denoted with an asterisk.

#### Curve Fitting:

No curves were fit for model

Motif Text Structures:

#### Motif

ID	Motif Graphic	Motif Text
A1	R 0 0 4 0 2 0 0 R	<6f8f>Neu5AcA2-<3or6><2f4f>GalB1-4<6f>GlcNAcB1-2<3f4f6f>ManA1-3(<2f4f>GalB1- 4<3f6f>GlcNAcB1-2<3f4f6f>ManA1-6)(<3f4f6f>GlcNAcB1-4)<2f>ManB
A0*	R R R R R R R R R R R R R R R R R R R	<2f4f>GalB1-4<3f6f>GlcNAcB1-2<3f4f6f>ManA1-6(<3f4f6f>GlcNAcB1-4)<2f>ManB
0		Non-Binders

Key:

10/16/2020



PHA-L

# Automated Analysis report for Calsepa evaluated at 2 ug/mL

#### List of Primary Calsepa Motifs



Minimal and complete motif definitions match the same set of glycans defined in the array. Components found in the complete motif but excluded in the minimal motif are not tested by the array. Monosaccharide identities and subsitution intolerance may or may not have been tested by the array, they are retained from the complete motif for readability.

### **Boxplot of Primary Motifs**

10/16/2020

Calsepa



### List of Fine-Specificity Calsepa Motifs

Motif ID	Nearest Common Name (Accuracy%**)	Motif Graphic Structure	Relative Binding	Number of Glycans	P-Value***
A1		$R \xrightarrow{4} B^{2} \bigcirc \alpha$ $R \xrightarrow{6} B^{4} B^{2} \bigcirc \alpha$ $R \xrightarrow{6} B^{4} B^{2} \bigcirc \alpha$	1.00	8	<0.001
A0*		$R \xrightarrow{6} \beta^{\beta 4} \xrightarrow{\beta 2} \alpha^{\alpha} R$	0.56	16	<0.001
B1	Terminal N-Glycan N- Acetyl Glucosamine Bisecting N-Glycan (90%)	$\mathbf{R} \xrightarrow{4} \mathbf{B}^{2} \mathbf{C}^{\alpha}$	0.69	4	<0.001

Nearest Con Name Motif ID (Accuracy <sup>o</sup>	nmon %**) Motif Graphic Structure	Relative Binding	Number of Glycans	P-Value***
B0*	$\mathbf{R}$	0.30	8	0.001
0 Non-Binders (	100%)	0.00	36	NA
Key:				
Hexose HexNAc Kdn Glc GlcNAc Neu5Ac Man GalNAc Neu5Gc Cal	Any Single Monosaccharide Any # of Monosaccharide ~SKIP~ Hydrogen or B-			

See Symbol Nomenclature for Glycans (SNFG) for complete key: https://www.ncbi.nlm.nih.gov/glycans/snfg.html (https://www.ncbi.nlm.nih.gov/glycans/snfg.html)

\*Motif indicates the remaining glycans not matched by motifs which are a subset. Motif definition needs to be taken in the context of the model.

\*\*Accuracy describes the consistency between common-name definition of the motif and the formal, text-based definition of the motif, in terms of percent agreement in the glycans containing the two motifs. Common Name label definitions given here (http://carbogrove.org/MotifLabels.php).

\*\*\*P-Value refers to difference from Non-Binders with multiple testing correction (Dunnet's Test)

Motifs with a red motif ID fail to show a logistic response to protein concentration in the range of concentrations analyzed. These motifs may be nonbinding motifs (motifs which define nonbinding exceptions) or simply fail to fit. Nonbinding motifs are determined based on concentration dependent response when available or the average binding of non-motif glycans otherwise.

#### **Boxplot of Fine-Specificity Motifs**





Figure 1. Glycan binding grouped by motif and motif family. Individual glycans are given as points on the plot.

#### Motif Intensity Map



Figure 2. Glycan intensity and motif distribution plot. The top half of the plot presents the observed glycan binding intensity of various glycans used in the array over their rank binding intensity; only the top glycans are shown. The second plot indicates the position of glycans containing the various motifs in the top plot with a yellow tick.

#### Motif Family Membership Map

B 1		A 1
0*		0*
# Family Sub	-Family Member *Family # Remainder	

\*Motif indicates the remaining glycans not matched by motifs which are a subset. Motif definition needs to be taken in the context of the model.

Figure 3. Treemap of glycan binding grouped by motif and family structure. The model structure can be represented as nested boxes where box size is proportional to the number of glycans with the motif and color changes with change in average relative binding of glycans with the motif. Only three layers of data splitting are included here, though further splitting may be possible.

#### Detailed Model Breakdown

Motif Glycan Examples:





Figure 4. Boxplots of glycan binding grouped by motif for each dataset in the model. Motifs are listed in ascending average binding intensity (for the selected concentration) and colored by family.

Model Structure:



\*Motif matches the remaining glycans not matched by earlier motifs in the model.

Figure 5. Tree representation of the regression tree model trained on array data. Data flows through the tree (topdown) and is split by the various motifs. The motif used the split the data at each point has the id "family+split number" except when further split. In the case of futher splits the id of the motif used to split the data is denoted with an asterisk.

Curve Fitting:

No curves were fit for model

Motif Text Structures:



Motif ID	Motif Graphic	Motif Text
B1	$\mathbf{R} \xrightarrow{4} \mathbf{B}^{\underline{\beta}} \mathbf{C} \mathbf{C}$	<3f4f6f>GlcNAcB1-2<3f4f6f>ManA1-3(<3f6f>GlcNAcB1-2<3f4f6f>ManA1-6) (<3f4f6f>GlcNAcB1-4)<2f>ManB
B0*	$\mathbf{R}^{\mathbf{\beta}}$	<3f4f6f>GlcNAcB1-2<3f4f6f>ManA1-3<2f>ManB
0		Non-Binders
Key:		
Hexose () Gic () Man () Gal ()	HexNAc Kdn GlcNAc Neu5Ac GalNAc Neu5Gc Fuc Sia	Any Single Monosaccharide Any # of Monosaccharide ~SKIP~ Hydrogen or Monosaccharide R-

# Automated Analysis report for MAL-I evaluated at 10 ug/mL

#### List of Primary MAL-I Motifs

Primary Motif ID	Primary Motif Minimal Structure	Primary Motif Complete Structure	Relative Binding	Number of Glycans
A	$ \overset{\alpha 3}{\overset{\beta 4}{\overset{\beta 4}{\overset{\beta 2}{\overset{\alpha 6}{\overset{\beta 6}$	$ \begin{array}{c} & \alpha^{3} \\ & \beta^{4} \\ & \beta^{2} \\ & \beta^{4} \\ & \beta^{2} \\ & \beta^{4} $	1.00	6
В	$R \xrightarrow{4} \square^{\beta 2} \square^{\alpha} \xrightarrow{\beta} \square^{\beta} R$	$R^{-4} \xrightarrow{\beta^2} \alpha_{6}^{\alpha} \xrightarrow{\beta^4} \beta^4$ $R^{-3} \xrightarrow{\beta^4} \beta^2 \xrightarrow{\alpha^3} \beta^4 \xrightarrow{\beta^4} \beta^4$	0.23	6
0			0.01	60

Minimal and complete motif definitions match the same set of glycans defined in the array. Components found in the complete motif but excluded in the minimal motif are not tested by the array. Monosaccharide identities and subsitution intolerance may or may not have been tested by the array, they are retained from the complete motif for readability.

### **Boxplot of Primary Motifs**



### List of Fine-Specificity MAL-I Motifs

Motif I	Nearest Common Name D (Accuracy%**)	Motif Graphic Structure	Relative Binding	Number of Glycans	P-Value***
A0		$ \overset{\alpha 3}{\overset{\beta 4}{\overset{\beta 2}{\overset{\alpha 6}{\overset{\beta 6}{\overset{1}{\overset{\beta 6}{\overset{\beta 6}{\overset{1}{\overset{1}{\overset{1}{\overset{1}{\overset{1}{\overset{1}{\overset{1}{$	1.00	6	-0.000
B0		$R \xrightarrow{4} \beta^{2} \alpha^{\alpha} \beta^{\beta} R$	0.22	6	0.033
0	Non-Binders (100%)	)	0.00	60	NA
Key:					
Hexose 🔘	HexNAc 🗌 Kdn 🔶 🗚	ny Single			
Glc 🔵	GlcNAc Neu5Ac	Any # of			
Man 🔵	GalNAc 📃 Neu5Gc 🧄 Mon	osaccharide ~SKIP~			
Gal 🔵	Fuc 🔺 Sia 📥 Mon	osaccharide R-			

See Symbol Nomenclature for Glycans (SNFG) for complete key: https://www.ncbi.nlm.nih.gov/glycans/snfg.html (https://www.ncbi.nlm.nih.gov/glycans/snfg.html)

\*Motif indicates the remaining glycans not matched by motifs which are a subset. Motif definition needs to be taken in the context of the model.

\*\*Accuracy describes the consistency between common-name definition of the motif and the formal, text-based definition of the motif, in terms of percent agreement in the glycans containing the two motifs. Common Name label definitions given here (http://carbogrove.org/MotifLabels.php).

\*\*\*P-Value refers to difference from Non-Binders with multiple testing correction (Dunnet's Test)

Motifs with a red motif ID fail to show a logistic response to protein concentration in the range of concentrations analyzed. These motifs may be nonbinding motifs (motifs which define nonbinding exceptions) or simply fail to fit. Nonbinding motifs are determined based on concentration dependent response when available or the average binding of non-motif glycans otherwise.

**Boxplot of Fine-Specificity Motifs** 





#### Motif Intensity Map





Figure 2. Glycan intensity and motif distribution plot. The top half of the plot presents the observed glycan binding intensity of various glycans used in the array over their rank binding intensity; only the top glycans are shown. The second plot indicates the position of glycans containing the various motifs in the top plot with a yellow tick.

#### Motif Family Membership Map



\*Motif indicates the remaining glycans not matched by motifs which are a subset. Motif definition needs to be taken in the context of the model.

Figure 3. Treemap of glycan binding grouped by motif and family structure. The model structure can be represented as nested boxes where box size is proportional to the number of glycans with the motif and color changes with change in average relative binding of glycans with the motif. Only three layers of data splitting are included here, though further splitting may be possible.

#### Detailed Model Breakdown

Motif Glycan Examples:





Figure 4. Boxplots of glycan binding grouped by motif for each dataset in the model. Motifs are listed in ascending average binding intensity (for the selected concentration) and colored by family.

Model Structure:



\*Motif matches the remaining glycans not matched by earlier motifs in the model.

Figure 5. Tree representation of the regression tree model trained on array data. Data flows through the tree (topdown) and is split by the various motifs. The motif used the split the data at each point has the id "family+split number" except when further split. In the case of futher splits the id of the motif used to split the data is denoted with an asterisk.

#### Curve Fitting:

No curves were fit for model

Motif Text Structures:

Motif ID	Motif Graphic	Motif Text
A0	$ \overset{\alpha 3}{\overset{\beta 4}{\overset{\beta 2}{\overset{\beta 2}{\overset{\alpha 6}{\overset{\beta 6}$	<6f8f>Neu5AcA2-3<2f4f6f>GalB1-4<3f6f>GlcNAcB1-2<3f4f6f>ManA1- 6<2f4f>ManB
В0	$R^{-\frac{4}{2}} \overset{\beta_2}{\longrightarrow} \overset{\alpha_6}{\longrightarrow} R}$	<2f4f6f>GalB1-4<3f6f>GlcNAcB1-2<3f4f6f>ManA1-3(<3f6f>GlcNAcB1- 2<3f4f6f>ManA1-6)<2f4f>ManB
0		Non-Binders
Key:		

11/14/2020



MAL-I

## SNA

# Automated Analysis report for SNA evaluated at 20 ug/mL

#### List of Primary SNA Motifs

Primary Motif ID	Primary Motif Minimal Structure	Primary Motif Complete Structure	Relative Binding	Number of Glycans
A	$\mathbf{R} = \mathbf{R} + $	$\mathbf{R}^{\underline{2}} \underbrace{\mathbf{A}}_{\underline{\beta}} \underbrace{\mathbf{A}}$	1.00	12
В		$ \begin{array}{c} & \alpha 6 \\ & \beta 4 \\ & \beta 2 \\ & R \\ & 3 \\ & R \end{array} \begin{array}{c} & \alpha \\ & \beta 4 \\ & R \end{array} $	0.26	10
0			0.00	50

Minimal and complete motif definitions match the same set of glycans defined in the array. Components found in the complete motif but excluded in the minimal motif are not tested by the array. Monosaccharide identities and subsitution intolerance may or may not have been tested by the array, they are retained from the complete motif for readability.

### **Boxplot of Primary Motifs**

10/16/2020



### List of Fine-Specificity SNA Motifs

Motif ID	Nearest Common Name (Accuracy%**)	Motif Graphic Structure	Relative Binding	Number of Glycans	P-Value***
A2	a6 Sialyl Type 2 LacNAc Terminal beta-Galactose (90%)		1.00	3	<0.001
A0*	a6 Sialyl Type 2 LacNAc Terminal beta-Galactose (90%)	$\mathbf{R}$	0.76	3	<0.001
A1	a6 Sialyl Type 2 LacNAc a3 Sialyl Galactose (92%)	$ \begin{array}{c}                                     $	0.50	6	<0.001
В0			0.18	10	0.001
0	Non-Binders (100%)		0.00	50	NA

SNA



Key:



See Symbol Nomenclature for Glycans (SNFG) for complete key: https://www.ncbi.nlm.nih.gov/glycans/snfg.html (https://www.ncbi.nlm.nih.gov/glycans/snfg.html)

\*Motif indicates the remaining glycans not matched by motifs which are a subset. Motif definition needs to be taken in the context of the model.

\*\*Accuracy describes the consistency between common-name definition of the motif and the formal, text-based definition of the motif, in terms of percent agreement in the glycans containing the two motifs. Common Name label definitions given here (http://carbogrove.org/MotifLabels.php).

\*\*\*P-Value refers to difference from Non-Binders with multiple testing correction (Dunnet's Test)

Motifs with a red motif ID fail to show a logistic response to protein concentration in the range of concentrations analyzed. These motifs may be nonbinding motifs (motifs which define nonbinding exceptions) or simply fail to fit. Nonbinding motifs are determined based on concentration dependent response when available or the average binding of non-motif glycans otherwise.



#### **Boxplot of Fine-Specificity Motifs**

Figure 1. Glycan binding grouped by motif and motif family. Individual glycans are given as points on the plot.
#### Motif Intensity Map



Figure 2. Glycan intensity and motif distribution plot. The top half of the plot presents the observed glycan binding intensity of various glycans used in the array over their rank binding intensity; only the top glycans are shown. The second plot indicates the position of glycans containing the various motifs in the top plot with a yellow tick.

#### Motif Family Membership Map





\*Motif indicates the remaining glycans not matched by motifs which are a subset. Motif definition needs to be taken in the context of the model.

Figure 3. Treemap of glycan binding grouped by motif and family structure. The model structure can be represented as nested boxes where box size is proportional to the number of glycans with the motif and color changes with change in average relative binding of glycans with the motif. Only three layers of data splitting are included here, though further splitting may be possible.

#### **Detailed Model Breakdown**

Motif Glycan Examples:



SNA



Figure 4. Boxplots of glycan binding grouped by motif for each dataset in the model. Motifs are listed in ascending average binding intensity (for the selected concentration) and colored by family.

file:///Z:/Shared Files/Glycan Arrays/Bisecting vs. Non-Bisecting Array/MF running for Li\_2020.10/SNA\_GeneratedReport\_1/SNA.html



#### Model Structure:



\*Motif matches the remaining glycans not matched by earlier motifs in the model.

Figure 5. Tree representation of the regression tree model trained on array data. Data flows through the tree (topdown) and is split by the various motifs. The motif used the split the data at each point has the id "family+split number" except when further split. In the case of further splits the id of the motif used to split the data is denoted with an asterisk.

#### Curve Fitting:

No curves were fit for model

Motif Text Structures:

#### Motif

ID	Motif Graphic	Motif Text
A1		<6f8f>Neu5AcA2-<3or6><2f4f>GalB1-4<6f>GlcNAcB1-2<3f4f6f>ManA1- 6(<6f8f>Neu5AcA2-6<2f3f4f>GalB1-4<3f6f>GlcNAcB1-2<3f4f6f>ManA1-3)<2f>ManB
A2		<6f8f>Neu5AcA2-6<2f3f4f>GalB1-4<3f6f>GlcNAcB1-2<3f4f6f>ManA1-3<2f4f>ManB
A0*	$ \begin{array}{c} \mathbf{R} \\ \mathbf{R}^{}}} \mathbf{B}^{}} \mathbf{B}^{}} \mathbf{B}^{}} \mathbf{R} \\ \mathbf{R}^{}} \mathbf{B}^{} \mathbf{B}^{\phantom}} \mathbf{B}^{} \mathbf{B}^{} \mathbf{B}^{} \mathbf{B}^{} \mathbf{B}^{} \mathbf{B}^{} \mathbf{B}^{} \mathbf{B}^{} \mathbf{B}^{\phantom} \mathbf{B}^{\phantom\phantom}} \mathbf{B}^{} \mathbf{B}^{\phantom\phantom} \mathbf{B}^{\phantom\phantom} \mathbf{B}^{\phantom\phantom} \mathbf{B}^{\phantom\phantom} \mathbf{B}^{\phantom\phantom} \mathbf{B}^{\phantom\phantom} \mathbf{B}^{\phantom\phantom}} \mathbf{B}^{\phantom\phantom} B$	<6f8f>Neu5AcA2-6<2f3f4f>GalB1-4<3f6f>GlcNAcB1-2<3f4f6f>ManA1-3<2f>ManB
B0		<6f8f>Neu5AcA2-6<2f3f4f>GalB1-4<3f6f>GlcNAcB1-2<3f4f6f>ManA1-6<2f>ManB

Gal 🔵

Fuc 🔺

Sia 🔶



R-

SNA

## AAL

# Automated Analysis report for AAL evaluated at 10 ug/mL

#### List of Primary AAL Motifs

Primary Motif ID	Primary Motif Minimal Structure	Primary Motif Complete Structure	Relative Binding	Number of Glycans
A	$\mathbf{R} \xrightarrow{4} \mathbf{\beta} \mathbf{R}$	$\mathbf{R}^{\mathbf{A}}_{\mathbf{A}}^{\mathbf{A}}_{\mathbf{A}}^{\mathbf{B}2} \mathbf{A}^{\mathbf{B}4}_{\mathbf{A}}^{\mathbf{B}4} \mathbf{A}^{\mathbf{B}4}$	1	40
0			0	32

Minimal and complete motif definitions match the same set of glycans defined in the array. Components found in the complete motif but excluded in the minimal motif are not tested by the array. Monosaccharide identities and subsitution intolerance may or may not have been tested by the array, they are retained from the complete motif for readability.

