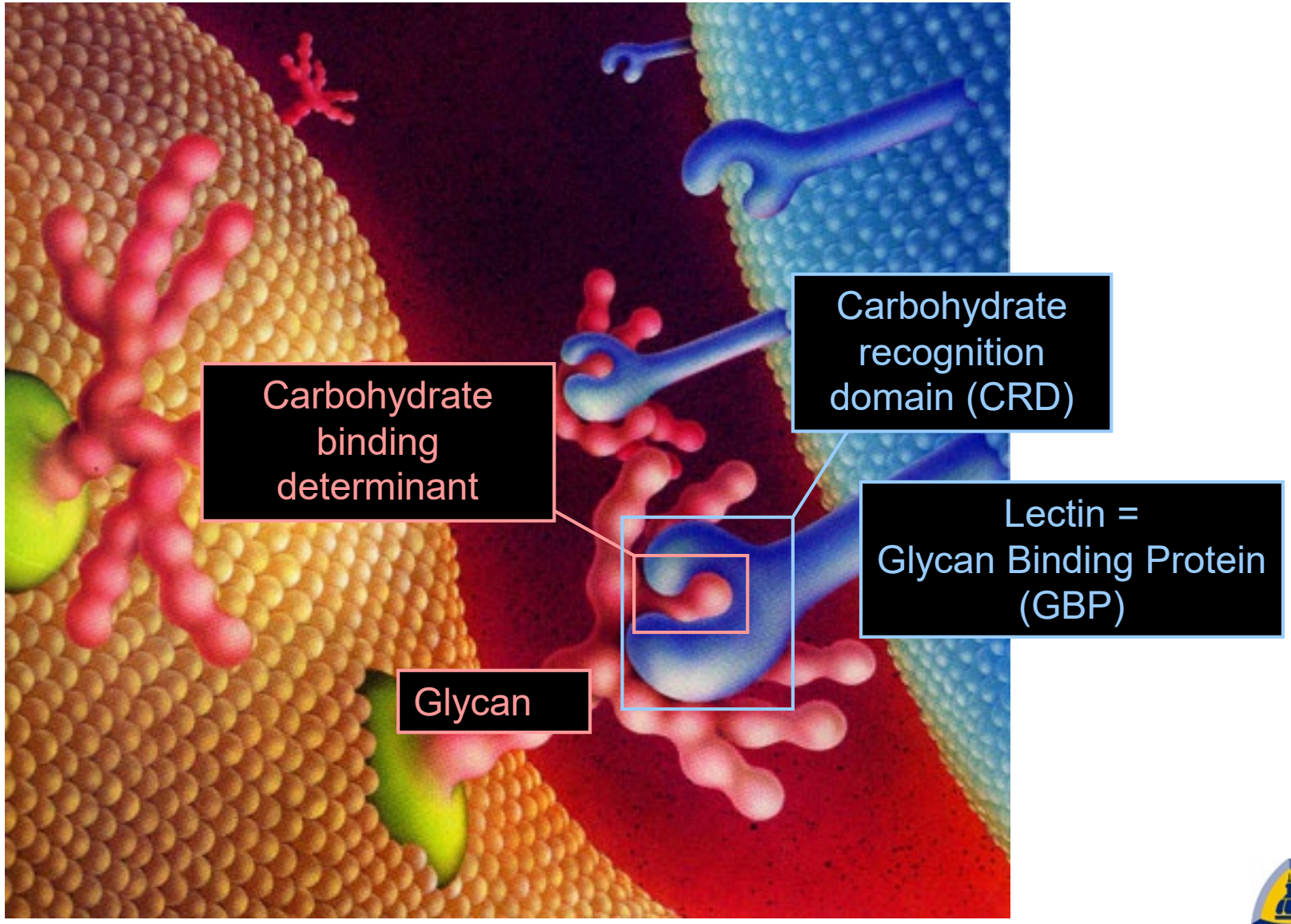
Protein-glycan recognition



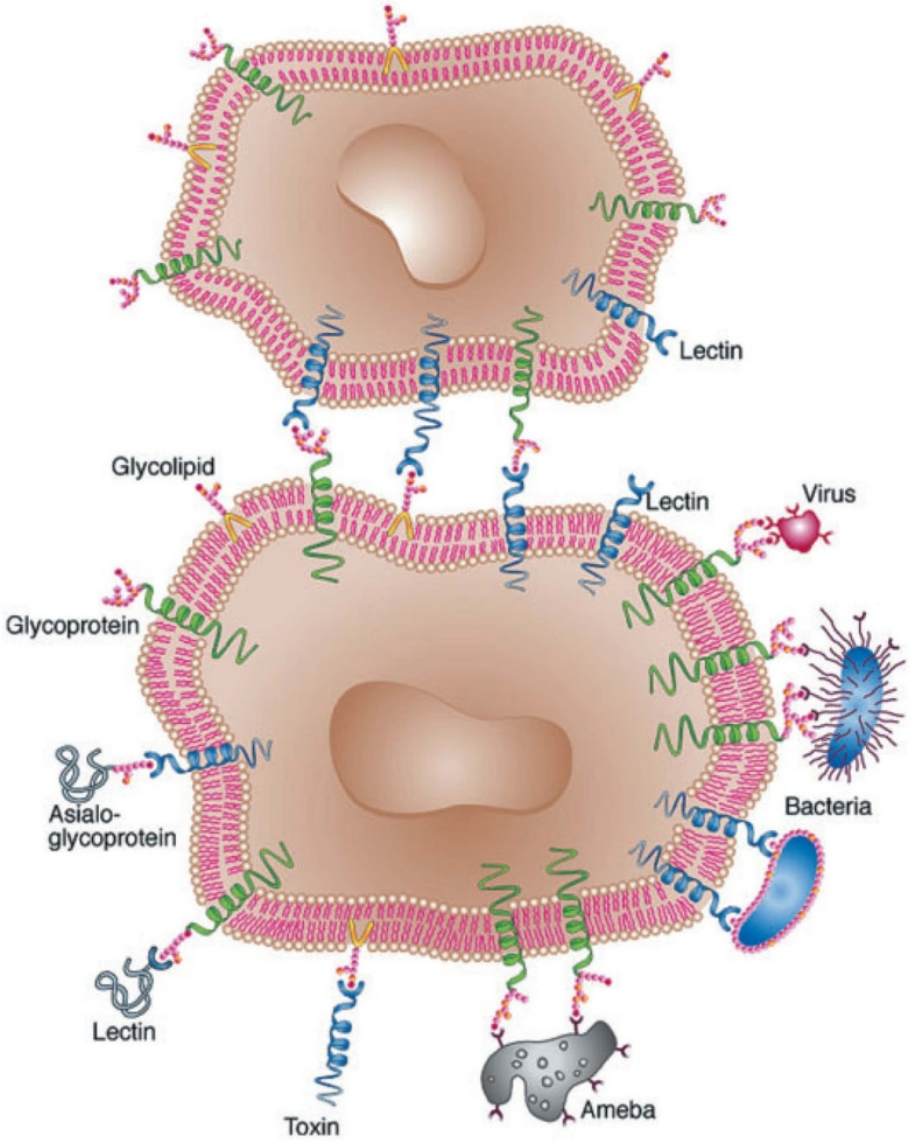
Sharon and Lis (1993) Scientific American

JOHNS HOPKINS MEDICINE

Objectives

- **Learn the diversity and families of glycan binding proteins**
- Learn the molecular/structural strategies used in protein-glycan binding and recognition
- Learn the methods and resources for determining glycan binding protein specificities
- Learn about plant lectins as tools in glycobiology research
- Learn about glycan binding strategies used by pathogens

Glycan binding protein diversity – phylogeny & function



Lectin	Role in
Microorganisms	
Amoeba	Infection
Bacteria	Infection
Influenza virus	Infection
Plants	
Various	Defense
Legumes	Symbiosis with nitrogen-fixing bacteria
Animals	
Calnexin, calreticulin, ERGIC-53	Control of glycoprotein biosynthesis
Collectins	Innate immunity
Dectin-1	Innate immunity
Galectins	Regulation of cell growth and apoptosis; regulation of the cell cycle; modulation of cell–cell and cell–substratum interactions
Macrophage mannose receptor	Innate immunity; clearance of sulfated glycoprotein hormones
Man-6-P receptors	Targeting of lysosomal enzymes
L-selectin	Lymphocyte homing
E- and P-selectins	Leukocyte trafficking to sites of inflammation
Siglecs	Cell-cell interactions in the immune and neural system
Spermadhesin	Sperm-egg interaction

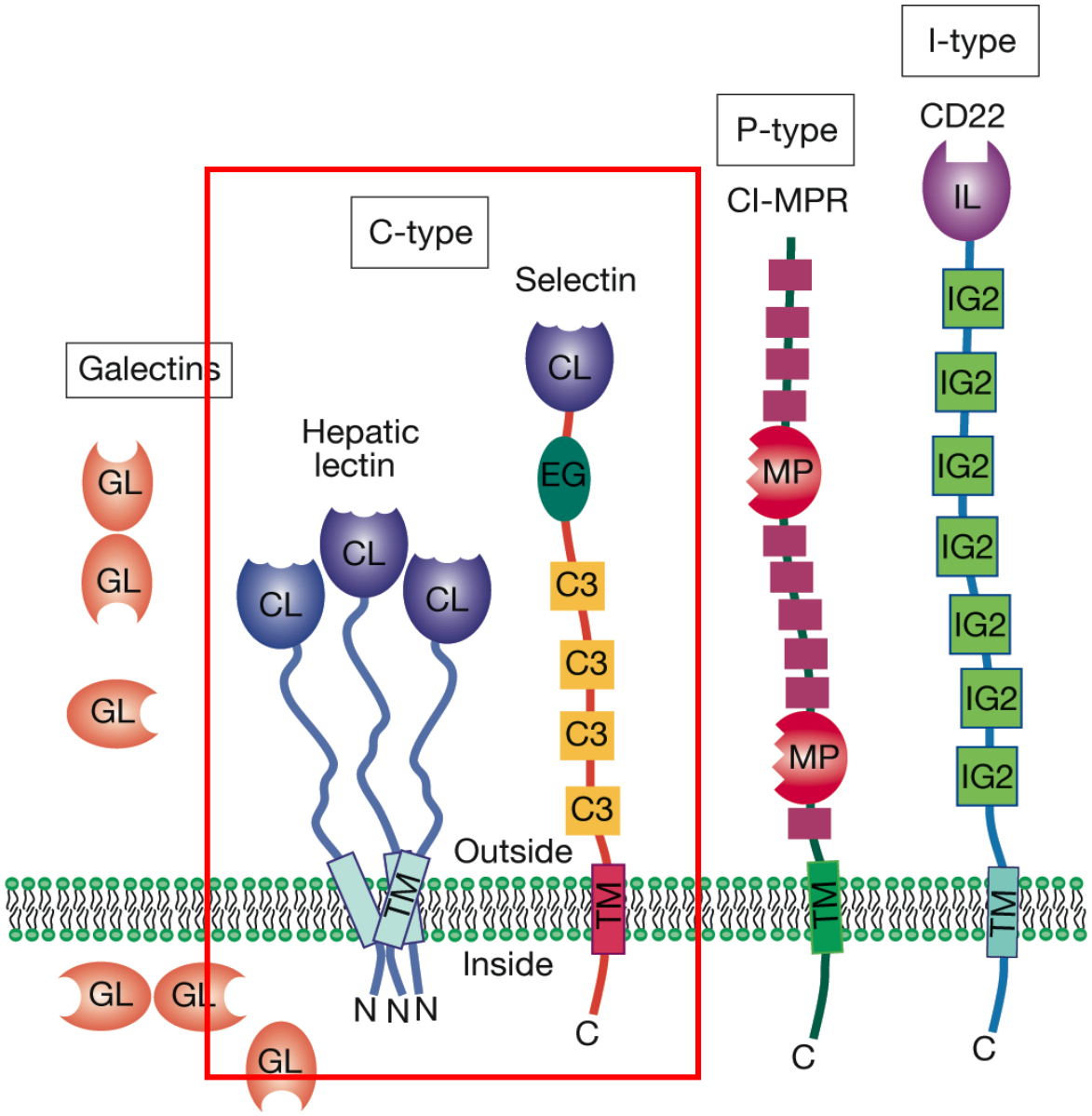
Sharon and Lis (2004) *Glycobiology* 14:53R