#### **Boxplot of Primary Motifs**

AAL



### List of Fine-Specificity AAL Motifs

Motif ID	Nearest Common Name (Accuracy%**)	Motif Graphic Structure	Relative Binding	Number of Glycans	P-Value***
A2*		$\mathbf{R} \xrightarrow{4} \mathbf{R} \xrightarrow{\beta 2} \mathbf{R} \xrightarrow{\alpha} \mathbf{R} \xrightarrow{6} \mathbf{R} \xrightarrow{6} \mathbf{R}$	1.00	16	<0.001
A1		$\mathbf{R} \xrightarrow{4} \mathbf{B}^{2} \mathbf{R} \xrightarrow{4} \mathbf{B}^{6} \mathbf{B}^{\beta} \mathbf{R}$	0.81	8	<0.001
A0*		$\mathbf{R} \xrightarrow{4} \beta \mathbf{R}$	0.80	16	<0.001
0	Non-Binders (100%)		0.00	32	NA

Key:



See Symbol Nomenclature for Glycans (SNFG) for complete key: https://www.ncbi.nlm.nih.gov/glycans/snfg.html (https://www.ncbi.nlm.nih.gov/glycans/snfg.html)

\*Motif indicates the remaining glycans not matched by motifs which are a subset. Motif definition needs to be taken in the context of the model.

\*\*Accuracy describes the consistency between common-name definition of the motif and the formal, text-based definition of the motif, in terms of percent agreement in the glycans containing the two motifs. Common Name label definitions given here (http://carbogrove.org/MotifLabels.php).

\*\*\*P-Value refers to difference from Non-Binders with multiple testing correction (Dunnet's Test)

Motifs with a red motif ID fail to show a logistic response to protein concentration in the range of concentrations analyzed. These motifs may be nonbinding motifs (motifs which define nonbinding exceptions) or simply fail to fit. Nonbinding motifs are determined based on concentration dependent response when available or the average binding of non-motif glycans otherwise.



#### **Boxplot of Fine-Specificity Motifs**

Figure 1. Glycan binding grouped by motif and motif family. Individual glycans are given as points on the plot.

#### Motif Intensity Map

AAL



Figure 2. Glycan intensity and motif distribution plot. The top half of the plot presents the observed glycan binding intensity of various glycans used in the array over their rank binding intensity; only the top glycans are shown. The second plot indicates the position of glycans containing the various motifs in the top plot with a yellow tick.

#### Motif Family Membership Map



\*Motif indicates the remaining glycans not matched by motifs which are a subset. Motif definition needs to be taken in the context of the model.

Figure 3. Treemap of glycan binding grouped by motif and family structure. The model structure can be represented as nested boxes where box size is proportional to the number of glycans with the motif and color changes with change in average relative binding of glycans with the motif. Only three layers of data splitting are included here, though further splitting may be possible.

#### **Detailed Model Breakdown**

Motif Glycan Examples:



0

ò

AAL



Motif ID

Ł

Å

A2

AAL

Figure 4. Boxplots of glycan binding grouped by motif for each dataset in the model. Motifs are listed in ascending average binding intensity (for the selected concentration) and colored by family.

Model Structure:



\*Motif matches the remaining glycans not matched by earlier motifs in the model.

Figure 5. Tree representation of the regression tree model trained on array data. Data flows through the tree (topdown) and is split by the various motifs. The motif used the split the data at each point has the id "family+split number" except when further split. In the case of futher splits the id of the motif used to split the data is denoted with an asterisk.

Curve Fitting:

No curves were fit for model

Motif Text Structures:

Motif ID	Motif Graphic	Motif Text
A1	$\mathbf{R} \stackrel{d}{\longrightarrow} \begin{bmatrix} \frac{\beta^2}{\alpha} \\ \frac{\beta^2}{\alpha} \end{bmatrix} \stackrel{d}{\longrightarrow} \mathbf{R}$	<6f8f>Neu5AcA2-<3or6><2f4f>GalB1-4<3f6f>GlcNAcB1-2<3f4f6f>ManA1- 3(<2f3f4f>FucA1-3<6f>GlcNAcB1-2<3f4f6f>ManA1-6)<2f>ManB
A2*	$R \xrightarrow{4} \beta^{2} \bigcirc \alpha$	<2f3f4f>FucA1-3<6f>GlcNAcB1-2<3f4f6f>ManA1-6<2f>ManB



AAL

## LTA (LTL)

# Automated Analysis report for LTA evaluated at 10 ug/mL

#### List of Primary LTA Motifs

Primary Motif ID	Primary Motif Minimal Structure	Primary Motif Complete Structure	Relative Binding	Number of Glycans
A	$\mathbf{R}$	$\mathbf{R}$	1	22
0			0	50

Minimal and complete motif definitions match the same set of glycans defined in the array. Components found in the complete motif but excluded in the minimal motif are not tested by the array. Monosaccharide identities and subsitution intolerance may or may not have been tested by the array, they are retained from the complete motif for readability.

#### **Boxplot of Primary Motifs**



### List of Fine-Specificity LTA Motifs

	Nearest Common				
Motif II	Name D (Accuracy%**)	Motif Graphic Structure	Relative Binding	Number of Glycans	P-Value***
A1		$ \begin{array}{c} & \beta 4 \\ & \beta 2 \\ & \alpha \\ & \beta \\ & \alpha \\ & \beta \\ $	1.00	6	<0.001
A0*	Terminal Lewis X (92%)	$\mathbf{R}$	0.68	16	<0.001
0	Non-Binders (100%)		0.00	50	NA
Key:					
Hexose () Gic () Man () Gal ()	HexNAc ☐ Kdn ◆ Any Si GlcNAc ■ Neu5Ac ◆ Any # GalNAc ■ Neu5Gc ◆ Monosacc Fuc ▲ Sia ◆ Monosacc	ngle charide of haride ~SKIP~ en or R- charide R-			

See Symbol Nomenclature for Glycans (SNFG) for complete key: https://www.ncbi.nlm.nih.gov/glycans/snfg.html (https://www.ncbi.nlm.nih.gov/glycans/snfg.html)

LTA

\*Motif indicates the remaining glycans not matched by motifs which are a subset. Motif definition needs to be taken in the context of the model.

\*\*Accuracy describes the consistency between common-name definition of the motif and the formal, text-based definition of the motif, in terms of percent agreement in the glycans containing the two motifs. Common Name label definitions given here (http://carbogrove.org/MotifLabels.php).

\*\*\*P-Value refers to difference from Non-Binders with multiple testing correction (Dunnet's Test)

Motifs with a red motif ID fail to show a logistic response to protein concentration in the range of concentrations analyzed. These motifs may be nonbinding motifs (motifs which define nonbinding exceptions) or simply fail to fit. Nonbinding motifs are determined based on concentration dependent response when available or the average binding of non-motif glycans otherwise.

# 20000 20000 10000 0 A Motif ID

#### **Boxplot of Fine-Specificity Motifs**



#### Motif Intensity Map

LTA



Figure 2. Glycan intensity and motif distribution plot. The top half of the plot presents the observed glycan binding intensity of various glycans used in the array over their rank binding intensity; only the top glycans are shown. The second plot indicates the position of glycans containing the various motifs in the top plot with a yellow tick.

#### Motif Family Membership Map



\*Motif indicates the remaining glycans not matched by motifs which are a subset. Motif definition needs to be taken in the context of the model.

Figure 3. Treemap of glycan binding grouped by motif and family structure. The model structure can be represented as nested boxes where box size is proportional to the number of glycans with the motif and color changes with change in average relative binding of glycans with the motif. Only three layers of data splitting are included here, though further splitting may be possible.

#### **Detailed Model Breakdown**

Motif Glycan Examples:



10/16/2020

LTA



Figure 4. Boxplots of glycan binding grouped by motif for each dataset in the model. Motifs are listed in ascending average binding intensity (for the selected concentration) and colored by family.

Model Structure:



LTA

\*Motif matches the remaining glycans not matched by earlier motifs in the model.

Figure 5. Tree representation of the regression tree model trained on array data. Data flows through the tree (topdown) and is split by the various motifs. The motif used the split the data at each point has the id "family+split number" except when further split. In the case of futher splits the id of the motif used to split the data is denoted with an asterisk.

Curve Fitting:

No curves were fit for model

Motif Text Structures:

Motif ID	Motif Graphic	Motif Text
A1	$\mathbf{R}^{\mathbf{A}\mathbf{A}\mathbf{B}\mathbf{A}\mathbf{B}\mathbf{C}\mathbf{C}\mathbf{C}\mathbf{C}\mathbf{C}\mathbf{C}\mathbf{C}\mathbf{C}\mathbf{C}C$	<2f3f4f6f>GalB1-4(<2f3f4f>FucA1-3)<6f>GlcNAcB1-2<3f4f6f>ManA1- 6<2f4f>ManB
A0*	$ \begin{array}{c} \beta 4 \\ \beta \\ 3 \\ \alpha \\ \end{array} $	<2f3f4f6f>GalB1-4(<2f3f4f>FucA1-3)<6f>GlcNAcB
0		Non-Binders

Key:

10/16/2020



# Automated Analysis report for RCA-I evaluated at 2 ug/mL

#### List of Primary RCA-I Motifs

Primary Motif ID	Primary Motif Minimal Structure	Primary Motif Complete Structure	Relative Binding	Number of Glycans
A	$\mathbf{R} \xrightarrow{6} \mathbf{\beta}^{\beta 4} \mathbf{\Gamma}^{\beta} \mathbf{R}$	$\mathbf{R} \stackrel{6}{\longrightarrow} \stackrel{\beta 4}{\longrightarrow} \stackrel{\beta 2}{\longrightarrow} \stackrel{\alpha}{\longrightarrow} \stackrel{\beta 4}{\longrightarrow} \stackrel{\beta 2}{\longrightarrow} \stackrel{\alpha}{\longrightarrow} \stackrel{\beta 4}{\longrightarrow} $	1.00	40
В	$ \begin{array}{c}                                     $	$\mathbf{R}^{\underline{\beta}4} \underbrace{\mathbf{R}^{\underline{\beta}2}}_{\mathbf{R}} \underbrace{\mathbf{R}^{\underline{\beta}4}}_{\mathbf{R}} \underbrace{\mathbf{R}^$	0.29	8
0			0.03	24

Minimal and complete motif definitions match the same set of glycans defined in the array. Components found in the complete motif but excluded in the minimal motif are not tested by the array. Monosaccharide identities and subsitution intolerance may or may not have been tested by the array, they are retained from the complete motif for readability.

### **Boxplot of Primary Motifs**



### List of Fine-Specificity RCA-I Motifs

Motif ID	Nearest Common Name (Accuracy%**)	Motif Graphic Structure	Relative Binding	Number of Glycans	P- Value***
A1		$ \begin{array}{c}       \beta 4 \\       R \\       R \\       R   \end{array}   $	1.00	12	<0.001
A0*		$R \xrightarrow{6} \beta^{\beta} \overline{\beta}^{\beta} R$	0.87	24	<0.001
A2	a6 Sialyl Type 2 LacNAc Terminal Lewis X (94%)	$ \overset{\alpha 6}{\longrightarrow} \overset{\beta 4}{\longrightarrow} \overset{\beta 2}{\longrightarrow} \overset{\alpha}{\longrightarrow} \overset{\alpha}{\longrightarrow} \overset{\alpha}{\longrightarrow} \overset{\alpha}{\longrightarrow} \overset{\beta 2}{\longrightarrow} \overset{\alpha}{\longrightarrow} \overset$	0.28	4	0.035
B0*		$ \begin{array}{c}       \beta 4 \\       \beta 2 \\       \alpha \\       R \\     $	0.41	4	0.001

Man 🔵 GalNAc 📃 Neu5Gc

Fuc 🔺

Sia

Gal 🔵

RCA-I

M	otif ID	Nearest Common Name (Accuracy%**)	Motif Graphic Structure	Relative Binding	Number of Glycans	P- Value***
	B1	Terminal Lewis X Bisecting N- Glycan (90%)	$ \begin{array}{c}  & \beta 4 \\  & 3 \\  & \alpha \\  & \alpha \\  & \alpha \\  & \beta 4 \\  & $	-0.02	4	1.000
	0	Non-Binders (100%)		0.00	24	NA
Key:						
Hexose 🔵 Glc 🔵	HexNAc 🗌 GlcNAc 📕	Kdn  Any Single Monosaccharide Neu5Ac	$\bigcirc$			

See Symbol Nomenclature for Glycans (SNFG) for complete key: https://www.ncbi.nlm.nih.gov/glycans/snfg.html (https://www.ncbi.nlm.nih.gov/glycans/snfg.html)

\*Motif indicates the remaining glycans not matched by motifs which are a subset. Motif definition needs to be taken in the context of the model.

\*\*Accuracy describes the consistency between common-name definition of the motif and the formal, text-based definition of the motif, in terms of percent agreement in the glycans containing the two motifs. Common Name label definitions given here (http://carbogrove.org/MotifLabels.php).

\*\*\*P-Value refers to difference from Non-Binders with multiple testing correction (Dunnet's Test)

Monosaccharide ~SKIP~

R-

Hydrogen or

Monosaccharide

Motifs with a red motif ID fail to show a logistic response to protein concentration in the range of concentrations analyzed. These motifs may be nonbinding motifs (motifs which define nonbinding exceptions) or simply fail to fit. Nonbinding motifs are determined based on concentration dependent response when available or the average binding of non-motif glycans otherwise.

#### **Boxplot of Fine-Specificity Motifs**





Figure 1. Glycan binding grouped by motif and motif family. Individual glycans are given as points on the plot.

#### Motif Intensity Map



Figure 2. Glycan intensity and motif distribution plot. The top half of the plot presents the observed glycan binding intensity of various glycans used in the array over their rank binding intensity; only the top glycans are shown. The second plot indicates the position of glycans containing the various motifs in the top plot with a yellow tick.

#### Motif Family Membership Map



\*Motif indicates the remaining glycans not matched by motifs which are a subset. Motif definition needs to be taken in the context of the model.

Figure 3. Treemap of glycan binding grouped by motif and family structure. The model structure can be represented as nested boxes where box size is proportional to the number of glycans with the motif and color changes with change in average relative binding of glycans with the motif. Only three layers of data splitting are included here, though further splitting may be possible.

#### **Detailed Model Breakdown**

Motif Glycan Examples:





All Concentration Plots:

V1.0\_BisectingN\_2ug/mL



Figure 4. Boxplots of glycan binding grouped by motif for each dataset in the model. Motifs are listed in ascending average binding intensity (for the selected concentration) and colored by family.

Model Structure:

RCA-I

\*Motif matches the remaining glycans not matched by earlier motifs in the model.

Figure 5. Tree representation of the regression tree model trained on array data. Data flows through the tree (topdown) and is split by the various motifs. The motif used the split the data at each point has the id "family+split number" except when further split. In the case of further splits the id of the motif used to split the data is denoted with an asterisk.

#### Curve Fitting:

No curves were fit for model

Motif Text Structures:

ID	Motif Graphic	Motif Text
A1	$\mathbf{R}^{\mathbf{\beta}4} \mathbf{R}^{\mathbf{\beta}2} \mathbf{R}^{\mathbf{\alpha}}$	<2f3f4f6f>GalB1-4<3f6f>GlcNAcB1-2<3f4f6f>ManA1-6<2f>ManB
A2	$ \begin{array}{c} \begin{array}{c} & & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ \end{array} \right) \begin{array}{c} & & & \\ & & \\ & & \\ & & \\ & & \\ & & \\ & & \\ & & \\ & & \\ \end{array} \right) \begin{array}{c} & & \\ & & \\ & & \\ & & \\ & & \\ & & \\ & & \\ & & \\ \end{array} \right) \begin{array}{c} & & \\ &$	<6f8f>Neu5AcA2-6<2f3f4f>GalB1-4<3f6f>GlcNAcB1-2<3f4f6f>ManA1- 6(<2f3f4f>FucA1-3<6f>GlcNAcB1-2<3f4f6f>ManA1-3)<2f>ManB
A0*	$\mathbf{R} \xrightarrow{6} \mathbf{\beta}^{\beta 4} \mathbf{\beta}^{\beta} \mathbf{R}$	<2f3f4f>GalB1-4<3f6f>GlcNAcB
B1	$ \begin{array}{c}  & \beta 4 \\  & 3 \\  & 3 \\  & \alpha \\  & \alpha \\  & \beta 4 \\  & \beta 2 \\  & \beta 4 \\  & $	<2f3f4f6f>GalB1-4(<2f3f4f>FucA1-3)<6f>GlcNAcB1-2<3f4f6f>ManA1- 6(<3f4f6f>GlcNAcB1-4)<2f>ManB

#### Motif



## ECL

# Automated Analysis report for ECL evaluated at 10 ug/mL

#### List of Primary ECL Motifs

Primary Motif ID	Primary Motif Minimal Structure	Primary Motif Complete Structure	Relative Binding	Number of Glycans
A	$\mathbf{R} = \frac{\mathbf{A}_{\mathbf{A}}_{\mathbf{A}_{\mathbf{A}}_{\mathbf{A}}_{\mathbf{A}}}}}}}}}}$	$\mathbf{R}^{\underline{2}}$	1.00	12
В	$ \begin{array}{c}             \beta^{4} \\             R \\             R \\         $	$\mathbf{R}^{\underline{\beta}4} \underbrace{\mathbf{\beta}2}_{\mathbf{R}} \underbrace{\mathbf{R}^{\underline{\beta}4}}_{\mathbf{R}} \underbrace{\mathbf{\beta}4}_{\mathbf{R}} \underbrace{\mathbf{\beta}4}$	0.34	10
0			0.00	50

Minimal and complete motif definitions match the same set of glycans defined in the array. Components found in the complete motif but excluded in the minimal motif are not tested by the array. Monosaccharide identities and subsitution intolerance may or may not have been tested by the array, they are retained from the complete motif for readability.

#### **Boxplot of Primary Motifs**



ECL

### List of Fine-Specificity ECL Motifs

	Nearest Common				
	Name	Motif Graphic		Number of	
Motif ID	(Accuracy%**)	Structure	Relative Binding	Glycans	P-Value***
A0		$\mathbf{R}$	1.00	12	<0.001
B0*		$ \begin{array}{c}                                     $	0.42	7	<0.001
B1	Terminal Type 2 LacNAc a3 Sialyl Galactose (90%)	$ \begin{array}{c}  & \beta 4 \\  & \beta 2 \\  & \beta 4 \\  & \beta 4 \\  & \beta 4 \\  & \beta 4 \\  & \beta 2 \\  & \beta 4 \\  & \beta 2 \\  & \beta 4 \\  & \beta 4 \\  & \beta 2 \\  & \beta 4 \\  & \beta$	0.16	3	0.579
0	Non-Binders (100%)		0.00	50	NA

Key:



See Symbol Nomenclature for Glycans (SNFG) for complete key: https://www.ncbi.nlm.nih.gov/glycans/snfg.html (https://www.ncbi.nlm.nih.gov/glycans/snfg.html)

ECL

\*Motif indicates the remaining glycans not matched by motifs which are a subset. Motif definition needs to be taken in the context of the model.

\*\*Accuracy describes the consistency between common-name definition of the motif and the formal, text-based definition of the motif, in terms of percent agreement in the glycans containing the two motifs. Common Name label definitions given here (http://carbogrove.org/MotifLabels.php).

\*\*\*P-Value refers to difference from Non-Binders with multiple testing correction (Dunnet's Test)

Motifs with a red motif ID fail to show a logistic response to protein concentration in the range of concentrations analyzed. These motifs may be nonbinding motifs (motifs which define nonbinding exceptions) or simply fail to fit. Nonbinding motifs are determined based on concentration dependent response when available or the average binding of non-motif glycans otherwise.



#### **Boxplot of Fine-Specificity Motifs**

Figure 1. Glycan binding grouped by motif and motif family. Individual glycans are given as points on the plot.

#### Motif Intensity Map



ECL



Figure 2. Glycan intensity and motif distribution plot. The top half of the plot presents the observed glycan binding intensity of various glycans used in the array over their rank binding intensity; only the top glycans are shown. The second plot indicates the position of glycans containing the various motifs in the top plot with a yellow tick.

#### Motif Family Membership Map

ECL

B 1	A 	
0*		Relative Binding 1.0 0.8 0.6 0.4 0.2
# Family Sub-Family #	# Member *Family Remainder	

\*Motif indicates the remaining glycans not matched by motifs which are a subset. Motif definition needs to be taken in the context of the model.

Figure 3. Treemap of glycan binding grouped by motif and family structure. The model structure can be represented as nested boxes where box size is proportional to the number of glycans with the motif and color changes with change in average relative binding of glycans with the motif. Only three layers of data splitting are included here, though further splitting may be possible.

#### Detailed Model Breakdown

Motif Glycan Examples:


ECL



Motif ID

ECL

Figure 4. Boxplots of glycan binding grouped by motif for each dataset in the model. Motifs are listed in ascending average binding intensity (for the selected concentration) and colored by family.

Model Structure:



\*Motif matches the remaining glycans not matched by earlier motifs in the model.

Figure 5. Tree representation of the regression tree model trained on array data. Data flows through the tree (topdown) and is split by the various motifs. The motif used the split the data at each point has the id "family+split number" except when further split. In the case of futher splits the id of the motif used to split the data is denoted with an asterisk.

Curve Fitting:

No curves were fit for model

Motif Text Structures:

Motif ID	Motif Graphic	Motif Text
A0	$R = \frac{4}{3} \frac{\beta^2}{\alpha} R$	<2f3f4f6f>GalB1-4<3f6f>GlcNAcB1-2<3f4f6f>ManA1-3<2f>ManB
B1		<6f8f>Neu5AcA2-<3or6><2f4f>GalB1-4<6f>GlcNAcB1-2<3f4f6f>ManA1- 3(<2f3f4f6f>GalB1-4<3f6f>GlcNAcB1-2<3f4f6f>ManA1-6)(<3f4f6f>GlcNAcB1-4) <2f>ManB

10/16/2020

Motif ID	Motif Graphic	Motif Text	
B0*	β <sup>4</sup> <sup>β2</sup> <sup>α</sup> <sup>6</sup> <sup>6</sup> <sup>6</sup> <sup>6</sup> <sup>8</sup> <sup>8</sup> <sup>8</sup> <sup>8</sup> <sup>8</sup> <sup>8</sup> <sup>8</sup> <sup>8</sup>	<2f3f4f6f>GalB1-4<3f6f>GlcNAcB1-2<3f4f6f>ManA1-6<2f>ManB	
0		Non-Binders	
Key:			
Hexose 🔿	HexNAc 🗌	Kdn Any Single	
Glc 🔵	GlcNAc 📃 N	J5Ac Any # of	
Man 🔵	GalNAc 📃 N	J5Gc Annosaccharide	
Gal 🔵	Fuc 🔺	Sia  Monosaccharide R-	

ECL

## WFA (WFL)

## Automated Analysis report for WFA evaluated at 10 ug/mL

### List of Primary WFA Motifs

Primary Motif ID	Primary Motif Minimal Structure	Primary Motif Complete Structure	Relative Binding	Number of Glycans
A	$\mathbf{R} = \mathbf{R} + $	$\mathbf{R}^{4}_{\mathbf{R}} \xrightarrow{\beta 2} \mathbf{\alpha}^{\alpha}_{\mathbf{R}} \xrightarrow{4 \ \beta 4} \mathbf{\beta 4}$	1.00	12
0			0.04	60

Minimal and complete motif definitions match the same set of glycans defined in the array. Components found in the complete motif but excluded in the minimal motif are not tested by the array. Monosaccharide identities and subsitution intolerance may or may not have been tested by the array, they are retained from the complete motif for readability.

### **Boxplot of Primary Motifs**

WFA



### List of Fine-Specificity WFA Motifs

	Nearest Common				
Motif ID	Name (Accuracy%**)	Motif Graphic Structure	Relative Binding	Number of Glycans	P-Value***
A0*		$R = \frac{4}{3} \frac{\beta}{\beta} R$	1.00	6	<0.001
A2	Terminal Type 2 LacNAc Terminal Lewis X (96%)	$\mathbf{R} \xrightarrow{4} \overset{\beta 2}{\overset{\alpha}{\overset{\alpha}{\overset{\alpha}{\overset{\beta}{\overset{\beta}{\overset{\beta}{\overset{\beta}{\beta$	0.48	3	0.001
A1	Terminal Type 2 LacNAc a3 Sialyl Galactose (90%)	$ \begin{array}{c}                                     $	0.11	3	0.736
0	Non-Binders (100%)		0.00	60	NA

Key:



See Symbol Nomenclature for Glycans (SNFG) for complete key: https://www.ncbi.nlm.nih.gov/glycans/snfg.html (https://www.ncbi.nlm.nih.gov/glycans/snfg.html)

WFA

\*Motif indicates the remaining glycans not matched by motifs which are a subset. Motif definition needs to be taken in the context of the model.

\*\*Accuracy describes the consistency between common-name definition of the motif and the formal, text-based definition of the motif, in terms of percent agreement in the glycans containing the two motifs. Common Name label definitions given here (http://carbogrove.org/MotifLabels.php).

\*\*\*P-Value refers to difference from Non-Binders with multiple testing correction (Dunnet's Test)

Motifs with a red motif ID fail to show a logistic response to protein concentration in the range of concentrations analyzed. These motifs may be nonbinding motifs (motifs which define nonbinding exceptions) or simply fail to fit. Nonbinding motifs are determined based on concentration dependent response when available or the average binding of non-motif glycans otherwise.



### **Boxplot of Fine-Specificity Motifs**

Figure 1. Glycan binding grouped by motif and motif family. Individual glycans are given as points on the plot.

### Motif Intensity Map



Figure 2. Glycan intensity and motif distribution plot. The top half of the plot presents the observed glycan binding intensity of various glycans used in the array over their rank binding intensity; only the top glycans are shown. The second plot indicates the position of glycans containing the various motifs in the top plot with a yellow tick.

### Motif Family Membership Map



\*Motif indicates the remaining glycans not matched by motifs which are a subset. Motif definition needs to be taken in the context of the model.

Figure 3. Treemap of glycan binding grouped by motif and family structure. The model structure can be represented as nested boxes where box size is proportional to the number of glycans with the motif and color changes with change in average relative binding of glycans with the motif. Only three layers of data splitting are included here, though further splitting may be possible.

### Detailed Model Breakdown

Motif Glycan Examples:



WFA





Figure 4. Boxplots of glycan binding grouped by motif for each dataset in the model. Motifs are listed in ascending average binding intensity (for the selected concentration) and colored by family.

Model Structure:



WFA

\*Motif matches the remaining glycans not matched by earlier motifs in the model.

Figure 5. Tree representation of the regression tree model trained on array data. Data flows through the tree (topdown) and is split by the various motifs. The motif used the split the data at each point has the id "family+split number" except when further split. In the case of further splits the id of the motif used to split the data is denoted with an asterisk.

#### Curve Fitting:

No curves were fit for model

Motif Text Structures:

Motif ID	Motif Graphic	Motif Text
A1		<6f8f>Neu5AcA2-<3or6><2f4f>GalB1-4<6f>GlcNAcB1-2<3f4f6f>ManA1- 6(<2f3f4f6f>GalB1-4<3f6f>GlcNAcB1-2<3f4f6f>ManA1-3)(<3f4f6f>GlcNAcB1-4) <2f>ManB
A2	$\mathbf{R} \xrightarrow{4} \mathbf{\beta}^{2} \mathbf{Q}_{\alpha}$ $\mathbf{R} \xrightarrow{3} \mathbf{\beta}^{\beta} \mathbf{R}$ $\mathbf{R} \xrightarrow{3} \mathbf{\beta}^{\beta} \mathbf{R}$	<2f3f4f6f>GalB1-4<3f6f>GlcNAcB1-2<3f4f6f>ManA1-3(<2f3f4f>FucA1-3<6f>GlcNAcB1- 2<3f4f6f>ManA1-6)<2f>ManB
A0*	$R = \frac{\beta^{\frac{\beta}{4}}}{3} R$	<2f3f4f6f>GalB1-4<3f6f>GlcNAcB1-2<3f4f6f>ManA1-3<2f>ManB
0		Non-Binders

11/14/2020

#### Key:



# Automated Analysis report for WGA evaluated at 5 ug/mL

### List of Primary WGA Motifs

Primary Motif ID	Primary Motif Minimal Structure	Primary Motif Complete Structure	Relative Binding	Number of Glycans
A	$\mathbf{R}$	$\mathbf{R}$	1.00	12
В	$\mathbf{R} = \mathbf{R} + $	$ \begin{array}{c} R \xrightarrow{2} 4^{46} \beta^{4} \\ R \xrightarrow{3} \beta^{2} \\ R \xrightarrow{3} \beta^{2} \end{array} $	0.48	24
С	$\mathbf{R} \xrightarrow{4} \mathbf{\beta}^{2} \mathbf{R} \xrightarrow{4} \mathbf{\beta}^{6} \mathbf{\beta} \mathbf{R}$	$\mathbf{R} \xrightarrow{4} \overset{\beta 2}{\longrightarrow} \overset{\alpha}{\longrightarrow} \overset{\alpha}{\longrightarrow} \overset{\beta 4}{\longrightarrow} \overset{\beta 4}{\longrightarrow} \overset{\beta 2}{\longrightarrow} \overset{\alpha}{\longrightarrow} \overset{\alpha}{\longrightarrow} \overset{\beta 4}{\longrightarrow} \overset{\beta 4}{\to} \overset{\beta 4}{\longrightarrow} \overset{\beta 4}{\overset} \overset{\beta 4}{\overset} \overset{\beta 4}{\overset} \overset{\beta 4}{\overset} \overset{\beta 4}{\overset} \overset{\beta 4}{\overset} $	0.27	8
0			0.03	28

Minimal and complete motif definitions match the same set of glycans defined in the array. Components found in the complete motif but excluded in the minimal motif are not tested by the array. Monosaccharide identities and subsitution intolerance may or may not have been tested by the array, they are retained from the complete motif for readability.

### **Boxplot of Primary Motifs**



### List of Fine-Specificity WGA Motifs

Motif ID	Nearest Common Name (Accuracy%**)	Motif Graphic Structure	Relative Binding	Number of Glycans	P- Value***
A0*	Terminal N-Glycan N-Acetyl Glucosamine Bisecting N-Glycan (93%)	$\mathbf{R}$	1.00	6	<0.001
A1		$\mathbf{R} \stackrel{\boldsymbol{\beta}}{\longrightarrow} \mathbf{R}$	0.00	6	1.000
B0		$ \begin{array}{c} \mathbf{R} \\ \mathbf$	0.26	24	<0.001
C1		$\mathbf{R} \xrightarrow{4} \mathbf{\beta}^{2} \mathbf{\alpha}^{\alpha}_{6} \mathbf{\beta}^{\beta} \mathbf{R}$	0.26	4	0.079

10/16/2020

WGA

Mot	if ID	Nearest Common Name (Accuracy%**)	Motif Graphic Structure	Relative Binding	Number of Glycans	P- Value***
C	0*	Terminal Type 2 LacNAc Bisecting N-Glycan (90%)	$\mathbf{R} \xrightarrow{4} \mathbf{\beta}^{2} \mathbf{\alpha}$ $\mathbf{R} \xrightarrow{4} \mathbf{\beta}^{6} \mathbf{\beta}$ $\mathbf{R} \xrightarrow{4} \mathbf{\beta}^{6} \mathbf{\beta}$ $\mathbf{R} \xrightarrow{4} \mathbf{\beta}^{2} \mathbf{\alpha}$	0.00	4	1.000
(	)	Non-Binders (100%)		0.00	28	NA
Key:	_	Ann Circle				



See Symbol Nomenclature for Glycans (SNFG) for complete key: https://www.ncbi.nlm.nih.gov/glycans/snfg.html (https://www.ncbi.nlm.nih.gov/glycans/snfg.html)

\*Motif indicates the remaining glycans not matched by motifs which are a subset. Motif definition needs to be taken in the context of the model.

\*\*Accuracy describes the consistency between common-name definition of the motif and the formal, text-based definition of the motif, in terms of percent agreement in the glycans containing the two motifs. Common Name label definitions given here (http://carbogrove.org/MotifLabels.php).

\*\*\*P-Value refers to difference from Non-Binders with multiple testing correction (Dunnet's Test)

Motifs with a red motif ID fail to show a logistic response to protein concentration in the range of concentrations analyzed. These motifs may be nonbinding motifs (motifs which define nonbinding exceptions) or simply fail to fit. Nonbinding motifs are determined based on concentration dependent response when available or the average binding of non-motif glycans otherwise.

### **Boxplot of Fine-Specificity Motifs**





### Motif Intensity Map



Figure 2. Glycan intensity and motif distribution plot. The top half of the plot presents the observed glycan binding intensity of various glycans used in the array over their rank binding intensity; only the top glycans are shown. The second plot indicates the position of glycans containing the various motifs in the top plot with a yellow tick.

### Motif Family Membership Map

1 0*	B	Relative Binding
A 1		1.00 0.75 0.50 0.25 0.00
0* # Family Sub-Family	# Member *Family Remainder	

\*Motif indicates the remaining glycans not matched by motifs which are a subset. Motif definition needs to be taken in the context of the model.

Figure 3. Treemap of glycan binding grouped by motif and family structure. The model structure can be represented as nested boxes where box size is proportional to the number of glycans with the motif and color changes with change in average relative binding of glycans with the motif. Only three layers of data splitting are included here, though further splitting may be possible.

### Detailed Model Breakdown

Motif Glycan Examples:





V1.0\_BisectingN\_5ug/mL



Figure 4. Boxplots of glycan binding grouped by motif for each dataset in the model. Motifs are listed in ascending average binding intensity (for the selected concentration) and colored by family.

Model Structure:

WGA Α N 14 В 1 24 1 L **B**0 С A1 A0\* 16 0 1 1 C0\* C1

\*Motif matches the remaining glycans not matched by earlier motifs in the model.

Figure 5. Tree representation of the regression tree model trained on array data. Data flows through the tree (topdown) and is split by the various motifs. The motif used the split the data at each point has the id "family+split number" except when further split. In the case of further splits the id of the motif used to split the data is denoted with an asterisk.