>80 mammalian lectins in >12 distinct families

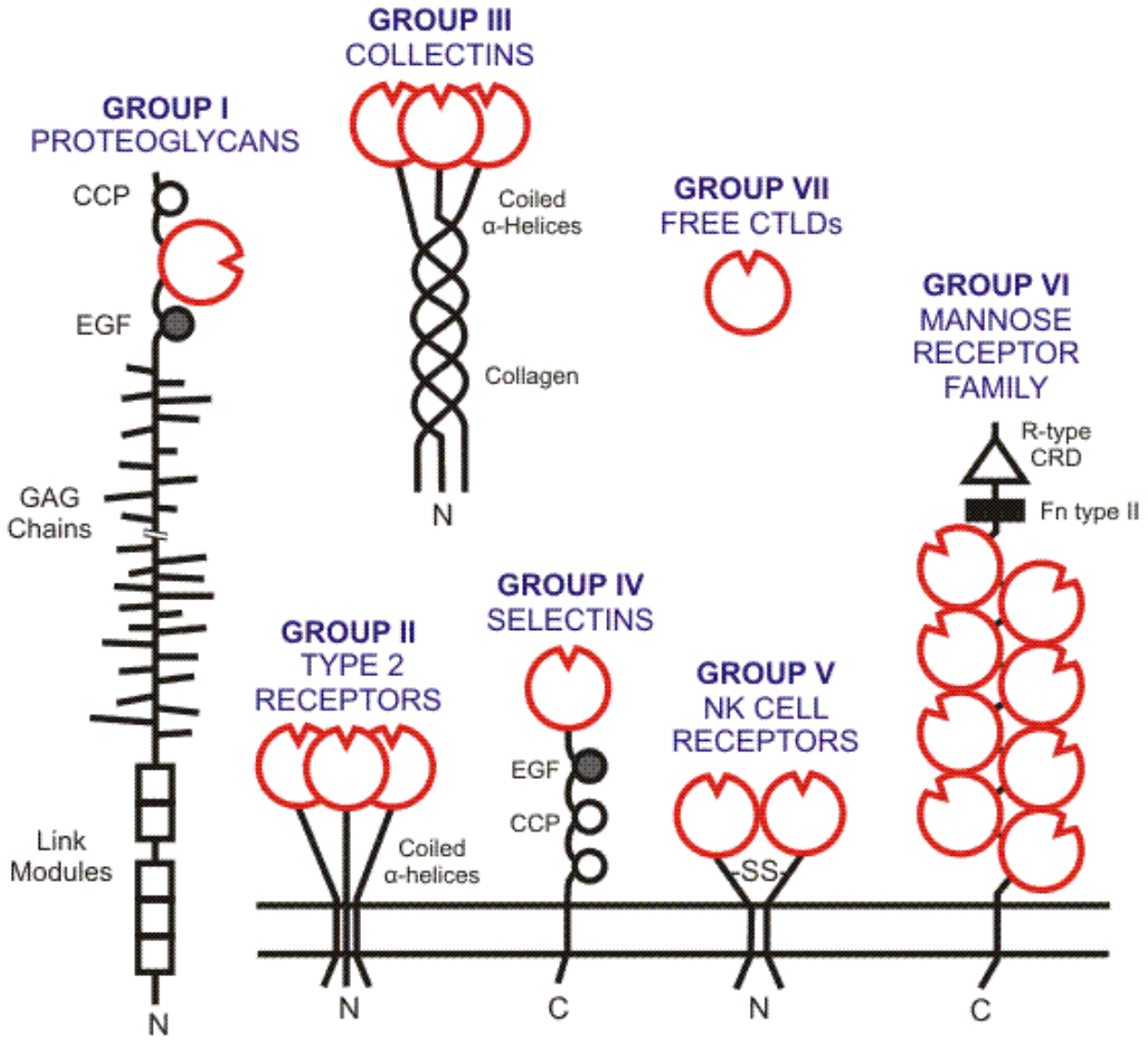
Lectin family	Saccharide ligands	Subcellular location	Examples of functions
Calnexin	Glc1Man9	ER	Protein sorting in the ER
M-type lectins	Man8	ER	ER -associated degradation
L-type lectins	various	ER, ERGIC, Golgi	Protein sorting in the ER
P-type lectins	Man 6-phosphate, others	Secretory pathway	post-Golgi glycoprotein trafficking
C-type lectins	Various, calcium-dependent	Cell membrane, extracellular	Cell adhesion (selectins), glycoprotein clearance, innate immunity (collectins).
Galectins	β-galactosides	Cytoplasm, extracellular	cell surface crosslinking
Siglecs (I-type lectins)	sialic acid	Cell membrane	molecular & cell recognition
R-type lectins	various	Golgi, Cell membrane	Enzyme targeting, hormone turnover
F-box lectins	GlcNAc2	Cytoplasm	degradation of misfolded proteins.
Ficolins	GlcNAc, GalNAc	cell membrane, extracellular	Innate immunity.
Chitinase-like lectins	chito-oligosaccharides	Extracellular	Collagen metabolism
F-type lectins	fucose termini	Extracellular	Innate immunity.
Intelectins	Gal, galactofuranose, pentoses	Extracellular/cell membrane	Innate immunity. Fertilization and embryogenesis.

Kurt Drickamer, Imperial College, http://www.imperial.ac.uk/research/animallectins/default.html

Mammalian lectin families

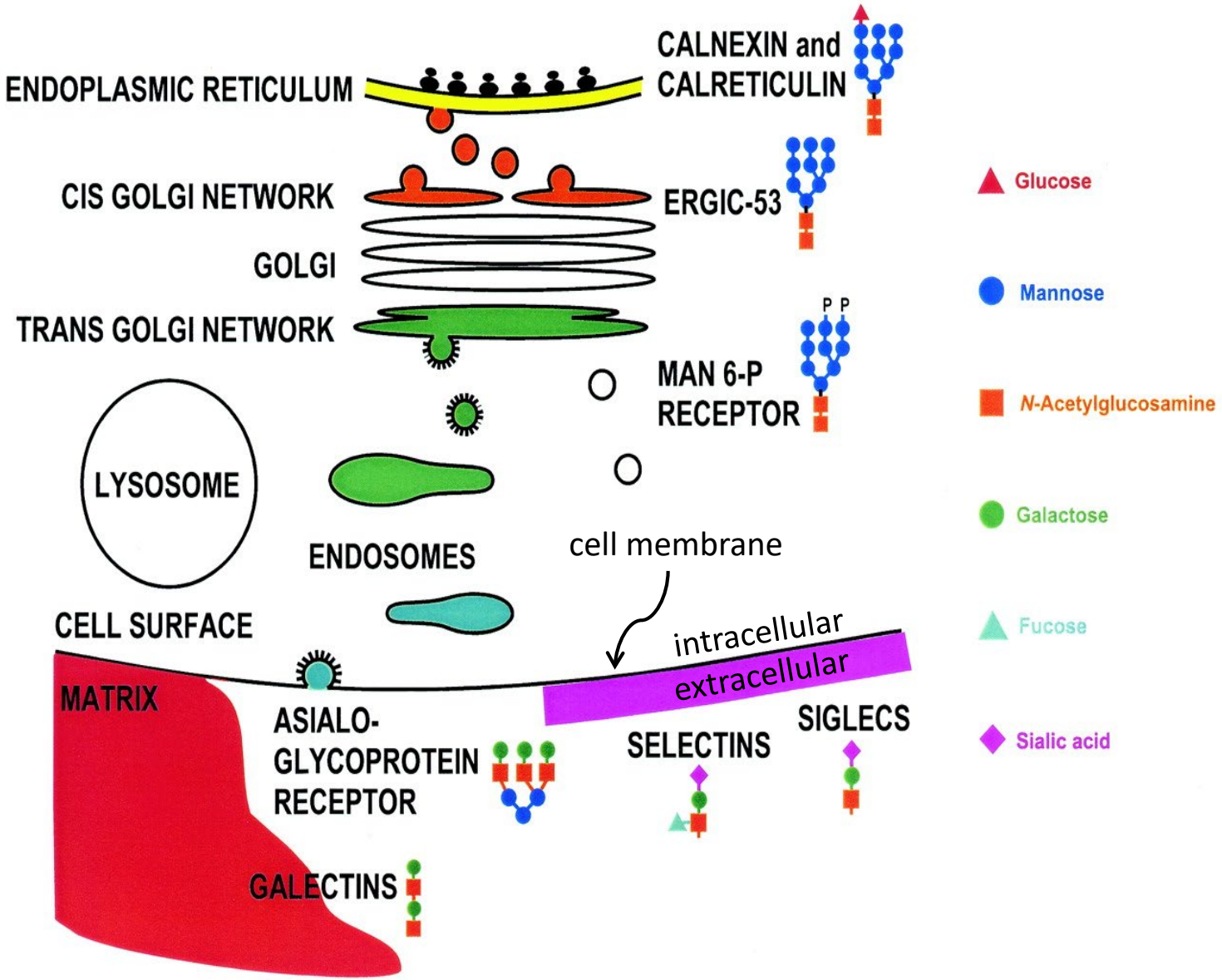


C-type lectins – diverse domain structures for diverse functions



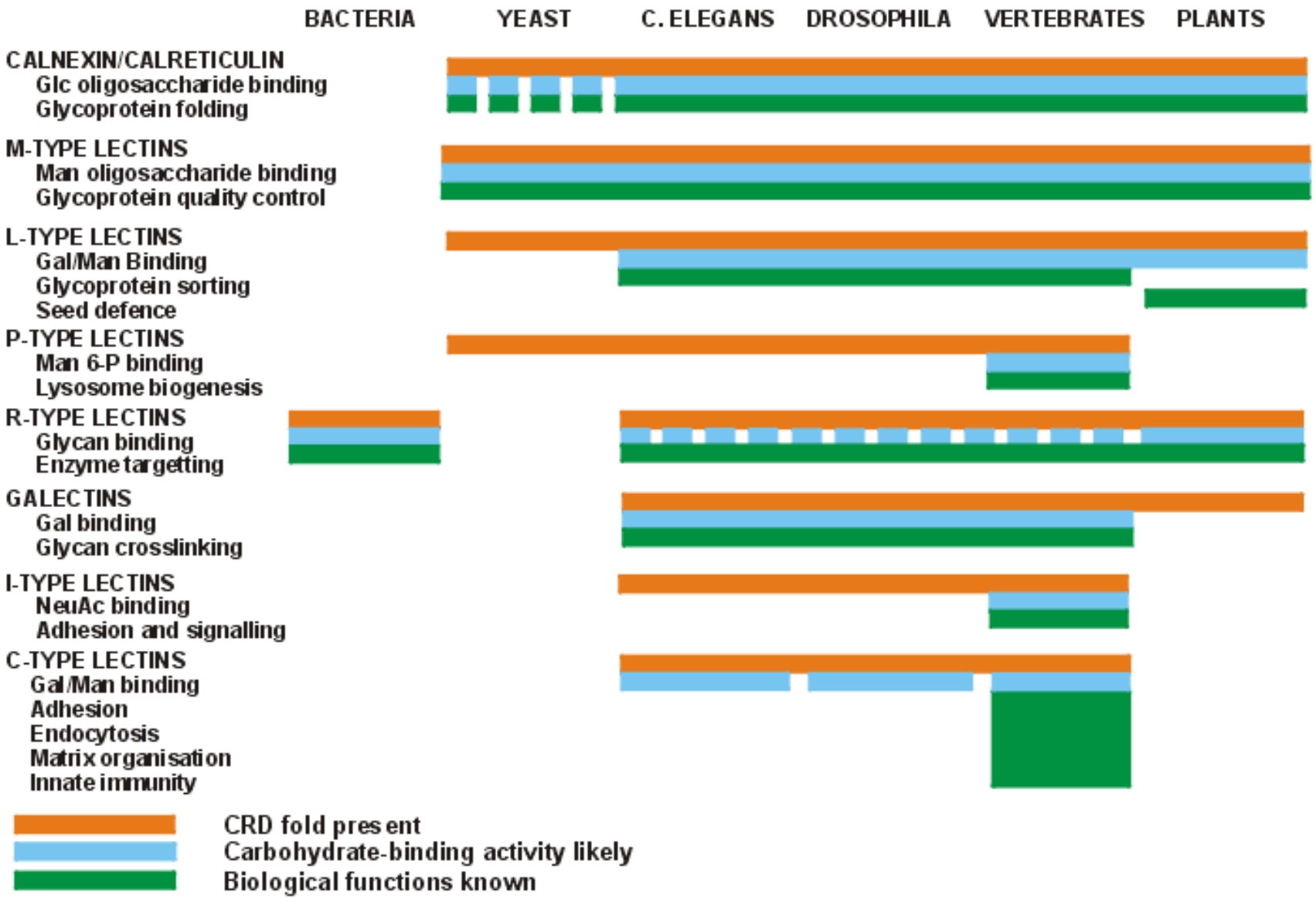
Kurt Drickamer, Imperial College London
http://www.imperial.ac.uk/research/animallectins

Animal lectins – cellular locations



Dodd, R.B. and
Drickamer, K. (2001)
Glycobiology **11**, 71R-79R

Lectin phylogeny



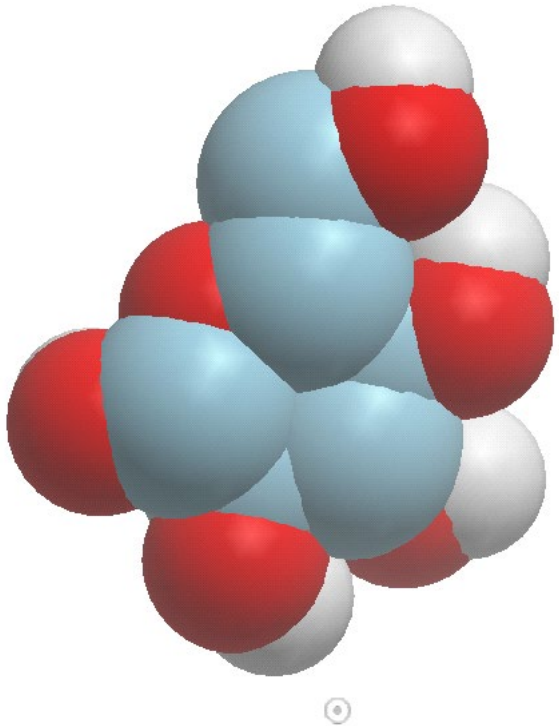
Kurt Drickamer, Imperial College,
http://www.imperial.ac.uk/research/animallectins/ctld/lectins.html

Objectives

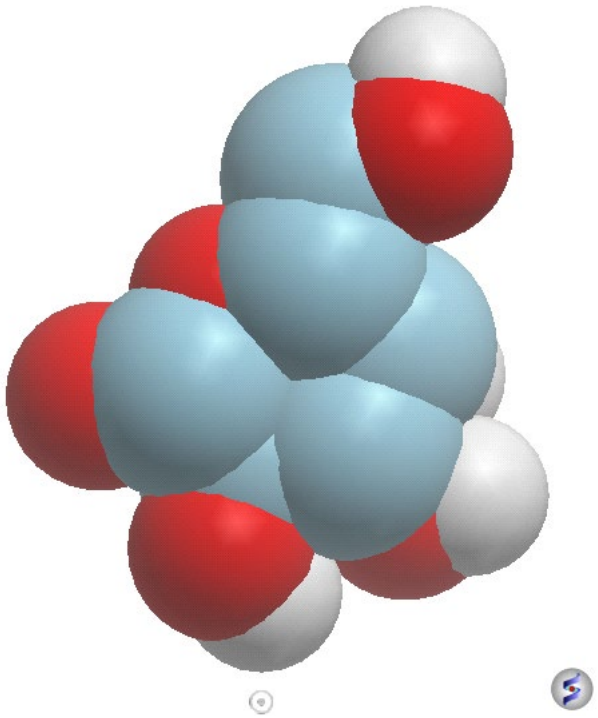
- Learn the diversity and families of glycan binding proteins
- Learn the molecular/structural strategies used in protein-glycan binding and recognition
- Learn the methods and resources for determining glycan binding protein specificities
- Learn about plant lectins as tools in glycobiology research
- Learn about glycan binding strategies used by pathogens