Curve Fitting:

No curves were fit for model

Motif Text Structures:

Motif ID	Motif Graphic	Motif Text
A1	$\mathbf{R} \stackrel{6}{\longrightarrow} \mathcal{B}$	<3f4f6f>GlcNAcB1-2<3f4f6f>ManA1-3<2f4f>ManB
A0*	$\mathbf{R}$	<3f4f6f>GlcNAcB1-2<3f4f6f>ManA1-3<2f>ManB
B0	$ \begin{array}{c}                                     $	<6f8f>Neu5AcA2-3<2f4f6f>GalB1-4<6f>GlcNAcB1-2<3f4f6f>ManA1-3<2f>ManB

10/16/2020

WGA

Motif ID	Motif Graphic	Motif Text
C1	$\mathbf{R}^{\underline{4}} \underbrace{\mathbf{B}^{\beta 2}}_{\mathbf{B}^{\alpha}} \underbrace{\mathbf{B}^{\alpha}}_{\mathbf{B}^{\beta}} \underbrace{\mathbf{B}^{\alpha}}_{\mathbf{B}^{\beta}} \mathbf{R}$	<2f3f4f6f>GalB1-4<3f6f>GlcNAcB1-2<3f4f6f>ManA1-3(<3f6f>GlcNAcB1- 2<3f4f6f>ManA1-6)<2f4f>ManB
C0*	$\mathbf{R} \xrightarrow{4} \mathbf{R} \xrightarrow{6} \mathbf{R}$	<2f3f4f6f>GalB1-4<3f6f>GlcNAcB1-2<3f4f6f>ManA1-3(<3f6f>GlcNAcB1- 2<3f4f6f>ManA1-6)<2f>ManB
0		Non-Binders
Key:		
Hexose Glc Man Gal	HexNAc C Kdn GlcNAc Neu5Ac GalNAc Neu5Gc Fuc Sia	Any Single Monosaccharide Any # of Monosaccharide ~SKIP~ Hydrogen or Monosaccharide R-

# Automated Analysis report for ConA evaluated at 1 ug/mL

### List of Primary ConA Motifs

Primary Motif ID	Primary Motif Minimal Structure	Primary Motif Complete Structure	Relative Binding	Number of Glycans
A	$\mathbf{R} \stackrel{4}{\longrightarrow} \stackrel{\beta 2}{\longrightarrow} \stackrel{\alpha}{\longrightarrow} \stackrel{\beta}{\longrightarrow} \mathbf{R}$	$\mathbf{R} \xrightarrow{4} \mathbf{\beta}^{2} \mathbf{\alpha}^{\alpha} \xrightarrow{6} \mathbf{\beta}^{4} \mathbf{\beta}^{4} \mathbf{\beta}^{4}$ $\mathbf{R} \xrightarrow{4} \mathbf{\beta}^{2} \mathbf{\alpha}^{\alpha} \xrightarrow{6} \mathbf{\beta}^{4} \mathbf{\beta}^{4} \mathbf{\beta}^{4}$ $\mathbf{R} \xrightarrow{6} \mathbf{\beta}^{4} \mathbf{\beta}^{2} \mathbf{\alpha}^{\alpha} \xrightarrow{6} \mathbf{\beta}^{4} \mathbf{\beta}^{4} \mathbf{\beta}^{4}$	1.00	24
В	$\mathbf{R} \xrightarrow{4} \mathbf{R} \xrightarrow{\beta 2} \mathbf{R} \xrightarrow{\alpha 6} \mathbf{R}$	$\mathbf{R} \xrightarrow{3} \mathbf{\beta}^{\beta 4} \mathbf{\beta}^{\beta 2} \mathbf{\alpha}^{\alpha}$ $\mathbf{R} \xrightarrow{3} \mathbf{\beta}^{\beta 4} \mathbf{\beta}^{\beta 4} \mathbf{\beta}^{\beta 4} \mathbf{\beta}^{\beta 4}$ $\mathbf{R} \xrightarrow{3} \mathbf{\beta}^{\beta 2} \mathbf{\alpha}^{\alpha}$ $\mathbf{R} \xrightarrow{3} \mathbf{\beta}^{\beta 2} \mathbf{\alpha}^{\alpha}$	0.30	12
0			0.01	36

Minimal and complete motif definitions match the same set of glycans defined in the array. Components found in the complete motif but excluded in the minimal motif are not tested by the array. Monosaccharide identities and subsitution intolerance may or may not have been tested by the array, they are retained from the complete motif for readability.

### **Boxplot of Primary Motifs**





### List of Fine-Specificity ConA Motifs

Motif ID	Nearest Common Name (Accuracy%**)	Motif Graphic Structure	Relative Binding	Number of Glycans	P-Value***
A0*		$\mathbf{R} \xrightarrow{4} \mathbf{R}^{\beta 2} \mathbf{R}^{\alpha 6} \mathbf{R}^{\beta \beta} \mathbf{R}$	1.00	13	<0.001
A2		$ \begin{array}{c} & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & $	0.32	8	<0.001
A1		$\mathbf{R}$	0.22	3	0.209
B0		$\mathbf{R} \xrightarrow{4} \mathbf{R} \xrightarrow{\beta 2} \mathbf{R} \xrightarrow{\alpha 6} \mathbf{R}$	0.20	12	0.010

Man 🔵

Gal 🔴

GalNAc 📃 Neu5Gc 🛆

Sia 🥢

Fuc 🔺

Hydrogen or

Monosaccharide

R-

ConA

Motif I	Nearest Comm Name D (Accuracy%*'	on <sup>(</sup> ) Motif Graphic Structure	Relative Binding	Number of Glycans	P-Value***
0	Non-Binders (10	0%)	0.00	36	NA
Key:					
Hexose () Glc () Man	HexNAc Kdn GlcNAc Neu5Ac GalNAc Neu5Gc	Any Single Monosaccharide Any # of Monosaccharide ~ <b>SKIP</b> ~			

See Symbol Nomenclature for Glycans (SNFG) for complete key: https://www.ncbi.nlm.nih.gov/glycans/snfg.html (https://www.ncbi.nlm.nih.gov/glycans/snfg.html)

\*Motif indicates the remaining glycans not matched by motifs which are a subset. Motif definition needs to be taken in the context of the model.

\*\*Accuracy describes the consistency between common-name definition of the motif and the formal, text-based definition of the motif, in terms of percent agreement in the glycans containing the two motifs. Common Name label definitions given here (http://carbogrove.org/MotifLabels.php).

\*\*\*P-Value refers to difference from Non-Binders with multiple testing correction (Dunnet's Test)

Motifs with a red motif ID fail to show a logistic response to protein concentration in the range of concentrations analyzed. These motifs may be nonbinding motifs (motifs which define nonbinding exceptions) or simply fail to fit. Nonbinding motifs are determined based on concentration dependent response when available or the average binding of non-motif glycans otherwise.

#### **Boxplot of Fine-Specificity Motifs**



Figure 1. Glycan binding grouped by motif and motif family. Individual glycans are given as points on the plot.

### Motif Intensity Map



Figure 2. Glycan intensity and motif distribution plot. The top half of the plot presents the observed glycan binding intensity of various glycans used in the array over their rank binding intensity; only the top glycans are shown. The second plot indicates the position of glycans containing the various motifs in the top plot with a yellow tick.

### Motif Family Membership Map

B	A	
	1 2	Relative Binding
		0.8 0.6
		0.4
		0.2
	0*	
# Family Sub-Family #	Member *Family Remainder	

\*Motif indicates the remaining glycans not matched by motifs which are a subset. Motif definition needs to be taken in the context of the model.

Figure 3. Treemap of glycan binding grouped by motif and family structure. The model structure can be represented as nested boxes where box size is proportional to the number of glycans with the motif and color changes with change in average relative binding of glycans with the motif. Only three layers of data splitting are included here, though further splitting may be possible.

### Detailed Model Breakdown

Motif Glycan Examples:

11/14/2020





Figure 4. Boxplots of glycan binding grouped by motif for each dataset in the model. Motifs are listed in ascending average binding intensity (for the selected concentration) and colored by family.

Model Structure:



\*Motif matches the remaining glycans not matched by earlier motifs in the model.

Figure 5. Tree representation of the regression tree model trained on array data. Data flows through the tree (topdown) and is split by the various motifs. The motif used the split the data at each point has the id "family+split number" except when further split. In the case of futher splits the id of the motif used to split the data is denoted with an asterisk.

Curve Fitting:

No curves were fit for model

Motif Text Structures:

#### 

Motif ID	Motif Grap	ohic	Motif Text
A 0*			
A0."		-R	<3161>GICINACB1-2<314161>ManA1-6<2141>ManB
B0		-R	<2f3f4f>FucA1-3<6f>GlcNAcB1-2<3f4f6f>ManA1-6<2f4f>ManB
0			Non-Binders
Key:			
Hexose 🔾	HexNAc 🗌	Kdn 🗸	Any Single
Glc 🔵	GlcNAc 📃	Neu5Ac	Any # of
Man 🔵	GalNAc 📃 I	Neu5Gc	
Gal 🔵	Fuc 🔺	Sia	Monosaccharide

# Automated Analysis report for GNL evaluated at 10 ug/mL

### List of Primary GNL Motifs

Primary Motif ID	Primary Motif Minimal Structure	Primary Motif Complete Structure	Relative Binding	Number of Glycans
A	$\mathbf{R} \stackrel{4}{\longrightarrow} \stackrel{\beta 2}{\longrightarrow} \stackrel{\alpha}{\longrightarrow} \stackrel{\beta}{\longrightarrow} \mathbf{R}$	$R \xrightarrow{4} \beta^{2} \alpha^{6} \beta^{4} \beta^{4}$ $R \xrightarrow{4} \beta^{2} \alpha^{3} \beta^{4} $	1.00	16
В	$R^{4}$	$R \xrightarrow{4}{3} \frac{\beta^2}{\alpha} \frac{\alpha}{3} \frac{6}{\beta^4} \frac{\beta^4}{\beta^4}$	0.54	4
0			0.01	52

Minimal and complete motif definitions match the same set of glycans defined in the array. Components found in the complete motif but excluded in the minimal motif are not tested by the array. Monosaccharide identities and subsitution intolerance may or may not have been tested by the array, they are retained from the complete motif for readability.

### **Boxplot of Primary Motifs**

11/14/2020

GNL



### List of Fine-Specificity GNL Motifs

	Nearest Common				
	Name	Motif Graphic		Number of	
Motif ID	(Accuracy%**)	Structure	Relative Binding	Glycans	P-Value***
A0*		$\mathbf{R}^{4} \mathbf{\beta}^{2} \mathbf{\beta}^{\alpha} \mathbf{\beta}^{\beta} \mathbf{R}$ $\mathbf{R}^{4} \mathbf{\beta}^{\beta} \mathbf{\beta}^{\alpha} \mathbf{\beta}^{\alpha} \mathbf{\beta}^{\beta} \mathbf{R}$	1.00	8	<0.001
A1		$\mathbf{R}^{-4} \mathbf{B}^{\beta 2} \mathbf{O}^{\alpha}_{\beta} \mathbf{B}^{\beta} \mathbf{R}$	0.02	8	0.999
В0		$R^{-\frac{\beta}{4}}$	0.30	4	0.499
0	Non-Binders (100%)		0.00	52	NA

Key:



See Symbol Nomenclature for Glycans (SNFG) for complete key: https://www.ncbi.nlm.nih.gov/glycans/snfg.html (https://www.ncbi.nlm.nih.gov/glycans/snfg.html)

GNL

\*Motif indicates the remaining glycans not matched by motifs which are a subset. Motif definition needs to be taken in the context of the model.

\*\*Accuracy describes the consistency between common-name definition of the motif and the formal, text-based definition of the motif, in terms of percent agreement in the glycans containing the two motifs. Common Name label definitions given here (http://carbogrove.org/MotifLabels.php).

\*\*\*P-Value refers to difference from Non-Binders with multiple testing correction (Dunnet's Test)

Motifs with a red motif ID fail to show a logistic response to protein concentration in the range of concentrations analyzed. These motifs may be nonbinding motifs (motifs which define nonbinding exceptions) or simply fail to fit. Nonbinding motifs are determined based on concentration dependent response when available or the average binding of non-motif glycans otherwise.



### **Boxplot of Fine-Specificity Motifs**

Figure 1. Glycan binding grouped by motif and motif family. Individual glycans are given as points on the plot.

### Motif Intensity Map





Figure 2. Glycan intensity and motif distribution plot. The top half of the plot presents the observed glycan binding intensity of various glycans used in the array over their rank binding intensity; only the top glycans are shown. The second plot indicates the position of glycans containing the various motifs in the top plot with a yellow tick.

### Motif Family Membership Map



\*Motif indicates the remaining glycans not matched by motifs which are a subset. Motif definition needs to be taken in the context of the model.

Figure 3. Treemap of glycan binding grouped by motif and family structure. The model structure can be represented as nested boxes where box size is proportional to the number of glycans with the motif and color changes with change in average relative binding of glycans with the motif. Only three layers of data splitting are included here, though further splitting may be possible.

### **Detailed Model Breakdown**

Motif Glycan Examples:





Motif ID

Figure 4. Boxplots of glycan binding grouped by motif for each dataset in the model. Motifs are listed in ascending average binding intensity (for the selected concentration) and colored by family.

Model Structure:
GNL



\*Motif matches the remaining glycans not matched by earlier motifs in the model.

Figure 5. Tree representation of the regression tree model trained on array data. Data flows through the tree (topdown) and is split by the various motifs. The motif used the split the data at each point has the id "family+split number" except when further split. In the case of futher splits the id of the motif used to split the data is denoted with an asterisk.

#### Curve Fitting:

No curves were fit for model

Motif Text Structures:

ID	Motif Graphic	Motif Text
A1	$\mathbf{R}^{\underline{4}} \mathbf{P}^{\underline{\beta} \underline{2}} \mathbf{P}^{\underline{\alpha} \underline{6}} \mathbf{P}^{\underline{\beta} \underline{\beta}} \mathbf{P}^{\underline{\alpha} \underline{6}} \mathbf{P}^{\underline{\beta} \underline{\beta}} \mathbf{P}^{\underline{\alpha} \underline{\beta} \underline{\beta}} \mathbf{P}^{\underline{\beta} \underline{\beta} \underline{\beta} \underline{\beta} \underline{\beta}} \mathbf{P}^{\underline{\beta} \underline{\beta} \underline{\beta} \underline{\beta} \underline{\beta} \underline{\beta} \underline{\beta} \underline{\beta}$	<6f8f>Neu5AcA2-<3or6><2f4f>GalB1-4<3f6f>GlcNAcB1-2<3f4f6f>ManA1- 3(<3f6f>GlcNAcB1-2<3f4f6f>ManA1-6)<2f4f>ManB
A0*	$R \xrightarrow{4} \beta^{2} \varphi^{\alpha} = \beta^{\alpha} \beta^{\beta} R$	<3f6f>GlcNAcB1-2<3f4f6f>ManA1-3(<3f6f>GlcNAcB1-2<3f4f6f>ManA1-6) <2f4f>ManB
В0	$R^{\frac{\beta}{4}}$	<2f3f4f6f>GalB1-4(<2f3f4f>FucA1-3)<6f>GlcNAcB1-2<3f4f6f>ManA1- 6(<3f6f>GlcNAcB1-2<3f4f6f>ManA1-3)<2f4f>ManB
0		Non-Binders

#### Motif

11/14/2020

#### Key:



GNL

## GSII

# Automated Analysis report for GSII evaluated at 10 ug/mL

### List of Primary GSII Motifs

Primary Motif ID	Primary Motif Minimal Structure	Primary Motif Complete Structure	Relative Binding	Number of Glycans
A	$\beta^{\beta}$ $\beta^{\alpha}$ $\beta^{\beta}$ $\beta^{\beta}$ $R$	$\mathbf{R}_{\mathbf{A}}^{\mathbf{A}} \mathbf{\beta}^{\mathbf{\beta}} \mathbf{\alpha}^{\mathbf{\beta}} \mathbf{\beta}^{\mathbf{\alpha}} \mathbf{\beta}^{\mathbf{\beta}} \mathbf{\beta}^{\beta$	1.00	6
В	$\mathbf{R}^{\underline{\beta} 2} \mathbf{R}^{\alpha}$	$\mathbf{R} \xrightarrow{4}{3} \mathbf{\beta}^{4} \mathbf{\beta}^{4}$ $\mathbf{R} \xrightarrow{4}{3} \mathbf{\beta}^{2} \mathbf{\beta}^{\alpha}$ $\mathbf{R} \xrightarrow{4}{3} \mathbf{\beta}^{2} \mathbf{\beta}^{\alpha}$	0.33	11
0			0.01	55

Minimal and complete motif definitions match the same set of glycans defined in the array. Components found in the complete motif but excluded in the minimal motif are not tested by the array. Monosaccharide identities and subsitution intolerance may or may not have been tested by the array, they are retained from the complete motif for readability.

## **Boxplot of Primary Motifs**



### List of Fine-Specificity GSII Motifs

Madif I	Nearest Common Name	Matif Quanhia Structure	Deletive Diredies	Number of	D \/al
	J (Accuracy%"")	Motif Graphic Structure	Relative Binding	Glycans	P-value"""
A0		$\beta^{\beta^2}$	1.00	6	<0.001
В0		$\mathbf{R}^{\mathbf{\beta} 2} \mathbf{R}^{\mathbf{\alpha}} \mathbf{R}^{\mathbf{\beta} \mathbf{\beta}} \mathbf{R}$	0.32	11	0.002
0	Non-Binders (100%	)	0.00	55	NA
Key:					
Hexose () Glc () Man () Gal ()	HexNAc Kdn () GicNAc Neu5Ac () GalNAc Neu5Gc () Fuc () Sia () Kdn () Morection Ho	Any Single nosaccharide Any # of nosaccharide ~SKIP~ lydrogen or nosaccharide R-			

See Symbol Nomenclature for Glycans (SNFG) for complete key: https://www.ncbi.nlm.nih.gov/glycans/snfg.html (https://www.ncbi.nlm.nih.gov/glycans/snfg.html)

file:///C:/Users/Jian/Desktop/Lei Li Bisecting N-glycan Manuscript/GSII\_GeneratedReport\_1/GSII.html

GSII

\*Motif indicates the remaining glycans not matched by motifs which are a subset. Motif definition needs to be taken in the context of the model.

\*\*Accuracy describes the consistency between common-name definition of the motif and the formal, text-based definition of the motif, in terms of percent agreement in the glycans containing the two motifs. Common Name label definitions given here (http://carbogrove.org/MotifLabels.php).

\*\*\*P-Value refers to difference from Non-Binders with multiple testing correction (Dunnet's Test)

Motifs with a red motif ID fail to show a logistic response to protein concentration in the range of concentrations analyzed. These motifs may be nonbinding motifs (motifs which define nonbinding exceptions) or simply fail to fit. Nonbinding motifs are determined based on concentration dependent response when available or the average binding of non-motif glycans otherwise.