Sugars as language: Strategic placement of molecular recognition determinants

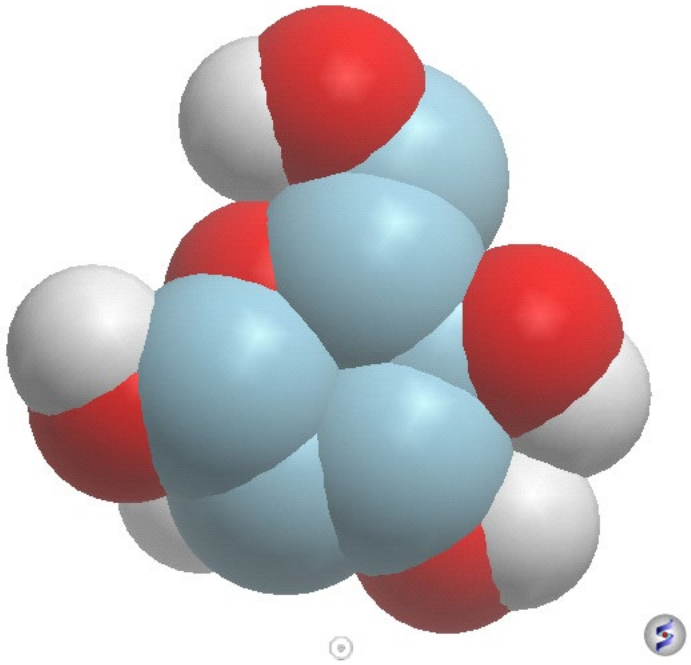
Glucose



Galactose

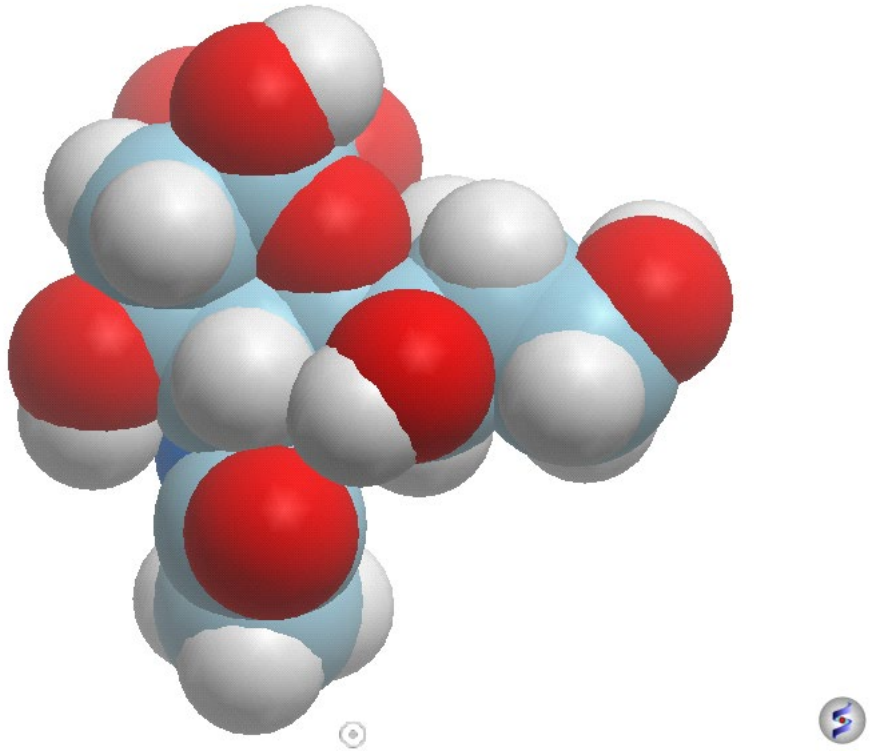


Mannose



Sugars as language: Strategic placement of molecular recognition determinants

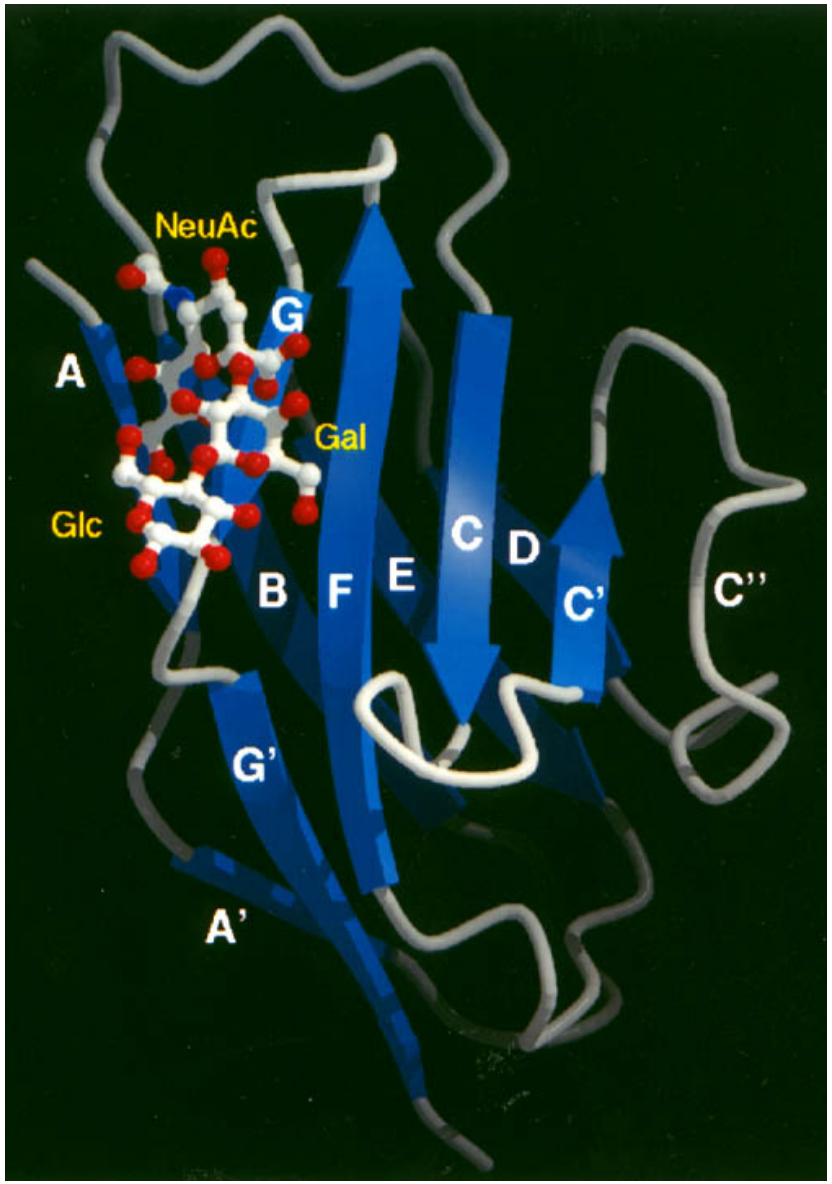
Sialic Acid (Neu5Ac)

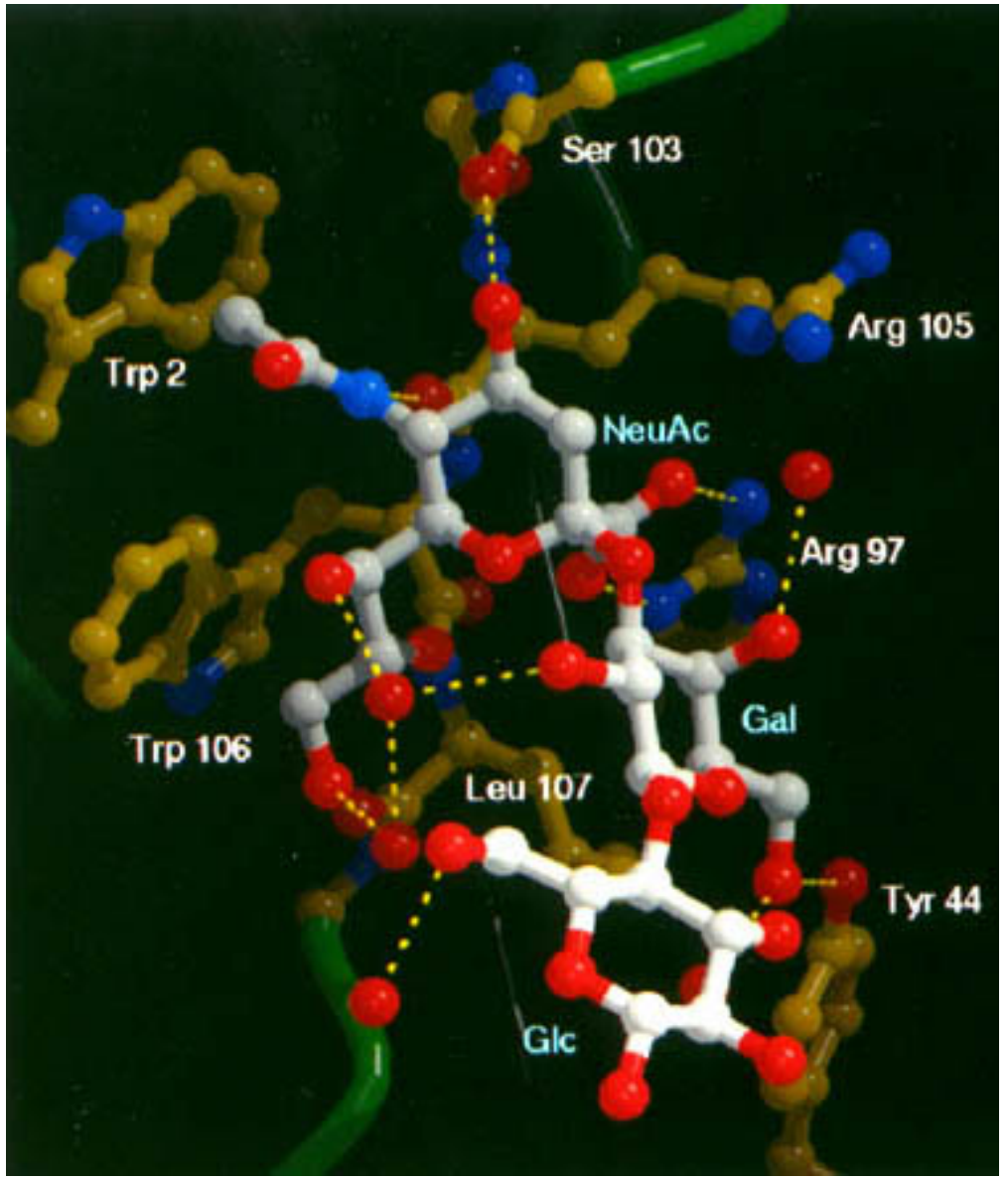


Principles of lectin-carbohydrate binding

- Precisely spaced cooperative hydrogen bonds
  (direct or through bound water molecules)
- Hydrophobic stacking
- Ionic interactions (for charged glycans)
- Calcium coordination (select lectins)
- Low site-affinity / high-avidity polyvalent binding

Example: Siglec-1

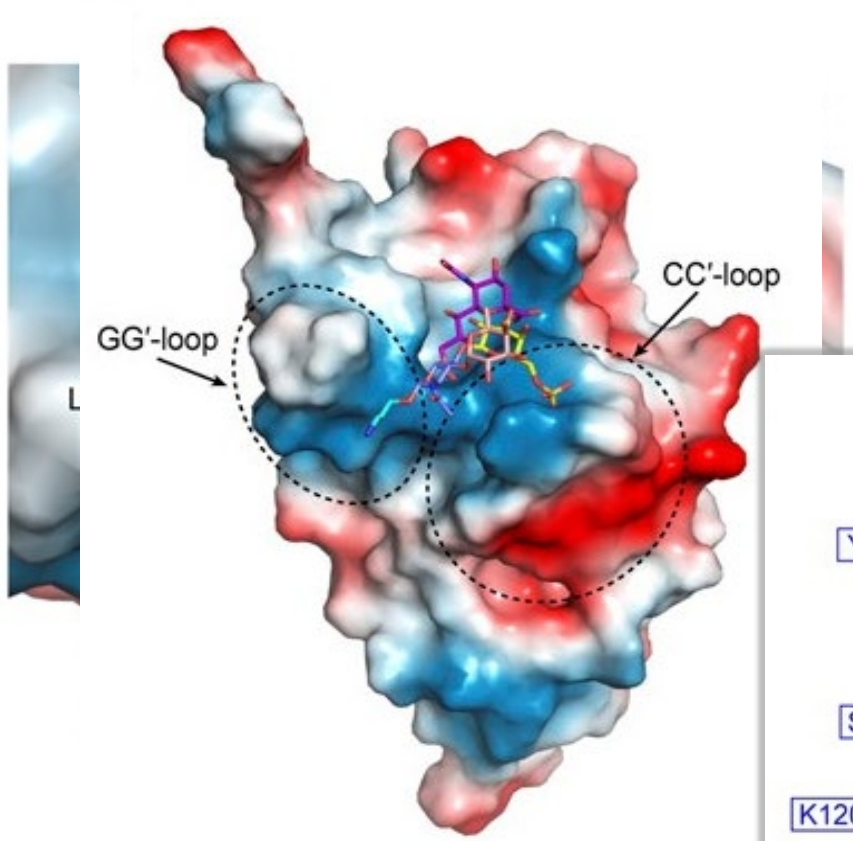




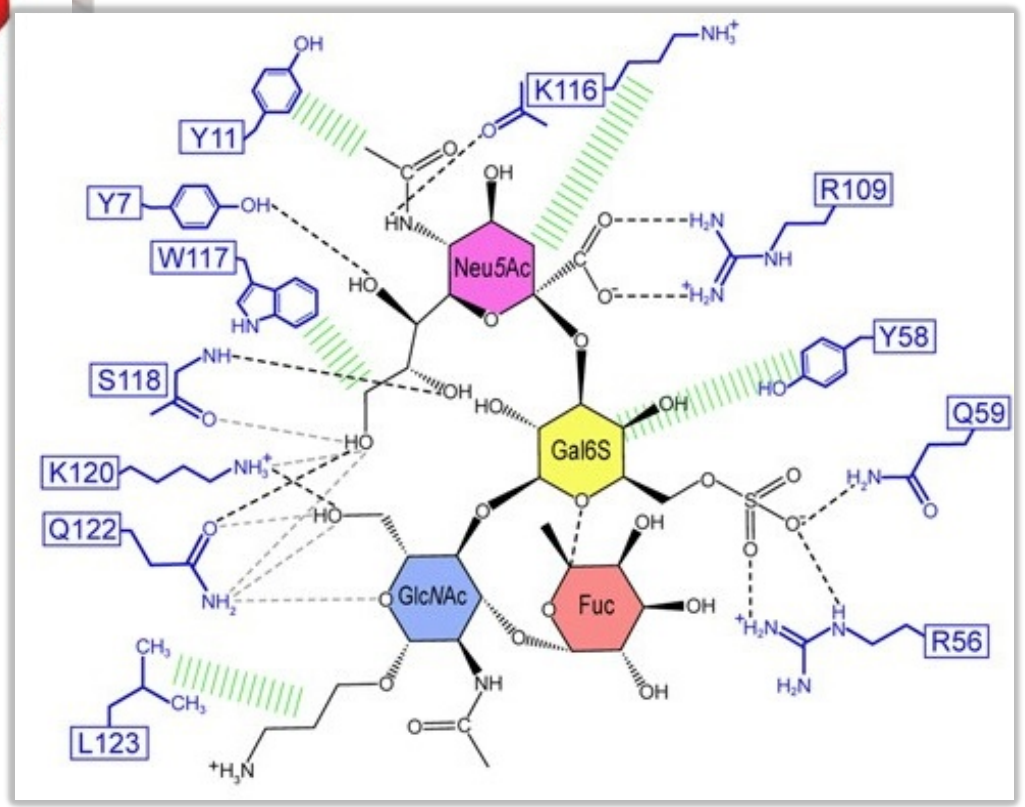
- **hydrogen bonds**
- **hydrophobic stacking**
- **ionic interactions**