#### **Boxplot of Fine-Specificity Motifs**



#### Motif Intensity Map



GSII



Figure 2. Glycan intensity and motif distribution plot. The top half of the plot presents the observed glycan binding intensity of various glycans used in the array over their rank binding intensity; only the top glycans are shown. The second plot indicates the position of glycans containing the various motifs in the top plot with a yellow tick.

### Motif Family Membership Map

A			В	
				Relative Binding
				0.8
				0.6
				0.4
# Family Sub-Fan #	nily Member #	*Family Remainder		

\*Motif indicates the remaining glycans not matched by motifs which are a subset. Motif definition needs to be taken in the context of the model.

Figure 3. Treemap of glycan binding grouped by motif and family structure. The model structure can be represented as nested boxes where box size is proportional to the number of glycans with the motif and color changes with change in average relative binding of glycans with the motif. Only three layers of data splitting are included here, though further splitting may be possible.

#### **Detailed Model Breakdown**

Motif Glycan Examples:



11/14/2020

GSII



Figure 4. Boxplots of glycan binding grouped by motif for each dataset in the model. Motifs are listed in ascending average binding intensity (for the selected concentration) and colored by family.

Model Structure:

GSII



\*Motif matches the remaining glycans not matched by earlier motifs in the model.

Figure 5. Tree representation of the regression tree model trained on array data. Data flows through the tree (topdown) and is split by the various motifs. The motif used the split the data at each point has the id "family+split number" except when further split. In the case of further splits the id of the motif used to split the data is denoted with an asterisk.

#### Curve Fitting:

No curves were fit for model

Motif Text Structures:



11/14/2020

#### Key:



# Automated Analysis report for Galectin3 evaluated at 50 ug/mL

#### List of Primary Galectin3 Motifs

Primary Motif ID	Primary Motif Minimal Structure	Primary Motif Complete Structure	Relative Binding	Number of Glycans
A	$R \xrightarrow{4} \beta^{2} \alpha$ $R \xrightarrow{3} \beta^{4} \beta^{2} \alpha$ $R \xrightarrow{3} \beta^{4} \beta^{2} \alpha$	$R \xrightarrow{4} 3^{\beta 2} \alpha$ $R \xrightarrow{4} 3^{6} \beta^{4} 3^{\beta 4} 3^{\beta 4$	1.00	16
В	$\mathbf{R} \xrightarrow{3} \mathbf{O}^{\beta 4} \mathbf{R}$	$R^{-\frac{3}{\beta^2}} \xrightarrow{\beta^2} \xrightarrow{\alpha}^{\beta^4} \xrightarrow{\beta^2} \xrightarrow{\alpha}^{\beta^4}$	0.25	24
0			-0.05	32

Minimal and complete motif definitions match the same set of glycans defined in the array. Components found in the complete motif but excluded in the minimal motif are not tested by the array. Monosaccharide identities and subsitution intolerance may or may not have been tested by the array, they are retained from the complete motif for readability.

### **Boxplot of Primary Motifs**





## List of Fine-Specificity Galectin3 Motifs

Motif ID	Nearest Common Name (Accuracy%**)	Motif Graphic Structure	Relative Binding	Number of Glycans	P-Value***
A1		$R \xrightarrow{3} \beta^{4} \xrightarrow{\beta^{2}} \alpha^{\alpha}$ $R \xrightarrow{3} \beta^{4} \xrightarrow{\beta^{2}} \alpha^{\alpha}$ $R \xrightarrow{3} \beta^{4} \xrightarrow{\beta^{2}} \alpha^{\alpha}$	1.00	4	<0.001
A0*		$R \xrightarrow{4} \beta^{\beta} 2 \qquad \alpha \qquad R \xrightarrow{4} \beta^{\beta} \beta \qquad R$ $R \xrightarrow{3} \beta^{\beta} \beta^{\beta} \qquad \beta^{\beta} \beta^{\beta} \qquad R$	0.53	8	<0.001
A2	a6 Sialyl Type 2 LacNAc a3 Sialyl Type 2 LacNAc (94%)	$\mathbf{R}^{\mathbf{\alpha}6} \mathbf{\beta}^{\mathbf{\beta}4} \mathbf{\beta}^{\mathbf{\beta}2} \mathbf{R}^{\mathbf{\alpha}\mathbf{\beta}\mathbf{\beta}\mathbf{\beta}\mathbf{\beta}\mathbf{\beta}\mathbf{\beta}\mathbf{\beta}\mathbf{\beta}\mathbf{\beta}\beta$	0.18	4	0.328

Motif ID	Nearest Commo Name (Accuracy%**)	on ) Motif Graphic Structure	Relative Binding	Number of Glycans	P-Value***
B0		$\mathbf{R} \xrightarrow{3} \beta^{\beta 4} \mathbf{R}$	0.17	24	0.018
0	Non-Binders (100%)		0.00	32	NA
Key:					
Hexose () H Glc () () Man () () Gal ()	IexNAc □ Kdn ♦ GlcNAc ■ Neu5Ac ♦ GalNAc □ Neu5Gc ♦ Fuc ▲ Sia ♦	Any Single Monosaccharide Any # of Monosaccharide ~SKIP~ Hydrogen or Monosaccharide R-			

See Symbol Nomenclature for Glycans (SNFG) for complete key: https://www.ncbi.nlm.nih.gov/glycans/snfg.html (https://www.ncbi.nlm.nih.gov/glycans/snfg.html)

\*Motif indicates the remaining glycans not matched by motifs which are a subset. Motif definition needs to be taken in the context of the model.

\*\*Accuracy describes the consistency between common-name definition of the motif and the formal, text-based definition of the motif, in terms of percent agreement in the glycans containing the two motifs. Common Name label definitions given here (http://carbogrove.org/MotifLabels.php).

\*\*\*P-Value refers to difference from Non-Binders with multiple testing correction (Dunnet's Test)

Motifs with a red motif ID fail to show a logistic response to protein concentration in the range of concentrations analyzed. These motifs may be nonbinding motifs (motifs which define nonbinding exceptions) or simply fail to fit. Nonbinding motifs are determined based on concentration dependent response when available or the average binding of non-motif glycans otherwise.

### **Boxplot of Fine-Specificity Motifs**





#### Motif Intensity Map





Figure 2. Glycan intensity and motif distribution plot. The top half of the plot presents the observed glycan binding intensity of various glycans used in the array over their rank binding intensity; only the top glycans are shown. The second plot indicates the position of glycans containing the various motifs in the top plot with a yellow tick.

### Motif Family Membership Map

A 2	B	
1		Relative Binding
		0.6 0.4 0.2
0*		
# Family Sub-Family	y Member *Family # Remainder	

\*Motif indicates the remaining glycans not matched by motifs which are a subset. Motif definition needs to be taken in the context of the model.

Figure 3. Treemap of glycan binding grouped by motif and family structure. The model structure can be represented as nested boxes where box size is proportional to the number of glycans with the motif and color changes with change in average relative binding of glycans with the motif. Only three layers of data splitting are included here, though further splitting may be possible.

#### **Detailed Model Breakdown**

Motif Glycan Examples:





Figure 4. Boxplots of glycan binding grouped by motif for each dataset in the model. Motifs are listed in ascending average binding intensity (for the selected concentration) and colored by family.

Model Structure:



\*Motif matches the remaining glycans not matched by earlier motifs in the model.

Figure 5. Tree representation of the regression tree model trained on array data. Data flows through the tree (topdown) and is split by the various motifs. The motif used the split the data at each point has the id "family+split number" except when further split. In the case of futher splits the id of the motif used to split the data is denoted with an asterisk.

Curve Fitting:

No curves were fit for model

Motif Text Structures:

Motif ID	Motif Graphic	Motif Text
A1	$R \xrightarrow{3} 6^{\frac{\beta}{4}} 8^{\frac{\beta}{2}} 6^{\frac{\beta}{4}} R$ $R \xrightarrow{3} 6^{\frac{\beta}{4}} 8^{\frac{\beta}{2}} 6^{\frac{\beta}{4}}$	<2f4f6f>GalB1-4<3f6f>GlcNAcB1-2<3f4f6f>ManA1-6(<2f4f6f>GalB1- 4<3f6f>GlcNAcB1-2<3f4f6f>ManA1-3)(<3f4f6f>GlcNAcB1-4)<2f>ManB
A2	$ \begin{array}{c} & & & & \\ & & & & \\ & & & \\ & & & & \\ & & & & \\ & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & \\ & & & & \\ & & & $	<6f8f>Neu5AcA2-6<2f3f4f>GalB1-4<3f6f>GlcNAcB1-2<3f4f6f>ManA1- 6(<2f4f6f>GalB1-4<3f6f>GlcNAcB1-2<3f4f6f>ManA1-3)<2f>ManB

Gal 🔵

Fuc 🔺

Galectin3

Motif ID	Motif Graphic	Motif Text
A0*	$R \xrightarrow{4} B^{2} O_{\alpha}$ $R \xrightarrow{5} B^{4} R$ $R \xrightarrow{5} O^{64} D^{2} O^{6}$	<2f4f6f>GalB1-4<3f6f>GlcNAcB1-2<3f4f6f>ManA1-3(<3f6f>GlcNAcB1- 2<3f4f6f>ManA1-6)<2f>ManB
B0	$R \xrightarrow{3} \beta^{\beta 4} \overline{\beta}^{\beta} R$	<2f4f6f>GalB1-4<3f6f>GlcNAcB
0		Non-Binders
Key:		
Hexose 🔿	HexNAc 🗌 Kdn	Any Single
Glc 🔵	GlcNAc Neu5Ac	Any # of
Man 🔵	GalNAc 📃 Neu5Gc	Monosaccharide ~SKIP~

BubGc Hydrogen or Sia A Monosaccharide

R-

# Automated Analysis report for Siglec3 evaluated at 20 ug/mL

### List of Primary Siglec3 Motifs

Primary Motif ID	Primary Motif Minimal Structure	Primary Motif Complete Structure	Relative Binding	Number of Glycans
A	$ \begin{array}{c}                                     $	$ \begin{array}{c}  & \alpha^{3} \\  & \beta^{4} \\  & \beta^{2} \\  & R \\  & R \\  & \alpha^{6} \\  & \beta^{4} \\  & \beta^{2} \\  & \beta^{4} \\ $	1.00	4
В	$R = \frac{1}{3} R = \frac{1}{3} R$	$\mathbf{R}^{\mathbf{A}/6}$	0.10	32
0			0.02	36

Minimal and complete motif definitions match the same set of glycans defined in the array. Components found in the complete motif but excluded in the minimal motif are not tested by the array. Monosaccharide identities and subsitution intolerance may or may not have been tested by the array, they are retained from the complete motif for readability.

### **Boxplot of Primary Motifs**

11/14/2020

Siglec3



## List of Fine-Specificity Siglec3 Motifs

Motif ID	Nearest Common Name (Accuracy%**)	Motif Graphic Structure	Relative Binding	Number of Glycans	P-Value***
A0	a6 Sialyl Type 2 LacNAc a3 Sialyl Type 2 LacNAc (94%)	$ \begin{array}{c}                                     $	1.00	4	<0.001
B0		$\mathbf{R}^{\mathbf{A}}^{\mathbf{B}}$	0.09	32	0.028
0	Non-Binders (100%)		0.00	36	NA
Key:					
Hexose 🔾	HexNAc 🗌 Kdn 🔶 Ar	ny Single			
Glc 🔵	GlcNAc Neu5Ac 🍐 A	ny # of			
Man 🔵	GalNAc Neu5Gc 🍐 Mono	saccharide ~SKIP~			
Gal 🔵	Fuc 🔺 🛛 Sia 📥 Mono	drogen or R-			

See Symbol Nomenclature for Glycans (SNFG) for complete key: https://www.ncbi.nlm.nih.gov/glycans/snfg.html (https://www.ncbi.nlm.nih.gov/glycans/snfg.html)

\*Motif indicates the remaining glycans not matched by motifs which are a subset. Motif definition needs to be taken in the context of the model.

\*\*Accuracy describes the consistency between common-name definition of the motif and the formal, text-based definition of the motif, in terms of percent agreement in the glycans containing the two motifs. Common Name label definitions given here (http://carbogrove.org/MotifLabels.php).

\*\*\*P-Value refers to difference from Non-Binders with multiple testing correction (Dunnet's Test)

Motifs with a red motif ID fail to show a logistic response to protein concentration in the range of concentrations analyzed. These motifs may be nonbinding motifs (motifs which define nonbinding exceptions) or simply fail to fit. Nonbinding motifs are determined based on concentration dependent response when available or the average binding of non-motif glycans otherwise.



#### **Boxplot of Fine-Specificity Motifs**

Figure 1. Glycan binding grouped by motif and motif family. Individual glycans are given as points on the plot.

## Motif Intensity Map



Figure 2. Glycan intensity and motif distribution plot. The top half of the plot presents the observed glycan binding intensity of various glycans used in the array over their rank binding intensity; only the top glycans are shown. The second plot indicates the position of glycans containing the various motifs in the top plot with a yellow tick.

#### Motif Family Membership Map

A						Β	Rel	ative Binding 1.00 0.75 0.50 0.25
# Fam	nily Su	ıb-Family	#	Member	*Family Remainder			

\*Motif indicates the remaining glycans not matched by motifs which are a subset. Motif definition needs to be taken in the context of the model.

Figure 3. Treemap of glycan binding grouped by motif and family structure. The model structure can be represented as nested boxes where box size is proportional to the number of glycans with the motif and color changes with change in average relative binding of glycans with the motif. Only three layers of data splitting are included here, though further splitting may be possible.

#### Detailed Model Breakdown

Motif Glycan Examples:

11/14/2020





Figure 4. Boxplots of glycan binding grouped by motif for each dataset in the model. Motifs are listed in ascending average binding intensity (for the selected concentration) and colored by family.

Model Structure:



\*Motif matches the remaining glycans not matched by earlier motifs in the model.

Figure 5. Tree representation of the regression tree model trained on array data. Data flows through the tree (topdown) and is split by the various motifs. The motif used the split the data at each point has the id "family+split number" except when further split. In the case of futher splits the id of the motif used to split the data is denoted with an asterisk.

#### Curve Fitting:

No curves were fit for model

Motif Text Structures:

ID	Motif Graphic	Motif Text
A0		<6f8f>Neu5AcA2-6<2f3f4f>GalB1-4<3f6f>GlcNAcB1-2<3f4f6f>ManA1- 3(<6f8f>Neu5AcA2-3<2f4f6f>GalB1-4<6f>GlcNAcB1-2<3f4f6f>ManA1-6)<2f>ManB
B0	$ \begin{array}{c} \mathbf{R} \\ \mathbf$	<6f8f>Neu5AcA2-<3or6><2f4f>GalB1-4<6f>GlcNAcB1-2<3f4f6f>ManA1-3<2f>ManB
0		Non-Binders

#### Motif

Key:

11/14/2020



Siglec3

# Automated Analysis report for Siglec10 evaluated at 20 ug/mL

#### List of Primary Siglec10 Motifs



Minimal and complete motif definitions match the same set of glycans defined in the array. Components found in the complete motif but excluded in the minimal motif are not tested by the array. Monosaccharide identities and subsitution intolerance may or may not have been tested by the array, they are retained from the complete motif for readability.

### **Boxplot of Primary Motifs**

11/14/2020

Siglec10



## List of Fine-Specificity Siglec10 Motifs

Motif IE	Nearest Common Name O (Accuracy%**)	Motif Graphic Structure	Relative Binding	Number of Glycans	P-Value***
A0	a6 Sialyl Type 2 LacNAc a3 Sialyl Type 2 LacNAc (94%)	$ \begin{array}{c}  & \alpha^{3} \\  & \beta^{4} \\  & R \\ $	1.00	4	<0.001
B0		$\mathbf{R}$	0.09	32	0.028
0	Non-Binders (100%)		0.00	36	NA
Key:					
Hexose () Glc () Man () Gal ()	HexNAc Kdn A GlcNAc Neu5Ac A GalNAc Neu5Gc A Fuc Sia A	Any Single nosaccharide Any # of iosaccharide ~SKIP~ ydrogen or nosaccharide R-			

Gal 🔵

See Symbol Nomenclature for Glycans (SNFG) for complete key: https://www.ncbi.nlm.nih.gov/glycans/snfg.html (https://www.ncbi.nlm.nih.gov/glycans/snfg.html)

\*Motif indicates the remaining glycans not matched by motifs which are a subset. Motif definition needs to be taken in the context of the model.

\*\*Accuracy describes the consistency between common-name definition of the motif and the formal, text-based definition of the motif, in terms of percent agreement in the glycans containing the two motifs. Common Name label definitions given here (http://carbogrove.org/MotifLabels.php).

\*\*\*P-Value refers to difference from Non-Binders with multiple testing correction (Dunnet's Test)

Motifs with a red motif ID fail to show a logistic response to protein concentration in the range of concentrations analyzed. These motifs may be nonbinding motifs (motifs which define nonbinding exceptions) or simply fail to fit. Nonbinding motifs are determined based on concentration dependent response when available or the average binding of non-motif glycans otherwise.



#### **Boxplot of Fine-Specificity Motifs**

Figure 1. Glycan binding grouped by motif and motif family. Individual glycans are given as points on the plot.

## Motif Intensity Map



Figure 2. Glycan intensity and motif distribution plot. The top half of the plot presents the observed glycan binding intensity of various glycans used in the array over their rank binding intensity; only the top glycans are shown. The second plot indicates the position of glycans containing the various motifs in the top plot with a yellow tick.

#### Motif Family Membership Map

A						В		
							Rel	ative Binding
								0.75
								0.50
								0.25
# Fa	imily #	Sub-Fami	y #	Member	*Family Remainder			

\*Motif indicates the remaining glycans not matched by motifs which are a subset. Motif definition needs to be taken in the context of the model.

Figure 3. Treemap of glycan binding grouped by motif and family structure. The model structure can be represented as nested boxes where box size is proportional to the number of glycans with the motif and color changes with change in average relative binding of glycans with the motif. Only three layers of data splitting are included here, though further splitting may be possible.

#### **Detailed Model Breakdown**

Motif Glycan Examples:





Figure 4. Boxplots of glycan binding grouped by motif for each dataset in the model. Motifs are listed in ascending average binding intensity (for the selected concentration) and colored by family.

Model Structure:



\*Motif matches the remaining glycans not matched by earlier motifs in the model.

Figure 5. Tree representation of the regression tree model trained on array data. Data flows through the tree (topdown) and is split by the various motifs. The motif used the split the data at each point has the id "family+split number" except when further split. In the case of further splits the id of the motif used to split the data is denoted with an asterisk.

#### Curve Fitting:

No curves were fit for model

Motif Text Structures:

ID	Motif Graphic	Motif Text
A0	$ \begin{array}{c}                                     $	<6f8f>Neu5AcA2-6<2f3f4f>GalB1-4<3f6f>GlcNAcB1-2<3f4f6f>ManA1- 3(<6f8f>Neu5AcA2-3<2f4f6f>GalB1-4<6f>GlcNAcB1-2<3f4f6f>ManA1-6)<2f>ManB
B0		<6f8f>Neu5AcA2-<3or6><2f4f>GalB1-4<6f>GlcNAcB1-2<3f4f6f>ManA1-3<2f>ManB
0		Non-Binders

#### Motif

Key:

11/14/2020



Siglec10

## **E-selectin**

# Automated Analysis report for E-selectin evaluated at 10 ug/mL

## List of Primary E-selectin Motifs



E-selectin

Minimal and complete motif definitions match the same set of glycans defined in the array. Components found in the complete motif but excluded in the minimal motif are not tested by the array. Monosaccharide identities and subsitution intolerance may or may not have been tested by the array, they are retained from the complete motif for readability.

#### **Boxplot of Primary Motifs**
10/16/2020

E-selectin



### List of Fine-Specificity E-selectin Motifs

Motif ID	Nearest Common Name (Accuracy%**)	Motif Graphic Structure	Relative Binding	Number of Glycans	P-Value***
Α4		$ \begin{array}{c}                                     $	1.00	5	<0.001
A2		$ \overset{\alpha 3}{\overset{\beta 4}{\overset{\beta 2}{\overset{\alpha}{\overset{\beta 2}{\overset{\alpha}{\overset{\beta 2}{\overset{\alpha}{\overset{\beta 2}{\overset{\alpha}{\overset{\beta 2}{\overset{\beta 2}{\overset{\alpha}{\overset{\beta 2}{\overset{\beta 2}{\overset{\alpha}{\overset{\beta 2}{\overset{\beta 2}{\overset{\beta 2}{\overset{\alpha}{\overset{\beta 2}{\overset{\beta 2}{\overset{\beta 2}{\overset{\alpha}{\overset{\beta 2}{\overset{\beta 2}}{\overset{\beta 2}}{\overset{\beta 2}{\overset{\beta 2}}{\overset{\beta 2}{\overset{\beta 2}}}}{\overset{\beta 2}{\overset{\beta 2}}{\overset{\beta 2}}}}}}}}}}}}}}}}}}}}}}}}}}}}}}}}}}}}$	0.92	9	<0.001
A5*	Sialyl Lewis X Bisecting N-Glycan (92%)	$\mathbf{R}$	0.70	5	<0.001

E-selectin

Nearest Common Name		n	Relative	Number of	
Motif II	D (Accuracy%**)	Motif Graphic Structure	Binding	Glycans	P-Value***
A1		$\mathbf{R} \xrightarrow{4} \mathbf{\beta}^{2} \mathbf{\beta}^{\alpha} \mathbf{\beta}^{\beta} \mathbf{\beta}^{\beta} \mathbf{R}$ $\mathbf{R} \xrightarrow{6} \mathbf{\beta}^{4} \mathbf{\beta}^{2} \mathbf{\beta}^{\alpha} \mathbf{\beta}^{\alpha} \mathbf{R}$ $\mathbf{R} \xrightarrow{6} \mathbf{\beta}^{4} \mathbf{\beta}^{2} \mathbf{\beta}^{\alpha} \mathbf{\beta}^{\alpha} \mathbf{R}$	0.41	6	0.023
A3		$\mathbf{R}^{\underline{\beta}4} \mathbf{R}^{\underline{\beta}2} \mathbf{R}^{\underline{\alpha}} \mathbf{R}^{\underline{\beta}2} \mathbf{R}^{\underline{\beta}$	0.01	5	1.000
A0*		$\mathbf{R} \xrightarrow{4} \mathbf{\beta} \mathbf{R}$	0.00	10	1.000
0	Non-Binders (100%)		0.00	32	NA
Key:					
Hexose () Glc () Man () Gal ()	HexNAC C Kdn GlcNAC Neu5AC GalNAC Neu5GC Fuc Sia	Any Single Monosaccharide Any # of Monosaccharide Hydrogen or Monosaccharide R-			

See Symbol Nomenclature for Glycans (SNFG) for complete key: https://www.ncbi.nlm.nih.gov/glycans/snfg.html (https://www.ncbi.nlm.nih.gov/glycans/snfg.html)

\*Motif indicates the remaining glycans not matched by motifs which are a subset. Motif definition needs to be taken in the context of the model.

\*\*Accuracy describes the consistency between common-name definition of the motif and the formal, text-based definition of the motif, in terms of percent agreement in the glycans containing the two motifs. Common Name label definitions given here (http://carbogrove.org/MotifLabels.php).

\*\*\*P-Value refers to difference from Non-Binders with multiple testing correction (Dunnet's Test)

Motifs with a red motif ID fail to show a logistic response to protein concentration in the range of concentrations analyzed. These motifs may be nonbinding motifs (motifs which define nonbinding exceptions) or simply fail to fit. Nonbinding motifs are determined based on concentration dependent response when available or the average binding of non-motif glycans otherwise.

#### **Boxplot of Fine-Specificity Motifs**



E-selectin



Figure 1. Glycan binding grouped by motif and motif family. Individual glycans are given as points on the plot.

#### Motif Intensity Map



E-selectin



Figure 2. Glycan intensity and motif distribution plot. The top half of the plot presents the observed glycan binding intensity of various glycans used in the array over their rank binding intensity; only the top glycans are shown. The second plot indicates the position of glycans containing the various motifs in the top plot with a yellow tick.

#### Motif Family Membership Map



\*Motif indicates the remaining glycans not matched by motifs which are a subset. Motif definition needs to be taken in the context of the model.

Figure 3. Treemap of glycan binding grouped by motif and family structure. The model structure can be represented as nested boxes where box size is proportional to the number of glycans with the motif and color changes with change in average relative binding of glycans with the motif. Only three layers of data splitting are included here, though further splitting may be possible.

#### Detailed Model Breakdown

Motif Glycan Examples:



E-selectin



All Concentration Plots:



Figure 4. Boxplots of glycan binding grouped by motif for each dataset in the model. Motifs are listed in ascending average binding intensity (for the selected concentration) and colored by family.

Model Structure:



\*Motif matches the remaining glycans not matched by earlier motifs in the model.

Figure 5. Tree representation of the regression tree model trained on array data. Data flows through the tree (topdown) and is split by the various motifs. The motif used the split the data at each point has the id "family+split number" except when further split. In the case of further splits the id of the motif used to split the data is denoted with an asterisk.

#### Curve Fitting:

No curves were fit for model

Motif Text Structures:

#### Motif

ID	Motif Graphic	Motif Text
A1	$\mathbf{R} \stackrel{4}{\xrightarrow{3}} \begin{array}{c} \beta 2 \\ \beta 2 \\ \beta 3 \\ \beta 4 \\ \beta 2 \\ \beta 4 \\ \beta 2 \\ \beta 2 \\ \beta 4 \\ \beta 2 \\ \beta 2 \\ \beta 4 \\ \beta 2 \\ \beta 4 \\ \beta 4 \\ \beta 2 \\ \beta 4 \\ \beta$	<2f4f>GalB1-4<3f6f>GlcNAcB1-2<3f4f6f>ManA1-3(<2f3f4f>FucA1-3<6f>GlcNAcB1- 2<3f4f6f>ManA1-6)(<3f4f6f>GlcNAcB1-4)<2f>ManB
A2		<6f8f>Neu5AcA2-3<2f4f6f>GalB1-4(<2f3f4f>FucA1-3)<6f>GlcNAcB1- 2<3f4f6f>ManA1-6<2f>ManB
A3	$\mathbf{R} \stackrel{\beta 4}{\longrightarrow} \mathbf{R} \stackrel{\beta 2}{\longrightarrow} \mathbf{R}$	<2f3f4f6f>GalB1-4(<2f3f4f>FucA1-3)<6f>GlcNAcB1-2<3f4f6f>ManA1- 6(<3f6f>GlcNAcB1-2<3f4f6f>ManA1-3)<2f>ManB

E-selectin

Motif ID	Motif Graphic	Motif Text
A4	Φ <sup>α3</sup> Φ <sup>β4</sup> Β <sup>β2</sup> α <sup>α</sup> <sup>β</sup> Β <sup>β</sup> -R	<6f8f>Neu5AcA2-3<2f4f6f>GalB1-4(<2f3f4f>FucA1-3)<6f>GlcNAcB1- 2<3f4f6f>ManA1-3<2f4f>ManB
A5*		<6f8f>Neu5AcA2-3<2f4f6f>GalB1-4(<2f3f4f>FucA1-3)<6f>GlcNAcB1- 2<3f4f6f>ManA1-3<2f>ManB
A0*	$\mathbf{R} \xrightarrow{4} \mathbf{\beta} \mathbf{R}$	<2f3f4f>FucA1-3<6f>GlcNAcB
0		Non-Binders
Key:		
Hexose 🔿	HexNAc 🗌 🛛 Kdn ┥	Any Single
Glc 🔵	GlcNAc 📃 Neu5Ac 🔌	Any # of
Man 🔵	GalNAc 📃 Neu5Gc 🧹	Monosaccharide ~SKIP~
Gal 🔵	Fuc 🔺 Sia 🧳	Monosaccharide R-

## CD15s

# Automated Analysis report for CD15s evaluated at 20 ug/mL

#### List of Primary CD15s Motifs

Primary Motif ID	Primary Motif Minimal Structure	Primary Motif Complete Structure	Relative Binding	Number of Glycans
A		$\mathbf{R}$	1.00	6
0			0.01	66

Minimal and complete motif definitions match the same set of glycans defined in the array. Components found in the complete motif but excluded in the minimal motif are not tested by the array. Monosaccharide identities and subsitution intolerance may or may not have been tested by the array, they are retained from the complete motif for readability.

#### **Boxplot of Primary Motifs**



Man 🔵

Gal 🔵

GalNAc 📃 Neu5Gc 🔶

Sia 🔶

Fuc 🔺

CD15s



#### List of Fine-Specificity CD15s Motifs

Hydrogen or

Monosaccharide

R-

	Nearest Common Name		Relative	Number of	
Motif ID	(Accuracy%**)	Motif Graphic Structure	Binding	Glycans	P-Value***
A0			1	6	<0.001
0	Non-Binders (100%)		0	66	NA
Key:					
Hexose O H	lexNAc □ Kdn ♦ Ar Mono GlcNAc ■ Neu5Ac ♦ A	ny Single osaccharide vny # of second second second			
Mon 🦱 🤇		sacchande			

See Symbol Nomenclature for Glycans (SNFG) for complete key: https://www.ncbi.nlm.nih.gov/glycans/snfg.html (https://www.ncbi.nlm.nih.gov/glycans/snfg.html)

\*Motif indicates the remaining glycans not matched by motifs which are a subset. Motif definition needs to be taken in the context of the model.

CD15s

\*\*Accuracy describes the consistency between common-name definition of the motif and the formal, text-based definition of the motif, in terms of percent agreement in the glycans containing the two motifs. Common Name label definitions given here (http://carbogrove.org/MotifLabels.php).

\*\*\*P-Value refers to difference from Non-Binders with multiple testing correction (Dunnet's Test)

Motifs with a red motif ID fail to show a logistic response to protein concentration in the range of concentrations analyzed. These motifs may be nonbinding motifs (motifs which define nonbinding exceptions) or simply fail to fit. Nonbinding motifs are determined based on concentration dependent response when available or the average binding of non-motif glycans otherwise.

## 

#### **Boxplot of Fine-Specificity Motifs**



#### Motif Intensity Map



CD15s



Figure 2. Glycan intensity and motif distribution plot. The top half of the plot presents the observed glycan binding intensity of various glycans used in the array over their rank binding intensity; only the top glycans are shown. The second plot indicates the position of glycans containing the various motifs in the top plot with a yellow tick.

#### Motif Family Membership Map

0/16/2020	CD15s
	Relative Binding
# Family Sub-Family # Mer	<sup>mber</sup> *Family Remainder

\*Motif indicates the remaining glycans not matched by motifs which are a subset. Motif definition needs to be taken in the context of the model.

Figure 3. Treemap of glycan binding grouped by motif and family structure. The model structure can be represented as nested boxes where box size is proportional to the number of glycans with the motif and color changes with change in average relative binding of glycans with the motif. Only three layers of data splitting are included here, though further splitting may be possible.

#### **Detailed Model Breakdown**

Motif Glycan Examples:



CD15s



Figure 4. Boxplots of glycan binding grouped by motif for each dataset in the model. Motifs are listed in ascending average binding intensity (for the selected concentration) and colored by family.

Model Structure:



\*Motif matches the remaining glycans not matched by earlier motifs in the model.

Figure 5. Tree representation of the regression tree model trained on array data. Data flows through the tree (topdown) and is split by the various motifs. The motif used the split the data at each point has the id "family+split number" except when further split. In the case of further splits the id of the motif used to split the data is denoted with an asterisk.

#### Curve Fitting:

No curves were fit for model

Motif Text Structures:



## RatDCIR2

# Automated Analysis report for RatDCIR2 evaluated at 20 ug/mL

#### List of Primary RatDCIR2 Motifs

Primary Motif ID	Primary Motif Minimal Structure	Primary Motif Complete Structure	Relative Binding	Number of Glycans
A	$\mathbf{R}$	$\mathbf{R} \xrightarrow{4}_{3} \xrightarrow{\beta2}_{\alpha} \xrightarrow{\alpha}_{\beta4} \xrightarrow{\beta4}_{\beta4} \xrightarrow{\beta4}_{\beta4}$	1	12
0			0	60

Minimal and complete motif definitions match the same set of glycans defined in the array. Components found in the complete motif but excluded in the minimal motif are not tested by the array. Monosaccharide identities and subsitution intolerance may or may not have been tested by the array, they are retained from the complete motif for readability.

#### **Boxplot of Primary Motifs**



A Motif ID

### List of Fine-Specificity RatDCIR2 Motifs

Motif IE	Nearest Common Name O (Accuracy%**)	Motif Graphic Structure	Relative Binding	Number of Glycans	P-Value***
A0*	Terminal N-Glycan N-Acetyl Glucosamine Bisecting N-Glycan (93%)	$\mathbf{R}$	1	6	<0.001
A1		$\mathbf{R} \stackrel{6}{\longrightarrow} \mathbf{R}$	0	6	1.000
0	Non-Binders (100%)		0	60	NA
Key:					
Hexose () Glc () Man () Gal ()	HexNAc Kdn A GlcNAc Neu5Ac A GalNAc Neu5Gc A Fuc Sia A	ny Single osaccharide vny # of osaccharide ~SKIP~ drogen or R- osaccharide R-			

file:///Z:/Shared Files/Glycan Arrays/Bisecting vs. Non-Bisecting Array/MF running for Li\_2020.10/RatDCIR2\_GeneratedReport\_1/RatDCIR2.html

RatDCIR2

See Symbol Nomenclature for Glycans (SNFG) for complete key: https://www.ncbi.nlm.nih.gov/glycans/snfg.html (https://www.ncbi.nlm.nih.gov/glycans/snfg.html)

\*Motif indicates the remaining glycans not matched by motifs which are a subset. Motif definition needs to be taken in the context of the model.

\*\*Accuracy describes the consistency between common-name definition of the motif and the formal, text-based definition of the motif, in terms of percent agreement in the glycans containing the two motifs. Common Name label definitions given here (http://carbogrove.org/MotifLabels.php).

\*\*\*P-Value refers to difference from Non-Binders with multiple testing correction (Dunnet's Test)

Motifs with a red motif ID fail to show a logistic response to protein concentration in the range of concentrations analyzed. These motifs may be nonbinding motifs (motifs which define nonbinding exceptions) or simply fail to fit. Nonbinding motifs are determined based on concentration dependent response when available or the average binding of non-motif glycans otherwise.



#### **Boxplot of Fine-Specificity Motifs**

Figure 1. Glycan binding grouped by motif and motif family. Individual glycans are given as points on the plot.

### Motif Intensity Map



RatDCIR2



Figure 2. Glycan intensity and motif distribution plot. The top half of the plot presents the observed glycan binding intensity of various glycans used in the array over their rank binding intensity; only the top glycans are shown. The second plot indicates the position of glycans containing the various motifs in the top plot with a yellow tick.

#### Motif Family Membership Map



\*Motif indicates the remaining glycans not matched by motifs which are a subset. Motif definition needs to be taken in the context of the model.

Figure 3. Treemap of glycan binding grouped by motif and family structure. The model structure can be represented as nested boxes where box size is proportional to the number of glycans with the motif and color changes with change in average relative binding of glycans with the motif. Only three layers of data splitting are included here, though further splitting may be possible.

#### **Detailed Model Breakdown**

Motif Glycan Examples:



RatDCIR2



Figure 4. Boxplots of glycan binding grouped by motif for each dataset in the model. Motifs are listed in ascending average binding intensity (for the selected concentration) and colored by family.

Model Structure:

6/8



\*Motif matches the remaining glycans not matched by earlier motifs in the model.

Figure 5. Tree representation of the regression tree model trained on array data. Data flows through the tree (topdown) and is split by the various motifs. The motif used the split the data at each point has the id "family+split number" except when further split. In the case of futher splits the id of the motif used to split the data is denoted with an asterisk.