May, A.P., et al. (1998)
Mol. Cell **1**, 719-728

Example: Siglec-8

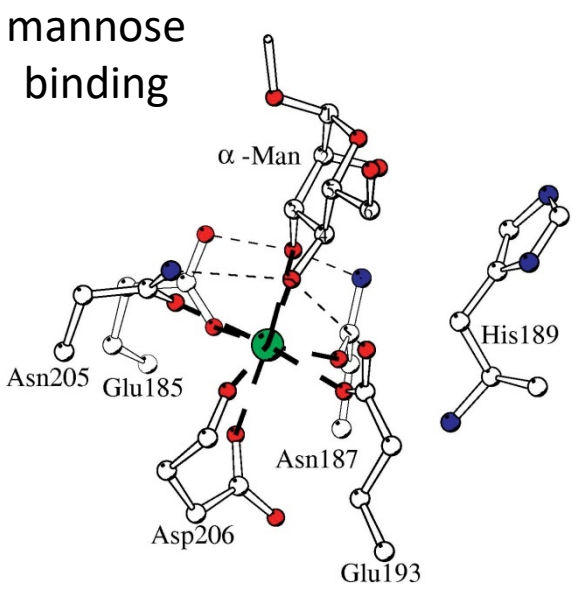


- **hydrogen bonds**
- **hydrophobic stacking**
- **ionic interactions**

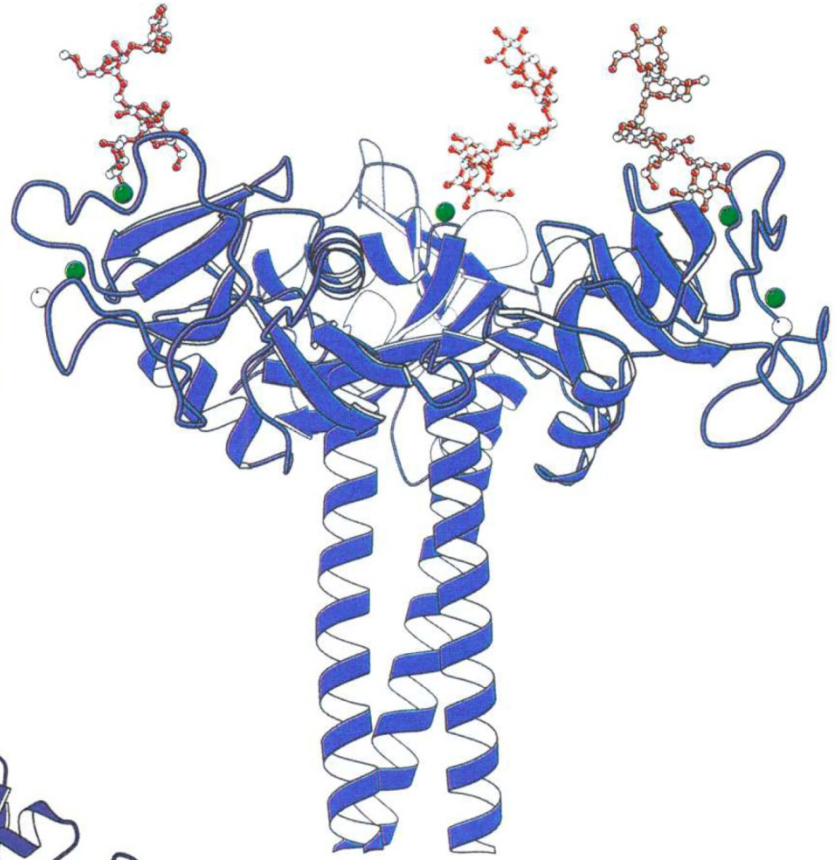


Pröpster et al. (2016) *PNAS* 113:E4170

Example: Mannose Binding Protein (MBP)



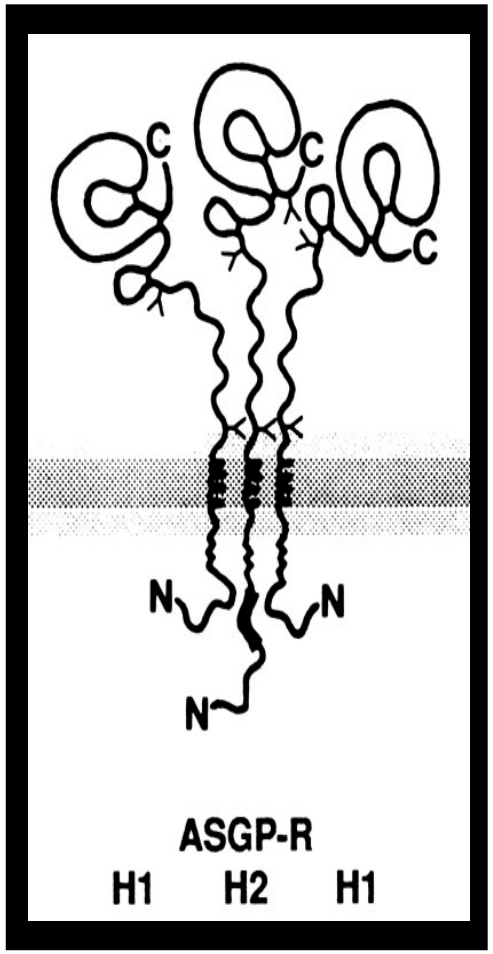
- **hydrogen bonds**
- **hydrophobic stacking**
- **calcium coordination**
- **multivalent binding**

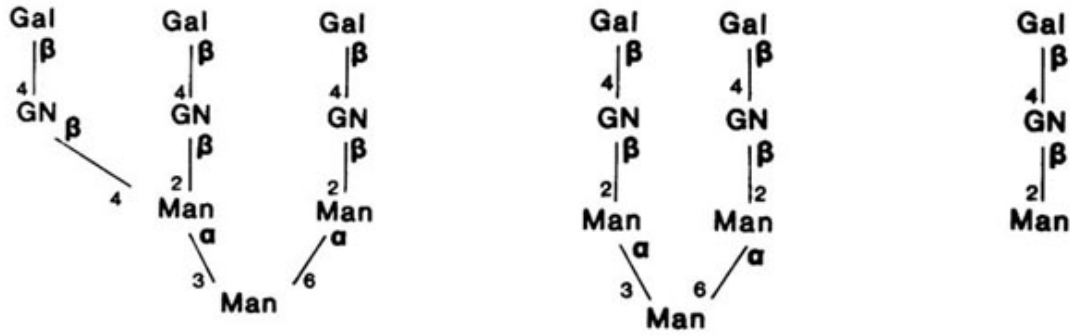


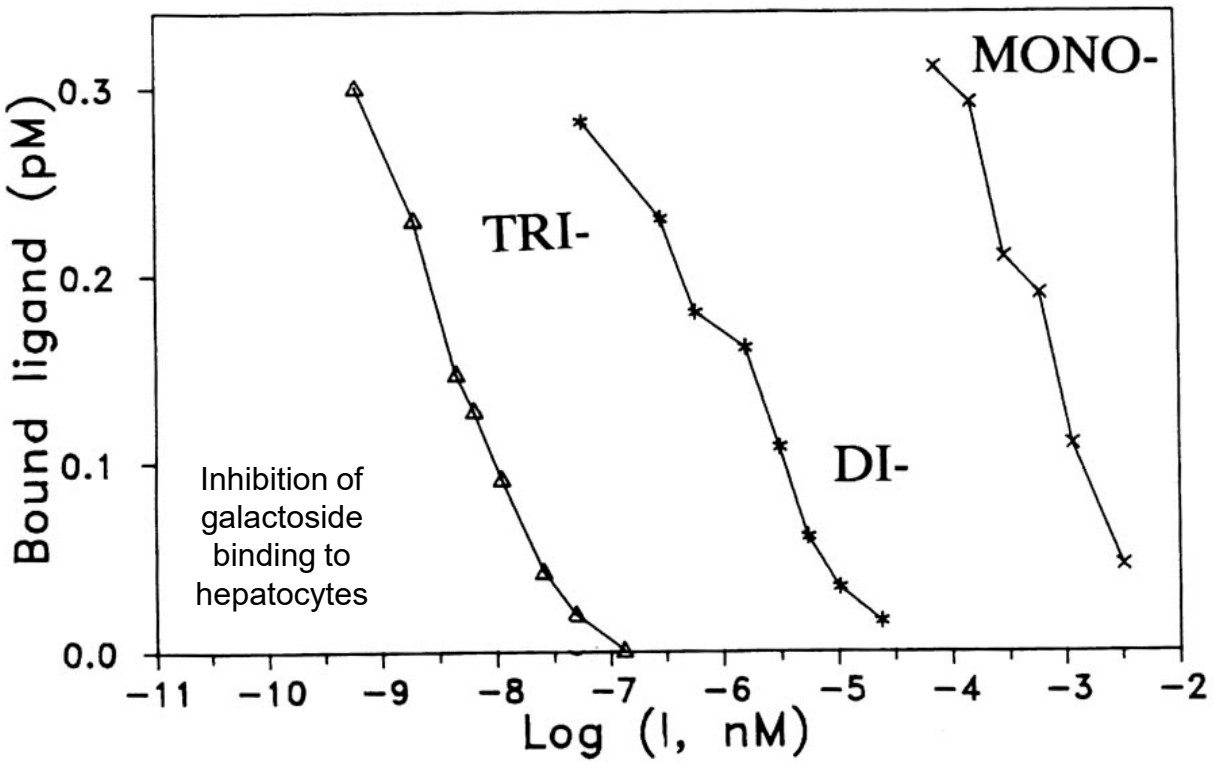
Weis & Drickamer (1994) *Structure* 2, 1227
Drickamer (1997) *Structure* 5, 465

Multivalent lectin binding – low site affinity, high multivalent avidity

The mammalian hepatic C-type lectin
A_2B trimer, each subunit with a galactose binding site