#### Curve Fitting:

No curves were fit for model

Motif Text Structures:



10/16/2020

Key:



RatDCIR2

## NR19240 (PR8\_H1)

# Automated Analysis report for NR19240 evaluated at 20 ug/mL

#### List of Primary NR19240 Motifs

Primary Motif ID	Primary Motif Minimal Structure	Primary Motif Complete Structure	Relative Binding	Number of Glycans
A	$\mathbf{A}^{\alpha}$ $\mathbf{A}^{\beta}$	$\mathbf{R}$	1.00	40
0			-0.01	32

Minimal and complete motif definitions match the same set of glycans defined in the array. Components found in the complete motif but excluded in the minimal motif are not tested by the array. Monosaccharide identities and subsitution intolerance may or may not have been tested by the array, they are retained from the complete motif for readability.

#### **Boxplot of Primary Motifs**

NR19240



### List of Fine-Specificity NR19240 Motifs

Motif ID	Nearest Common Name (Accuracy%**)	Motif Graphic Structure	Relative Binding	Number of Glycans	P-Value***
A0*		$\mathbf{A}^{\alpha 3} \mathbf{A}^{\beta} \mathbf{R}$	1.00	13	<0.001
A1		$ \begin{array}{c}                                     $	0.66	24	<0.001
A2	a3 Sialyl Type 2 LacNAc Terminal beta-Galactose (90%)	$\mathbf{R} \xrightarrow{6} \beta_{4} \beta_{2} \qquad \alpha_{6} \beta_{6} \beta_{7} \mathbf{R}$	0.65	3	<0.001
0	Non-Binders (100%)		0.00	32	NA

Key:



See Symbol Nomenclature for Glycans (SNFG) for complete key: https://www.ncbi.nlm.nih.gov/glycans/snfg.html (https://www.ncbi.nlm.nih.gov/glycans/snfg.html)

NR19240

\*Motif indicates the remaining glycans not matched by motifs which are a subset. Motif definition needs to be taken in the context of the model.

\*\*Accuracy describes the consistency between common-name definition of the motif and the formal, text-based definition of the motif, in terms of percent agreement in the glycans containing the two motifs. Common Name label definitions given here (http://carbogrove.org/MotifLabels.php).

\*\*\*P-Value refers to difference from Non-Binders with multiple testing correction (Dunnet's Test)

Motifs with a red motif ID fail to show a logistic response to protein concentration in the range of concentrations analyzed. These motifs may be nonbinding motifs (motifs which define nonbinding exceptions) or simply fail to fit. Nonbinding motifs are determined based on concentration dependent response when available or the average binding of non-motif glycans otherwise.



#### **Boxplot of Fine-Specificity Motifs**

Figure 1. Glycan binding grouped by motif and motif family. Individual glycans are given as points on the plot.

#### Motif Intensity Map



**Top Ordered Glycans** 

Figure 2. Glycan intensity and motif distribution plot. The top half of the plot presents the observed glycan binding intensity of various glycans used in the array over their rank binding intensity; only the top glycans are shown. The second plot indicates the position of glycans containing the various motifs in the top plot with a yellow tick.

#### Motif Family Membership Map

2	Α	Relative Binding 1.0 0.9 0.8 0.7
# Family Sub-Fan #	nily Member *Family # Remainder	

\*Motif indicates the remaining glycans not matched by motifs which are a subset. Motif definition needs to be taken in the context of the model.

Figure 3. Treemap of glycan binding grouped by motif and family structure. The model structure can be represented as nested boxes where box size is proportional to the number of glycans with the motif and color changes with change in average relative binding of glycans with the motif. Only three layers of data splitting are included here, though further splitting may be possible.

#### **Detailed Model Breakdown**

Motif Glycan Examples:



NR19240



Figure 4. Boxplots of glycan binding grouped by motif for each dataset in the model. Motifs are listed in ascending average binding intensity (for the selected concentration) and colored by family.

Model Structure:



\*Motif matches the remaining glycans not matched by earlier motifs in the model.

Figure 5. Tree representation of the regression tree model trained on array data. Data flows through the tree (topdown) and is split by the various motifs. The motif used the split the data at each point has the id "family+split number" except when further split. In the case of futher splits the id of the motif used to split the data is denoted with an asterisk.

Curve Fitting:

No curves were fit for model

Motif Text Structures:



NR19240



## NR19441 (NY18H1)

# Automated Analysis report for NR19441 evaluated at 20 ug/mL

#### List of Primary NR19441 Motifs

Primary Motif ID	Primary Motif Minimal Structure	Primary Motif Complete Structure	Relative Binding	Number of Glycans
А	$\mathbf{A}^{\alpha 6} \mathbf{A}^{\beta 4} \mathbf{A}^{\beta 2} \mathbf{A}^{\alpha 6} \mathbf{A}^{\beta 6} \mathbf{A}^{\beta 6} \mathbf{R}$	$\mathbf{\Phi}^{\alpha 6} \mathbf{\Phi}^{\beta 4} \mathbf{\Phi}^{\beta 2} \mathbf{\Phi}^{\mathfrak{B}/6} \mathbf{\Phi}^{\beta 4} \mathbf{\Phi}^{$	1.00	11
0			0.01	61

Minimal and complete motif definitions match the same set of glycans defined in the array. Components found in the complete motif but excluded in the minimal motif are not tested by the array. Monosaccharide identities and subsitution intolerance may or may not have been tested by the array, they are retained from the complete motif for readability.

#### **Boxplot of Primary Motifs**



NR19441

#### List of Fine-Specificity NR19441 Motifs

Nearest Common		on			
Name			Relative Number of		
Motif II	D (Accuracy%**	<sup>t</sup> ) Motif Graphic Structure	Binding	Glycans	P-Value***
A0		$\mathbf{A}^{\alpha 6} \mathbf{A}^{\beta 4} \mathbf{A}^{\beta 2} \mathbf{A}^{\alpha \beta / 6} \mathbf{A}^{\beta} \mathbf{R}$	1	11	<0.001
0	Non-Binders (10	0%)	0	61	NA
Key:					
Hexose 🔿	HexNAc 🗌 🛛 Kdn 🔶	Any Single			
Glc 🔵	GlcNAc 📕 Neu5Ac 🔶	Any # of			
Man 🔵	GalNAc 📃 Neu5Gc 🔶	Monosaccharide ~SKIP~			
Gal 🔵	Fuc 🔺 Sia 🍎	Hydrogen or R- Monosaccharide			

See Symbol Nomenclature for Glycans (SNFG) for complete key: https://www.ncbi.nlm.nih.gov/glycans/snfg.html (https://www.ncbi.nlm.nih.gov/glycans/snfg.html)

\*Motif indicates the remaining glycans not matched by motifs which are a subset. Motif definition needs to be taken in the context of the model.

\*\*Accuracy describes the consistency between common-name definition of the motif and the formal, text-based definition of the motif, in terms of percent agreement in the glycans containing the two motifs. Common Name label definitions given here (http://carbogrove.org/MotifLabels.php).

\*\*\*P-Value refers to difference from Non-Binders with multiple testing correction (Dunnet's Test)

Motifs with a red motif ID fail to show a logistic response to protein concentration in the range of concentrations analyzed. These motifs may be nonbinding motifs (motifs which define nonbinding exceptions) or simply fail to fit. Nonbinding motifs are determined based on concentration dependent response when available or the average binding of non-motif glycans otherwise.

#### **Boxplot of Fine-Specificity Motifs**





#### Motif Intensity Map



NR19441



Top Ordered Glycans

Figure 2. Glycan intensity and motif distribution plot. The top half of the plot presents the observed glycan binding intensity of various glycans used in the array over their rank binding intensity; only the top glycans are shown. The second plot indicates the position of glycans containing the various motifs in the top plot with a yellow tick.

#### Motif Family Membership Map
A	Relative Binding
# Family Sub-Family # Member *Family # Remainder	

\*Motif indicates the remaining glycans not matched by motifs which are a subset. Motif definition needs to be taken in the context of the model.

Figure 3. Treemap of glycan binding grouped by motif and family structure. The model structure can be represented as nested boxes where box size is proportional to the number of glycans with the motif and color changes with change in average relative binding of glycans with the motif. Only three layers of data splitting are included here, though further splitting may be possible.

### Detailed Model Breakdown

Motif Glycan Examples:



10/16/2020

NR19441





All Concentration Plots:



Figure 4. Boxplots of glycan binding grouped by motif for each dataset in the model. Motifs are listed in ascending average binding intensity (for the selected concentration) and colored by family.

Model Structure:



\*Motif matches the remaining glycans not matched by earlier motifs in the model.

Figure 5. Tree representation of the regression tree model trained on array data. Data flows through the tree (topdown) and is split by the various motifs. The motif used the split the data at each point has the id "family+split number" except when further split. In the case of futher splits the id of the motif used to split the data is denoted with an asterisk.

#### Curve Fitting:

No curves were fit for model

Motif Text Structures:

Motif ID	Motif Graphic	c Motif Text
A0		<pre>&lt;6f8f&gt;Neu5AcA2-6&lt;2f3f4f&gt;GalB1-4&lt;3f6f&gt;GlcNAcB1-2&lt;3f4f6f&gt;ManA1-&lt;3or6&gt; &lt;2f4f&gt;ManB</pre>
0		Non-Binders
Key:		
Hexose 🔾	HexNAc 🗌 Kdn 🔶	Any Single
Glc 🔵	GlcNAc 📕 Neu5Ac 🔶	Any # of
Man 🔵	GalNAc 📃 Neu5Gc 🔶	Hydrogen or
Gal 🔵	Fuc 🔺 Sia 🔶	Monosaccharide R-

# NR41637 (SP27\_H1)

# Automated Analysis report for NR41637 evaluated at 20 ug/mL

### List of Primary NR41637 Motifs

Primary Motif ID	Primary Motif Minimal Structure	Primary Motif Complete Structure	Relative Binding	Number of Glycans
A	$\mathbf{A}^{\alpha  3} \mathbf{A}^{\beta} \mathbf{R}$	$\mathbf{R}^{\alpha 3} \mathbf{R}^{\beta 4} \mathbf{R}^{\alpha 3} \mathbf{R}^{\beta 4} \mathbf{R}^{\beta 4$	1.00	40
0			-0.02	32

Minimal and complete motif definitions match the same set of glycans defined in the array. Components found in the complete motif but excluded in the minimal motif are not tested by the array. Monosaccharide identities and subsitution intolerance may or may not have been tested by the array, they are retained from the complete motif for readability.

### **Boxplot of Primary Motifs**



### List of Fine-Specificity NR41637 Motifs

Motif ID	Nearest Common Name (Accuracy%**)	Motif Graphic Structure	Relative Binding	Number of Glycans	P-Value***
A3	Sialyl Lewis X a3 Sialyl Type 2 LacNAc (94%)	$ \begin{array}{c} & \alpha 3 \\ & \beta 4 \\ & 3 \\ & \alpha 3 \\ & \alpha 3 \\ & \alpha 3 \\ & \beta 4 \\ & \beta 2 \\ & \alpha 3 \\ & \beta 4 \\ & \beta 2 \\ & \alpha 3 \\ & \beta 2 \\ & \alpha 3 \\ & \beta 4 \\ & \beta 2 \\ & \alpha 3 \\ & \beta 2 \\ & \beta 2 \\ & \alpha 3 \\ & \beta 2 \\ $	1.00	4	<0.001
A5*	a3 Sialyl Galactose Terminal N-Glycan N- Acetyl Glucosamine (90%)	$ \begin{array}{c} & \alpha 3 \\ & \beta 4 \\ & 3 \\ & R \\ $	0.73	5	<0.001
A0*		$\mathbf{A}^{\alpha 3} \mathbf{A}^{\beta} \mathbf{R}$	0.55	10	<0.001
A2*		$ \begin{array}{c} & \alpha^{3} \\ & \beta^{4} \\ & \beta^{2} \\ & R $	0.35	9	<0.001

2/10

	Nearest Common Name		Relative	Number of	
Motif ID	(Accuracy%**)	Motif Graphic Structure	Binding	Glycans	P-Value***
A4		$ \begin{array}{c} & \alpha^{3} \\ & \beta^{4} \\ & \beta^{2} \\ & \beta^{4} \\ & \beta^{6} \\ & \beta^{6} \\ & R \end{array} $	0.34	3	0.190
A6	a3 Sialyl Type 2 LacNAc Terminal beta-Galactose (92%)	$\mathbf{R} \xrightarrow{6} \mathbf{\beta}_{4} \xrightarrow{3} \mathbf{\beta}_{2} \xrightarrow{\alpha} \mathbf{R} \xrightarrow{4} \xrightarrow{6} \mathbf{\beta}_{5} \xrightarrow{\alpha} \mathbf{R}$	0.04	6	1.000
A1	a3 Sialyl Type 2 LacNAc Terminal beta-Galactose (90%)	$ \begin{array}{c} & \alpha^{3} \\ & \beta^{4} \\ & \beta^{2} \\ & \alpha \\ & \beta^{4} \\ & \beta^{2} \\ & $	0.01	3	1.000
0	Non-Binders (100%)		0.00	32	NA
Key:					
Hexose 🔿 占	lexNAc 🗌 Kdn 🔶 🔤 Any	y Single			



See Symbol Nomenclature for Glycans (SNFG) for complete key: https://www.ncbi.nlm.nih.gov/glycans/snfg.html (https://www.ncbi.nlm.nih.gov/glycans/snfg.html)

\*Motif indicates the remaining glycans not matched by motifs which are a subset. Motif definition needs to be taken in the context of the model.

\*\*Accuracy describes the consistency between common-name definition of the motif and the formal, text-based definition of the motif, in terms of percent agreement in the glycans containing the two motifs. Common Name label definitions given here (http://carbogrove.org/MotifLabels.php).

\*\*\*P-Value refers to difference from Non-Binders with multiple testing correction (Dunnet's Test)

Motifs with a red motif ID fail to show a logistic response to protein concentration in the range of concentrations analyzed. These motifs may be nonbinding motifs (motifs which define nonbinding exceptions) or simply fail to fit. Nonbinding motifs are determined based on concentration dependent response when available or the average binding of non-motif glycans otherwise.

### **Boxplot of Fine-Specificity Motifs**



Figure 1. Glycan binding grouped by motif and motif family. Individual glycans are given as points on the plot.

### Motif Intensity Map

4/10



Figure 2. Glycan intensity and motif distribution plot. The top half of the plot presents the observed glycan binding intensity of various glycans used in the array over their rank binding intensity; only the top glycans are shown. The second plot indicates the position of glycans containing the various motifs in the top plot with a yellow tick.

### Motif Family Membership Map



\*Motif indicates the remaining glycans not matched by motifs which are a subset. Motif definition needs to be taken in the context of the model.

Figure 3. Treemap of glycan binding grouped by motif and family structure. The model structure can be represented as nested boxes where box size is proportional to the number of glycans with the motif and color changes with change in average relative binding of glycans with the motif. Only three layers of data splitting are included here, though further splitting may be possible.

### **Detailed Model Breakdown**

Motif Glycan Examples:



6/10



All Concentration Plots:

V1.0\_BisectingN\_20ug/mL



Figure 4. Boxplots of glycan binding grouped by motif for each dataset in the model. Motifs are listed in ascending average binding intensity (for the selected concentration) and colored by family.

Model Structure:



\*Motif matches the remaining glycans not matched by earlier motifs in the model.

Figure 5. Tree representation of the regression tree model trained on array data. Data flows through the tree (topdown) and is split by the various motifs. The motif used the split the data at each point has the id "family+split number" except when further split. In the case of further splits the id of the motif used to split the data is denoted with an asterisk.

Curve Fitting:

No curves were fit for model

Motif Text Structures:

Motif ID	Motif Graphic	Motif Text
A1	$ \begin{array}{c} 0 \xrightarrow{3} \xrightarrow{3} \xrightarrow{6} \xrightarrow{6} \xrightarrow{6} \xrightarrow{6} \xrightarrow{6} \xrightarrow{6} \xrightarrow{6} \xrightarrow{7} \xrightarrow{7} \xrightarrow{7} \xrightarrow{7} \xrightarrow{7} \xrightarrow{7} \xrightarrow{7} 7$	<6f8f>Neu5AcA2-3<2f4f6f>GalB1-4<3f6f>GlcNAcB1-2<3f4f6f>ManA1-6(<2f3f4f>GalB1- 4<6f>GlcNAcB1-2<3f4f6f>ManA1-3)(<3f4f6f>GlcNAcB1-4)<2f>ManB
A2*		<6f8f>Neu5AcA2-3<2f4f6f>GalB1-4<6f>GlcNAcB1-2<3f4f6f>ManA1-6(<2f3f4f>GalB1- 4<6f>GlcNAcB1-2<3f4f6f>ManA1-3)<2f>ManB
A3		<6f8f>Neu5AcA2-3<2f4f6f>GalB1-4(<2f3f4f>FucA1-3)<6f>GlcNAcB1-2<3f4f6f>ManA1- 6(<6f8f>Neu5AcA2-3<2f4f6f>GalB1-4<6f>GlcNAcB1-2<3f4f6f>ManA1-3)<2f>ManB
A4		<6f8f>Neu5AcA2-3<2f4f6f>GalB1-4<3f6f>GlcNAcB1-2<3f4f6f>ManA1- 6(<3f4f6f>GlcNAcB1-4)<2f>ManB

10/16/2020

NR41637

Motif ID	Motif Graphi	Motif Text
A5*	Ф <sup>23</sup>	<6f8f>Neu5AcA2-3<2f4f6f>GalB1-4<6f>GlcNAcB1-2<3f4f6f>ManA1-6<2f>ManB
A6		<6f8f>Neu5AcA2-3<2f4f6f>GalB1-4<3f6f>GlcNAcB1-2<3f4f6f>ManA1-3(<2f3f4f>GalB1- 4<6f>GlcNAcB1-2<3f4f6f>ManA1-6)<2f>ManB
A0*	$\mathbf{\mathbf{A}}^{\alpha3}\mathbf{\mathbf{A}}^{\beta}\mathbf{\mathbf{R}}$	<6f8f>Neu5AcA2-3<2f4f6f>GalB
0		Non-Binders
Key:		
Hexose () Glc () Man () Gal ()	HexNAc GlcNAc GalNAc Fuc	Kdn Any Single Monosaccharide Any # of Neu5Gc Monosaccharide ~SKIP~ Hydrogen or Sia Monosaccharide R-

## NR42486 (CR32\_H1)

# Automated Analysis report for NR42486 evaluated at 20 ug/mL

### List of Primary NR42486 Motifs

Primary Motif ID	Primary Motif Minimal Structure	Primary Motif Complete Structure	Relative Binding	Number of Glycans
А	$\mathbf{A}^{\alpha 6} \mathbf{O}^{\beta 4} \mathbf{1}^{\beta 2} \mathbf{6}^{6 6 6 \mathbf{\beta}} \mathbf{\mathbf{R}}$	$\mathbf{a}^{\alpha 6} \mathbf{a}^{\beta 4} \mathbf{a}^{\beta 2} \mathbf{a}^{\beta 6} \mathbf{a}^{\beta 4} \mathbf{a}^{\beta 4$	1	11
0			0	61

Minimal and complete motif definitions match the same set of glycans defined in the array. Components found in the complete motif but excluded in the minimal motif are not tested by the array. Monosaccharide identities and subsitution intolerance may or may not have been tested by the array, they are retained from the complete motif for readability.