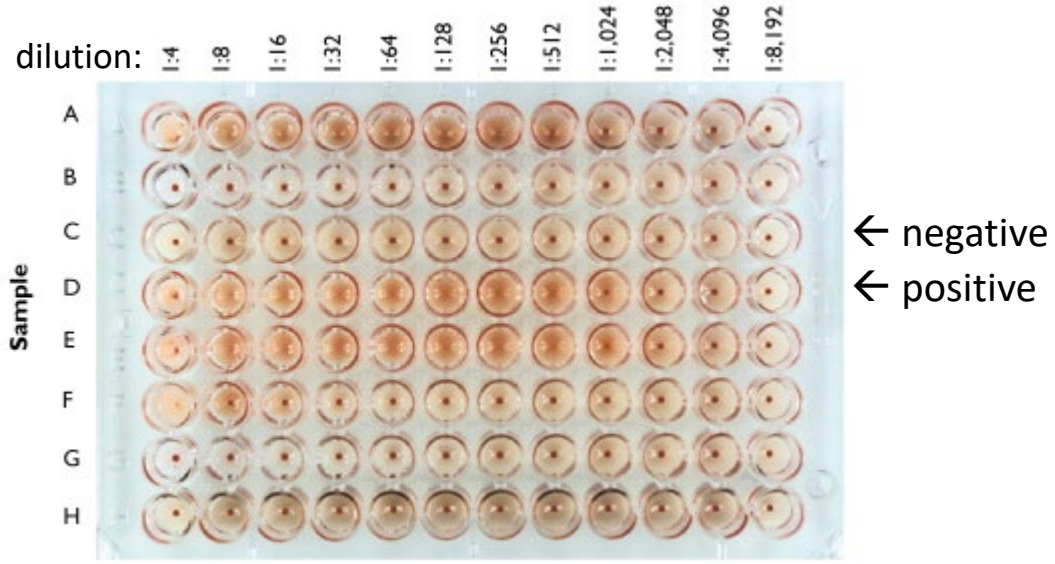


Lee et al (1983) *J Biol Chem* 258, 199

Objectives

- Learn the diversity and families of glycan binding proteins
- Learn the molecular/structural strategies used in protein-glycan binding and recognition
- Learn the methods and resources for determining glycan binding protein specificities
- Learn about plant lectins as tools in glycobiology research
- Learn about glycan binding strategies used by pathogens

Lectins were originally identified as "hemagglutinins"



← negative
← positive

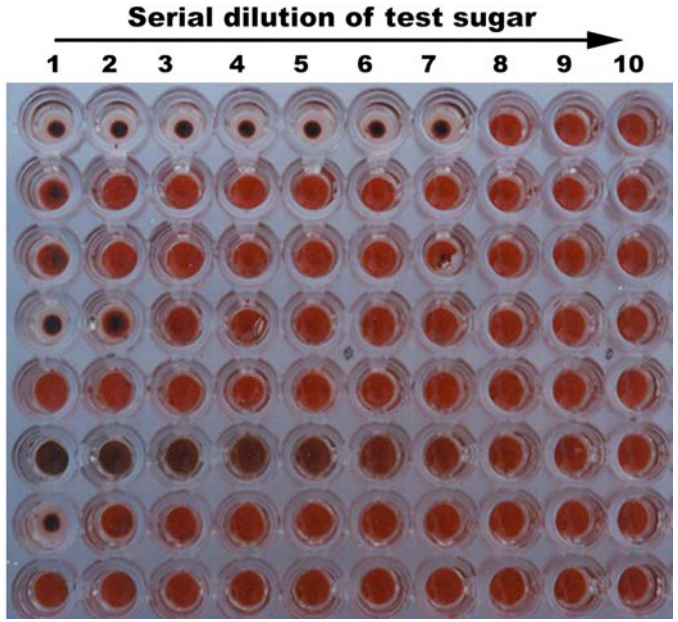
Influenza hemagglutinin: red cells fall to the bottom *unless* they are crosslinked.

http://www.virology.ws/2009/05/27/influenza-hemagglutination-inhibition-assay/

Sugar binding specificity: inhibition of galectin-induced hemagglutination by lactose

Ola et al (2007) *Cell Biology International* 31, 578

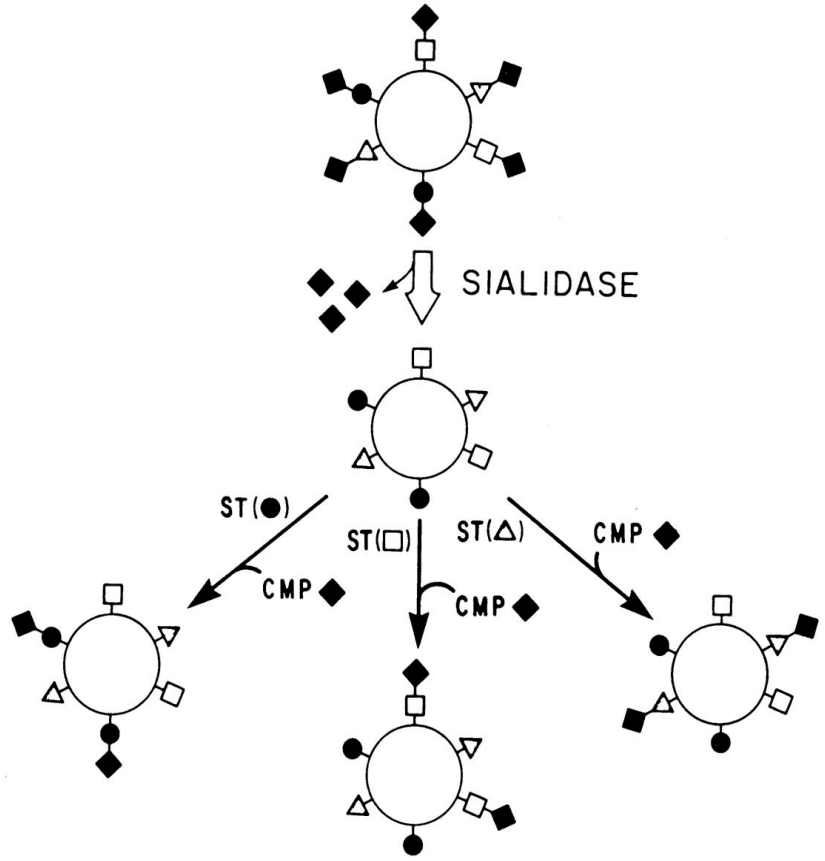
Carbohydrate	MIC (mM)
Lactose	0.8
Galactose	100
Methyl-β-D-galactopyranoside	100
p-nitrophenyl β-D-galactopyranoside	50
p-nitrophenyl α-D-galactopyranoside	NI
D-galactosamine	NI
Methyl-α-D-galactopyranoside	100
Glucose	NI



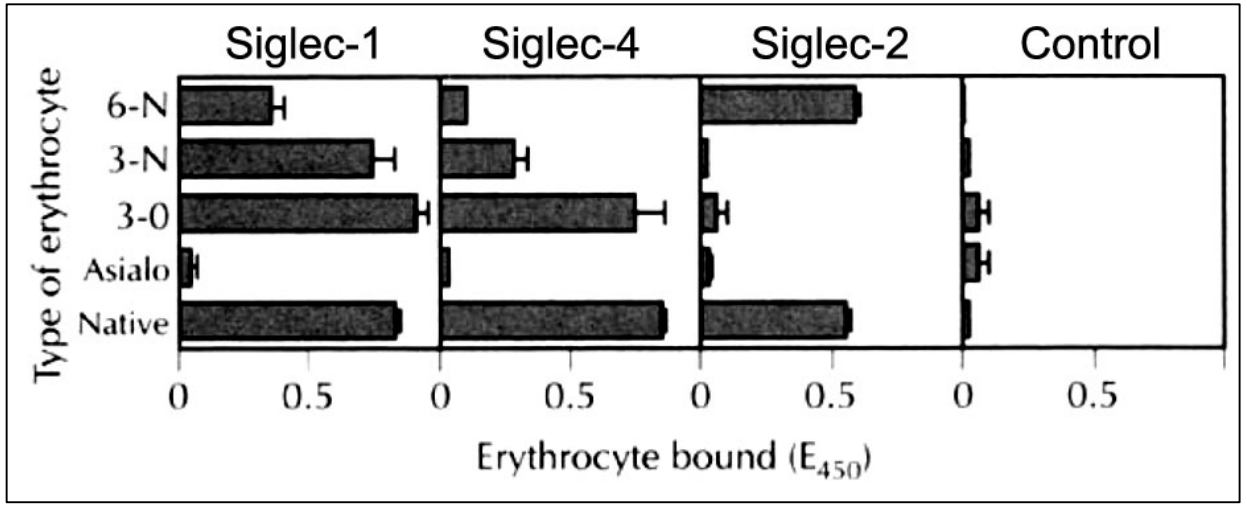
← inhibition

← no inhibition

Determine binding specificity by structurally-specific reconstitution

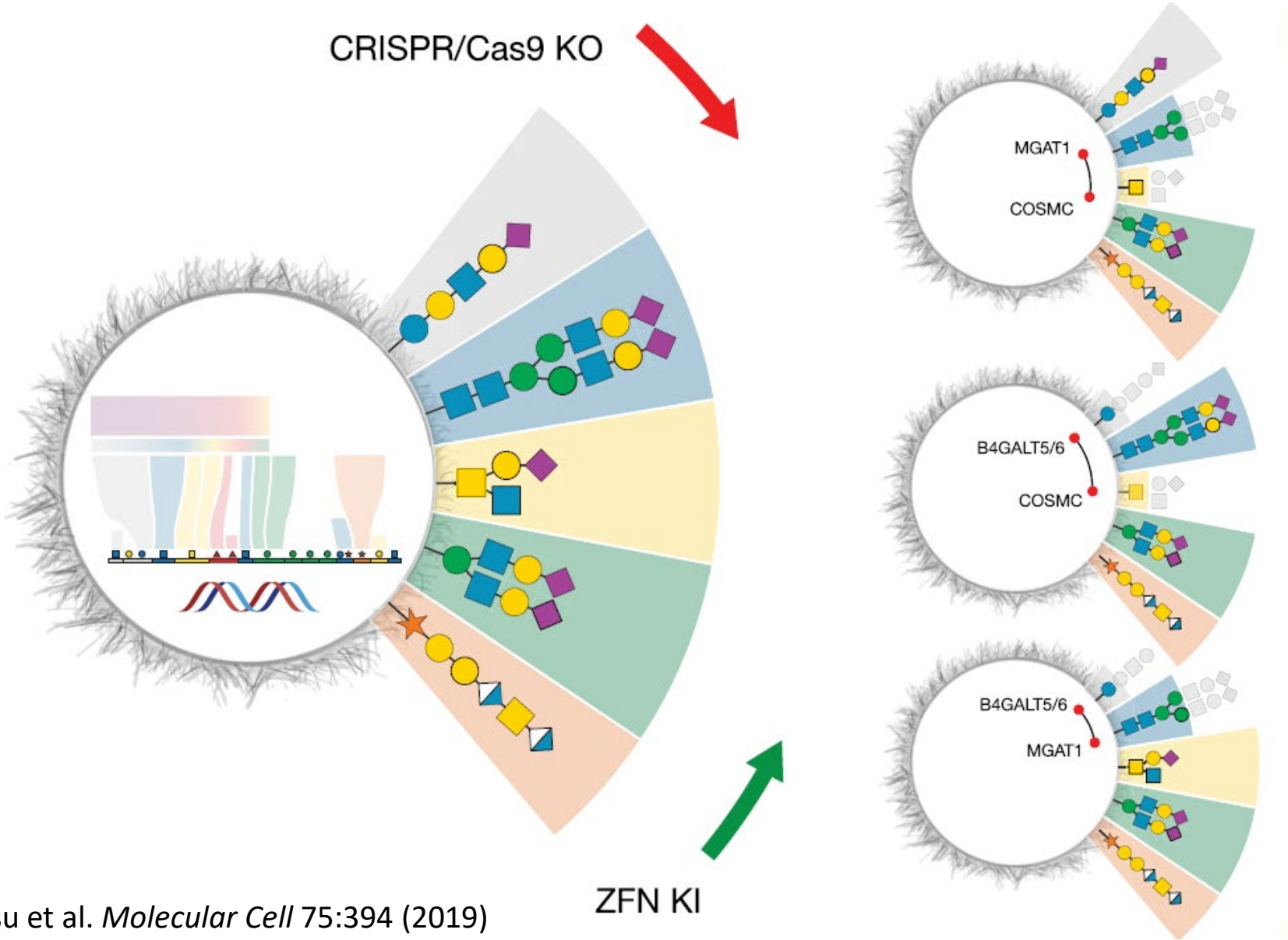


Paulson JC and Rogers GN (1987)
Methods Enzymol 138, 162



Kelm, et al. (1994) Curr. Biol. 4, 965

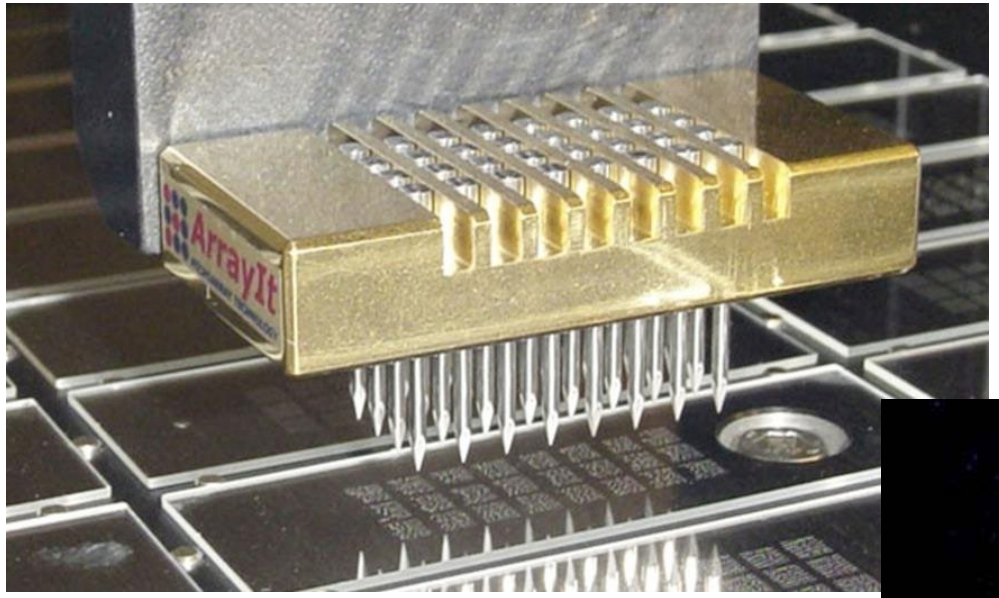
Glycome genetic modulation



Test GBP binding

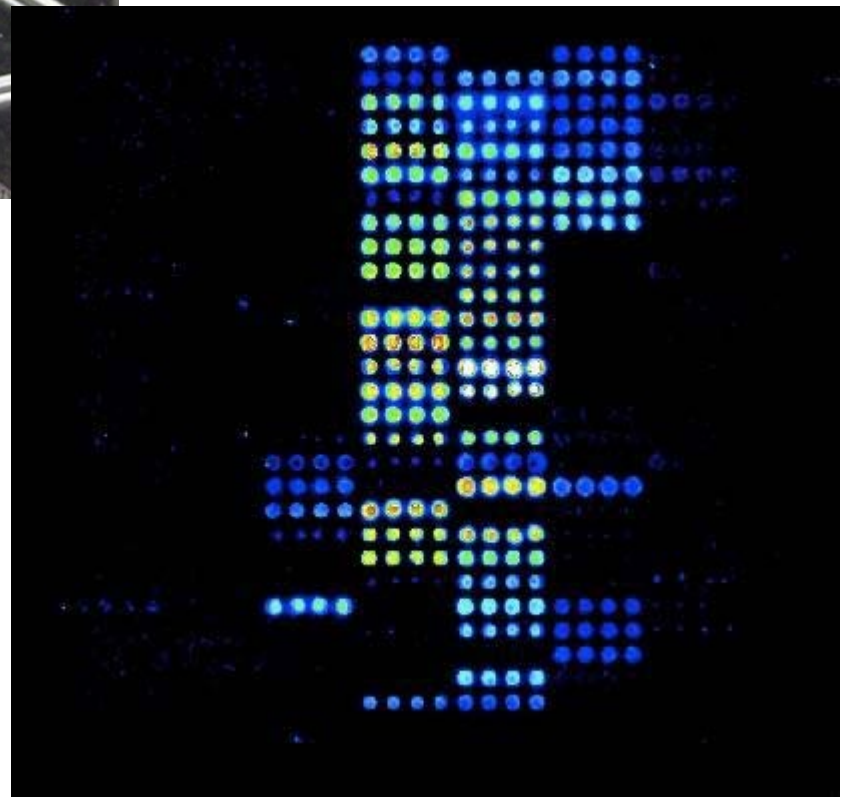
Narimatsu et al. *Molecular Cell* 75:394 (2019)

Glycan printed microarray



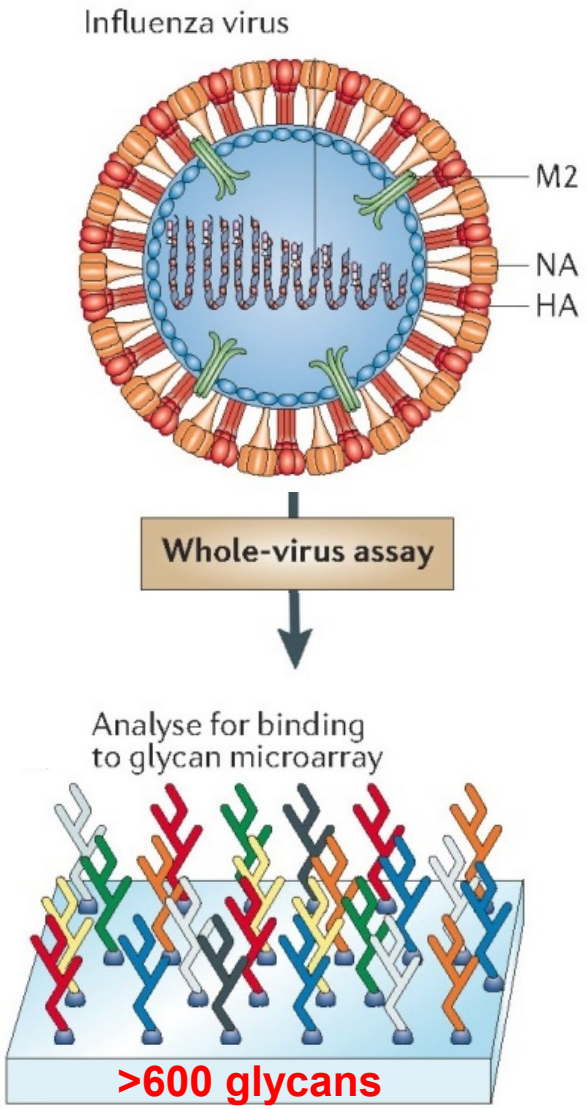
Stamp derivatized glycans onto glass slides

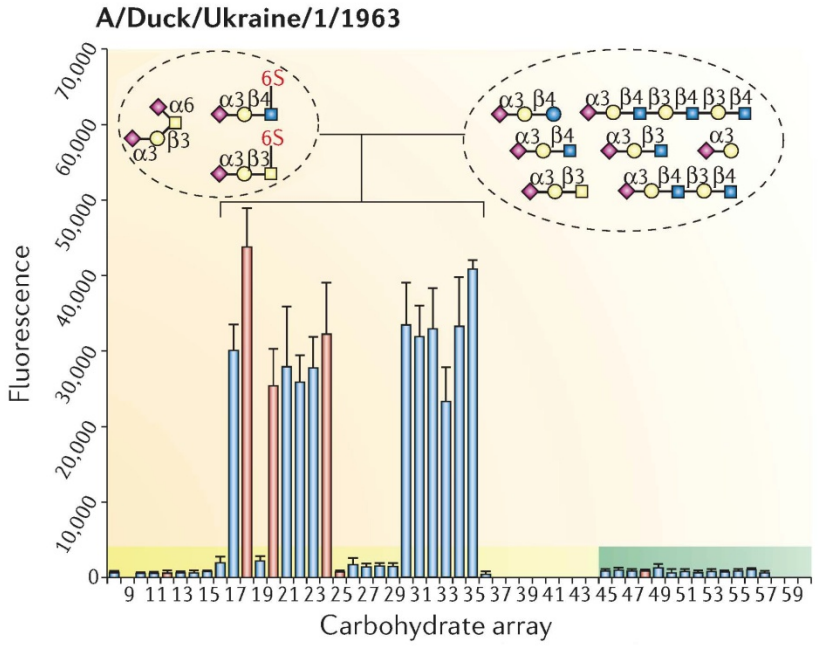
Overlay with labeled GBP, wash, detect

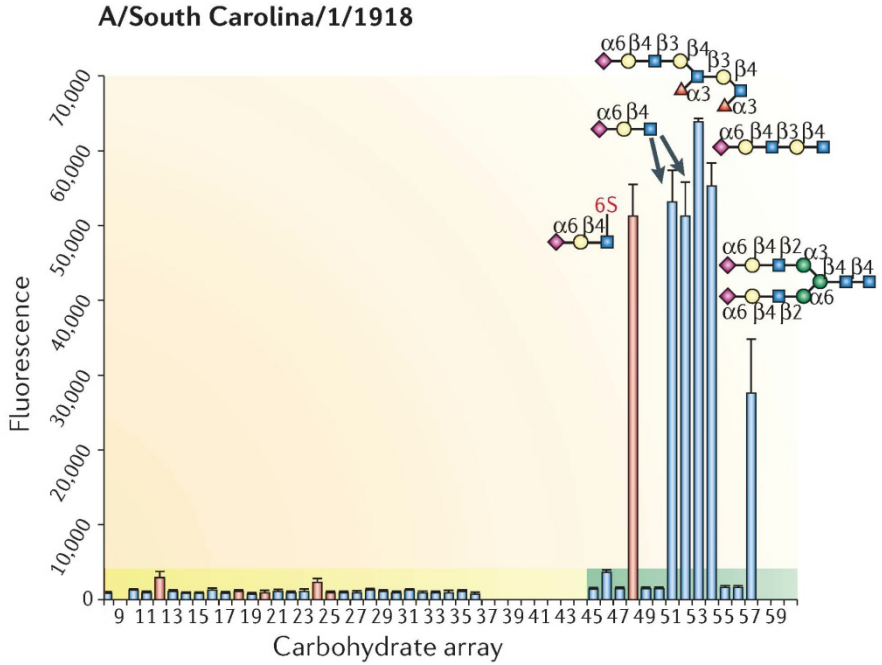


Ola Blixt, Copenhagen Center for Glycomics
http://glycomics.ku.dk/research/glyco-biomarkers/microarray/