### **Boxplot of Primary Motifs**



### List of Fine-Specificity NR42486 Motifs

	Nearest Comm Name	ion	Relative	Number of	
Motif II	D (Accuracy%*	*) Motif Graphic Structure	Binding	Glycans	P-Value***
A0		$\mathbf{A}^{\alpha 6} \mathbf{A}^{\beta 4} \mathbf{A}^{\beta 2} \mathbf{A}^{\alpha 6} \mathbf{A}^{\beta - \alpha} \mathbf{R}$	1	11	<0.001
0	Non-Binders (10	0%)	0	61	NA
Key:					
Hexose 🔾	HexNAc 🗌 🛛 Kdn 🔶	Any Single			
Glc 🔵	GlcNAc 📕 Neu5Ac 🔶	Any # of			
Man 🔵 Gal 🔵	GalNAc Neu5Gc	Monosaccharide ~ <sup>SKIP~</sup> Hydrogen or Monosaccharide <b>R</b> -			

See Symbol Nomenclature for Glycans (SNFG) for complete key: https://www.ncbi.nlm.nih.gov/glycans/snfg.html (https://www.ncbi.nlm.nih.gov/glycans/snfg.html)

\*Motif indicates the remaining glycans not matched by motifs which are a subset. Motif definition needs to be taken in the context of the model.

\*\*Accuracy describes the consistency between common-name definition of the motif and the formal, text-based definition of the motif, in terms of percent agreement in the glycans containing the two motifs. Common Name label definitions given here (http://carbogrove.org/MotifLabels.php).

\*\*\*P-Value refers to difference from Non-Binders with multiple testing correction (Dunnet's Test)

Motifs with a red motif ID fail to show a logistic response to protein concentration in the range of concentrations analyzed. These motifs may be nonbinding motifs (motifs which define nonbinding exceptions) or simply fail to fit. Nonbinding motifs are determined based on concentration dependent response when available or the average binding of non-motif glycans otherwise.

### **Boxplot of Fine-Specificity Motifs**





### Motif Intensity Map



Top Ordered Glycans

Figure 2. Glycan intensity and motif distribution plot. The top half of the plot presents the observed glycan binding intensity of various glycans used in the array over their rank binding intensity; only the top glycans are shown. The second plot indicates the position of glycans containing the various motifs in the top plot with a yellow tick.

### Motif Family Membership Map

A	Relative Binding
# Family Sub-Family # Member *Family Remainder	

\*Motif indicates the remaining glycans not matched by motifs which are a subset. Motif definition needs to be taken in the context of the model.

Figure 3. Treemap of glycan binding grouped by motif and family structure. The model structure can be represented as nested boxes where box size is proportional to the number of glycans with the motif and color changes with change in average relative binding of glycans with the motif. Only three layers of data splitting are included here, though further splitting may be possible.

### Detailed Model Breakdown

Motif Glycan Examples:



10/16/2020

NR42486





All Concentration Plots:



Figure 4. Boxplots of glycan binding grouped by motif for each dataset in the model. Motifs are listed in ascending average binding intensity (for the selected concentration) and colored by family.

Model Structure:



\*Motif matches the remaining glycans not matched by earlier motifs in the model.

Figure 5. Tree representation of the regression tree model trained on array data. Data flows through the tree (topdown) and is split by the various motifs. The motif used the split the data at each point has the id "family+split number" except when further split. In the case of further splits the id of the motif used to split the data is denoted with an asterisk.

#### Curve Fitting:

No curves were fit for model

Motif Text Structures:

Motif ID	Motif Graphic	c Motif Text
A0		<pre>&lt;6f8f&gt;Neu5AcA2-6&lt;2f3f4f&gt;GalB1-4&lt;3f6f&gt;GlcNAcB1-2&lt;3f4f6f&gt;ManA1-&lt;3or6&gt; &lt;2f4f&gt;ManB</pre>
0		Non-Binders
Key:		
Hexose 🔿	HexNAc 🗌 Kdn 🔶	Any Single
Glc 🔵	GlcNAc 📕 Neu5Ac 🔶	Any # of
Man 🔵	GalNAc 📃 Neu5Gc 🔶	Monosaccharide ~SKIP~
Gal 🔴	Fuc 🔺 Sia 🔶	Monosaccharide

## NR44365 (A1\_H7)

# Automated Analysis report for NR44365 evaluated at 20 ug/mL

### List of Primary NR44365 Motifs

Primary Motif ID	Primary Motif Minimal Structure	Primary Motif Complete Structure	Relative Binding	Number of Glycans
A	$\mathbf{A}^{\alpha 3} \mathbf{A}^{\beta} \mathbf{R}$	$\mathbf{R}$	1.00	40
В	$\mathbf{R}$	$\mathbf{R}^{\underline{2}} \underbrace{4/6}_{\beta 4} \underbrace{\beta 4}_{\beta 4} \beta $	0.49	8
0			0.01	24

Minimal and complete motif definitions match the same set of glycans defined in the array. Components found in the complete motif but excluded in the minimal motif are not tested by the array. Monosaccharide identities and subsitution intolerance may or may not have been tested by the array, they are retained from the complete motif for readability.

## **Boxplot of Primary Motifs**





### List of Fine-Specificity NR44365 Motifs

Motif ID	Nearest Common Name (Accuracy%**)	Motif Graphic Structure	Relative Binding	Number of Glycans	P-Value***
A2*	a3 Sialyl Type 2 LacNAc Bisecting N-Glycan (90%)		1.00	8	<0.001
A3		$ \begin{array}{c} & \alpha 3 \\ & \beta 4 \\ & \beta 2 \\ & \alpha \\ & \alpha \\ & \alpha \\ & \beta 2 \\ & \beta 2 \\ & \alpha \\ & \beta 2 \\ & $	0.92	4	<0.001
A0*		$\mathbf{A}^{\alpha  3} \mathbf{A}^{\beta} \mathbf{R}$	0.74	18	<0.001
A1		$ \begin{array}{c} & \alpha^{3} \\ & \beta^{4} \\ & \beta^{2} \\ & \beta^{4} \\ & \beta^{2} \\ & \beta^{2} \\ \end{array} $	0.72	4	<0.001

Nearest Common

NR44365

Motif IE	Name O (Accuracy%**)	Motif Graphic Structure	Relative Binding	Number of Glycans	P-Value***
A4	Sialyl Lewis X Bisecting N-Glycan (93%)	$ \begin{array}{c} & \alpha 3 \\ & \beta 4 \\ & \beta 2 \\ & \alpha \\ & \alpha \\ & \beta 4 \\ & \beta 2 \\ & \alpha \\ & \beta 4 \\ & \beta $	0.52	6	<0.001
B0		$\mathbf{R} = \mathbf{R} + $	0.37	8	<0.001
0	Non-Binders (100%)		0.00	24	NA
Key:					
Hexose 🔘	HexNAc 🗌 Kdn 🔶 🗚	ny Single			
Glc 🔵	GlcNAc Neu5Ac 🧄 🖉	osaccharide <u>/</u> Any # of			
Man 🔵	GalNAc 🔄 Neu5Gc 🧄 Mono	osaccharide ~SKIP~			
Gal 🔵	Fuc 🔺 Sia 🃥 Mon	orogen or R- osaccharide			

See Symbol Nomenclature for Glycans (SNFG) for complete key: https://www.ncbi.nlm.nih.gov/glycans/snfg.html (https://www.ncbi.nlm.nih.gov/glycans/snfg.html)

\*Motif indicates the remaining glycans not matched by motifs which are a subset. Motif definition needs to be taken in the context of the model.

\*\*Accuracy describes the consistency between common-name definition of the motif and the formal, text-based definition of the motif, in terms of percent agreement in the glycans containing the two motifs. Common Name label definitions given here (http://carbogrove.org/MotifLabels.php).

\*\*\*P-Value refers to difference from Non-Binders with multiple testing correction (Dunnet's Test)

Motifs with a red motif ID fail to show a logistic response to protein concentration in the range of concentrations analyzed. These motifs may be nonbinding motifs (motifs which define nonbinding exceptions) or simply fail to fit. Nonbinding motifs are determined based on concentration dependent response when available or the average binding of non-motif glycans otherwise.

### **Boxplot of Fine-Specificity Motifs**



Figure 1. Glycan binding grouped by motif and motif family. Individual glycans are given as points on the plot.

### Motif Intensity Map



Figure 2. Glycan intensity and motif distribution plot. The top half of the plot presents the observed glycan binding intensity of various glycans used in the array over their rank binding intensity; only the top glycans are shown. The second plot indicates the position of glycans containing the various motifs in the top plot with a yellow tick.

### Motif Family Membership Map



\*Motif indicates the remaining glycans not matched by motifs which are a subset. Motif definition needs to be taken in the context of the model.

Figure 3. Treemap of glycan binding grouped by motif and family structure. The model structure can be represented as nested boxes where box size is proportional to the number of glycans with the motif and color changes with change in average relative binding of glycans with the motif. Only three layers of data splitting are included here, though further splitting may be possible.

### **Detailed Model Breakdown**

Motif Glycan Examples:









Figure 4. Boxplots of glycan binding grouped by motif for each dataset in the model. Motifs are listed in ascending average binding intensity (for the selected concentration) and colored by family.

Model Structure:



\*Motif matches the remaining glycans not matched by earlier motifs in the model.

Figure 5. Tree representation of the regression tree model trained on array data. Data flows through the tree (topdown) and is split by the various motifs. The motif used the split the data at each point has the id "family+split number" except when further split. In the case of futher splits the id of the motif used to split the data is denoted with an asterisk.

### Curve Fitting:

No curves were fit for model

Motif Text Structures:

### Motif

ID	Motif Graphic	Motif Text
A1	$ \begin{array}{c} & \overset{\alpha 3}{\longrightarrow} \overset{\beta 4}{\longrightarrow} \overset{\beta 2}{\longrightarrow} \overset{\alpha _{6}}{\longrightarrow} \overset{\alpha _{6}}{\longrightarrow} \overset{\alpha _{6}}{\longrightarrow} \overset{\alpha _{6}}{\longrightarrow} \overset{\alpha _{7}}{\longrightarrow} \alpha $	<6f8f>Neu5AcA2-3<2f4f6f>GalB1-4<3f6f>GlcNAcB1-2<3f4f6f>ManA1- 6(<3f6f>GlcNAcB1-2<3f4f6f>ManA1-3)<2f4f>ManB
A2*		<6f8f>Neu5AcA2-3<2f4f6f>GalB1-4<3f6f>GlcNAcB1-2<3f4f6f>ManA1-6<2f>ManB
A3	$ \begin{array}{c} \bullet^{\alpha 3} \bullet^{\beta 4} \blacksquare^{\beta 2} \bullet^{\alpha} \\ \stackrel{a}{\overset{a}{\overset{a}{\overset{a}{\overset{b}{\overset{a}{\overset{b}{\overset{b}{\overset$	<6f8f>Neu5AcA2-3<2f4f6f>GalB1-4(<2f3f4f>FucA1-3)<6f>GlcNAcB1-2<3f4f6f>ManA1- 6(<3f6f>GlcNAcB1-2<3f4f6f>ManA1-3)<2f4f>ManB
A4		<6f8f>Neu5AcA2-3<2f4f6f>GalB1-4(<2f3f4f>FucA1-3)<6f>GlcNAcB1-2<3f4f6f>ManA1- 6(<3f4f6f>GlcNAcB1-4)<2f>ManB



# RatDCIR2\_DeepSearch2

## Automated Analysis report for RatDCIR2\_DeepSearch2 evaluated at 20 ug/mL

### List of Primary RatDCIR2\_DeepSearch2 Motifs

Primary Motif ID	Primary Motif Minimal Structure	Primary Motif Complete Structure	Relative Binding	Number of Glycans
A	$\mathbf{R} \xrightarrow{\mathbf{A}}_{6} \xrightarrow{\beta} \mathbf{R}$	$\mathbf{R}$	1	12
0			0	60

Minimal and complete motif definitions match the same set of glycans defined in the array. Components found in the complete motif but excluded in the minimal motif are not tested by the array. Monosaccharide identities and subsitution intolerance may or may not have been tested by the array, they are retained from the complete motif for readability.

### **Boxplot of Primary Motifs**

11/19/2020



### List of Fine-Specificity RatDCIR2\_DeepSearch2 Motifs

Motif ID	Nearest Common Name (Accuracy%**)	Motif Graphic Structure	Relative Binding	Number of Glycans	P- Value***
A3*		$\mathbf{R}$	1.00	1	<0.001
A2	Terminal N- Glycan N-Acetyl Glucosamine Terminal Type 2 LacNAc (94%)	$\beta = \beta = \alpha$	0.57	2	<0.001
A1	Terminal N- Glycan N-Acetyl Glucosamine 5- Acetyl Neuramic Acid (92%)	$\mathbf{R}^{\mathbf{d}/6} \mathbf{R}^{\mathbf{d}} \mathbf{R}^{\mathbf{d}}$	0.11	6	<0.001

RatDCIR2\_DeepSearch2

Motif ID	Nearest Common Name (Accuracy%**)	Motif Graphic Structure	Relative Binding	Number of Glycans	P- Value***
A0*	Terminal N- Glycan N-Acetyl Glucosamine Terminal beta- Galactose (90%)	$\mathbf{R}$	0.00	3	1.000
0	Non-Binders (100%)		0.00	60	NA
Key:					
Hexose HexNAc Gic GicNAc Man GalNAc Gal Fuc	<ul> <li>Kdn ◆</li> <li>Neu5Ac ◆</li> <li>Neu5Gc ◆</li> <li>Sia ◆</li> <li>Any Single Monosaccha Any # of</li> <li>Monosacchar Hydrogen c</li> </ul>	e ride 〜 ide 〜SKIP〜 or R- ide R-			

See Symbol Nomenclature for Glycans (SNFG) for complete key: https://www.ncbi.nlm.nih.gov/glycans/snfg.html (https://www.ncbi.nlm.nih.gov/glycans/snfg.html)

\*Motif indicates the remaining glycans not matched by motifs which are a subset. Motif definition needs to be taken in the context of the model.

\*\*Accuracy describes the consistency between common-name definition of the motif and the formal, text-based definition of the motif, in terms of percent agreement in the glycans containing the two motifs. Common Name label definitions given here (http://carbogrove.org/MotifLabels.php).

\*\*\*P-Value refers to difference from Non-Binders with multiple testing correction (Dunnet's Test)

Motifs with a red motif ID fail to show a logistic response to protein concentration in the range of concentrations analyzed. These motifs may be nonbinding motifs (motifs which define nonbinding exceptions) or simply fail to fit. Nonbinding motifs are determined based on concentration dependent response when available or the average binding of non-motif glycans otherwise.

### **Boxplot of Fine-Specificity Motifs**

11/19/2020



Figure 1. Glycan binding grouped by motif and motif family. Individual glycans are given as points on the plot.

### Motif Intensity Map



**Top Ordered Glycans** 

Figure 2. Glycan intensity and motif distribution plot. The top half of the plot presents the observed glycan binding intensity of various glycans used in the array over their rank binding intensity; only the top glycans are shown. The second plot indicates the position of glycans containing the various motifs in the top plot with a yellow tick.

### Motif Family Membership Map



\*Motif indicates the remaining glycans not matched by motifs which are a subset. Motif definition needs to be taken in the context of the model.

Figure 3. Treemap of glycan binding grouped by motif and family structure. The model structure can be represented as nested boxes where box size is proportional to the number of glycans with the motif and color changes with change in average relative binding of glycans with the motif. Only three layers of data splitting are included here, though further splitting may be possible.

### Detailed Model Breakdown

Motif Glycan Examples:






Figure 4. Boxplots of glycan binding grouped by motif for each dataset in the model. Motifs are listed in ascending average binding intensity (for the selected concentration) and colored by family.

Model Structure:



\*Motif matches the remaining glycans not matched by earlier motifs in the model.

Figure 5. Tree representation of the regression tree model trained on array data. Data flows through the tree (topdown) and is split by the various motifs. The motif used the split the data at each point has the id "family+split number" except when further split. In the case of futher splits the id of the motif used to split the data is denoted with an asterisk.

## Curve Fitting:

No curves were fit for model

Motif Text Structures:

Motif ID	Motif Graphic	Motif Text
A1	$\mathbf{R}^{\mathbf{G}/6}$	<6f8f>Neu5AcA2-<3or6><2f4f>GalB1-4<6f>GlcNAcB1-2<3f4f6f>ManA1- 6(<3f4f6f>GlcNAcB1-2<3f4f6f>ManA1-3)<2f>ManB
A2	$R^{\beta 4}$	<2f3f4f6f>GalB1-4<6f>GlcNAcB1-2<3f4f6f>ManA1-6(<3f4f6f>GlcNAcB1- 2<3f4f6f>ManA1-3)(<3f4f6f>GlcNAcB1-4)<2f>ManB