Pathogen binding specificity

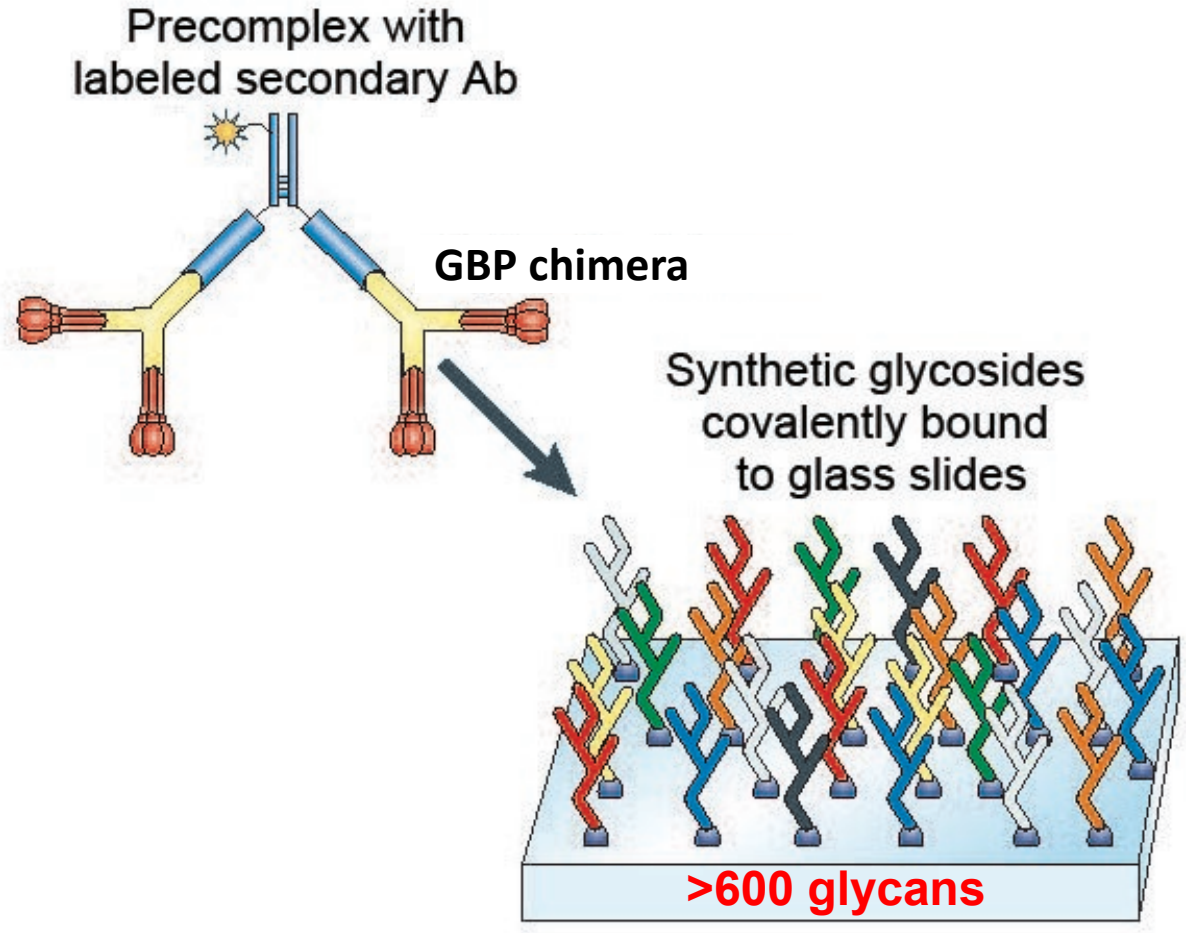






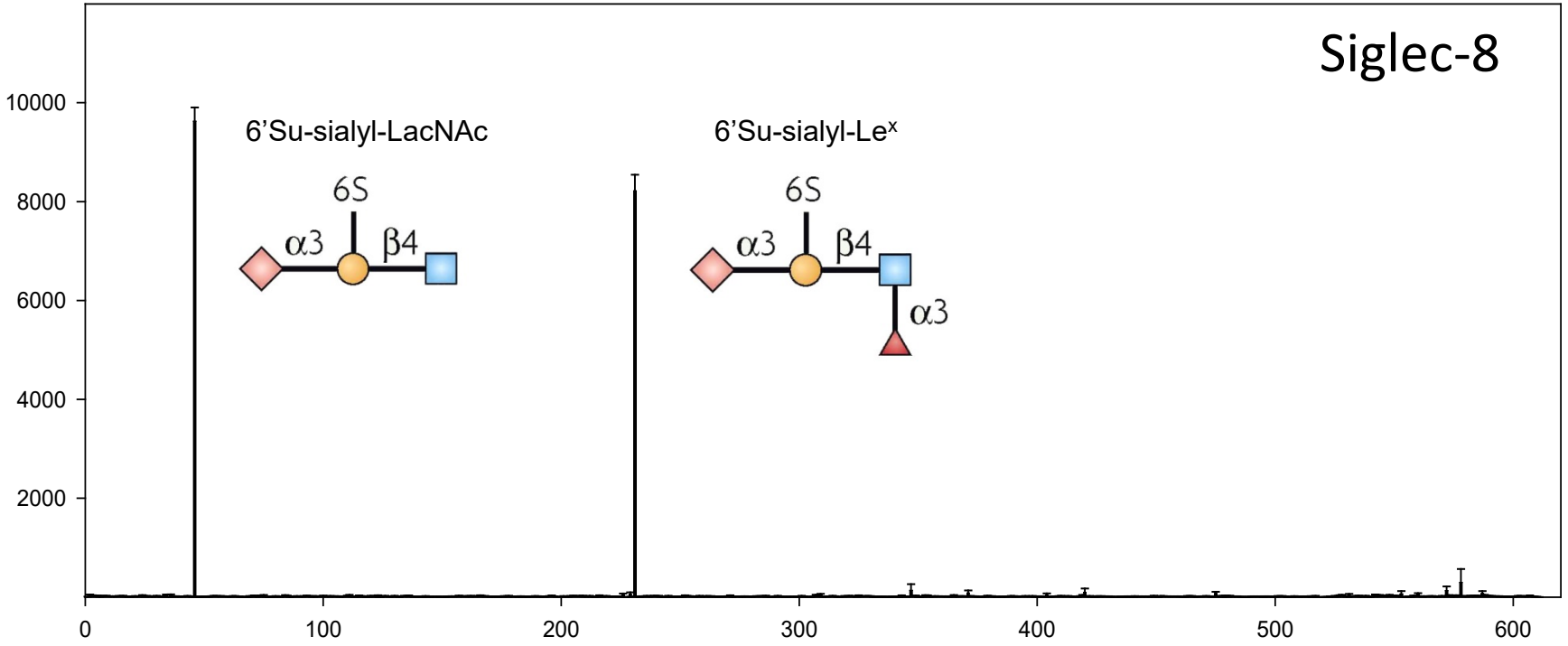
Stevens et al (2006) *Nat Rev Microbiol* 4:857

Glycan binding protein specificity by glycan array



Modified from Stevens et al (2006)
Nature Rev Microbiol 4:857

Examples of glycan array binding data



Bochner et al (2005) *J Biol Chem* 280:4307
& Bochner (2011) http://www.functionalglycomics.org/

Glycan array public resource: http://functionalglycomics.org

Cyanovirin-N Ulex Europaeus Agglutinin (UEA-I) Influenza A Puerto Rico

Objectives

- Learn the diversity and families of glycan binding proteins
- Learn the molecular/structural strategies used in protein-glycan binding and recognition
- Learn the methods and resources for determining glycan binding protein specificities
- Learn about plant lectins as tools in glycobiology research
- Learn about glycan binding strategies used by pathogens

Plant lectins

- Found in a variety of legumes and other plants, often in seeds
- Discovered by hemagglutination
- Proposed to function in plant pathogen and/or symbiont recognition
- Broad variety of glycan specificities - many have yet to be defined in detail
  (see Banana Lectin, Wheat Germ Agglutinin)
- Often sturdy proteins suitable for research, biotechnology and diagnostic applications

Plant lectins: sturdy & commercially available

https://vectorlabs.com/media/contentmanager/content/docs/brochures/LectinBrochure_2017.F.pdf

Lectin	Common Abbreviation	Source	Glycoprotein	Metal Ions Present	Mitogenic Activity	Blood Group Specificity	Preferred Sugar Specificity	Inhibitor or Eluting Sugar
Agaricus bisporus	ABL	*Agaricus bisporus* white button mushrooms	–	No	No	Non-specific	Gal(β-1,3) GalNAc	Fetuin
Aleuria aurantia	AAL	*Aleuria aurantia mushrooms*	No	--	No	Non-specific	Fucα6GlcNAc	L-Fuc
Bauhinia purpurea	BPL, BPA	*Bauhinia purpurea alba* (Camel's Foot Tree) seeds	Yes	No	Yes	A,B,O (-SA)	Galβ3GalNAc	Lactose
Concanavalin A	Con A	*Canavalia ensiformis* (Jack Bean) seeds	No	Ca^{++}, Mn^{++}	Yes	Non-specific	αMan, αGlc	MeαMan+ MeαGlc
Succinylated Concanavalin A	Succinylated Con A	*Canavalia ensiformis* (Jack Bean) seeds	No	Ca^{++}, Mn^{++}	Yes	None	αMan, αGlc	MeαMan+ MeαGlc
Datura stramonium	DSL	*Datura stramonium* (Thorn Apple, Jimson Weed) seeds	Yes	No	Yes	A, B, O	$(GlcNAc)_{2-4}$	Chitin hydrolysate
Dolichos biflorus	DBA	*Dolicos biflorus* (Horse Gram) seeds	Yes	Ca^{++}, Mn^{++}, Mg^{++} Zn^{++}	No	A_1>>A_2	αGalNAc	GalNAc
Erythrina cristagalli	ECL, ECA	*Erythrina cristagalli* (Coral Tree) seeds	Yes	Ca^{++}, Mn^{++}, Zn^{++}	Yes	A (–SA)	Galβ4GlcNAc	Lactose
Galanthus nivalis	GNL	*Galanthus nivalis* (Snowdrop) bulbs	No	No	No	Rabbit	αMan	MeαMan
Griffonia (Bandeiraea) simplicifolia I	GSL I, BSL I	*Griffonia (Bandeiraea) simplicifolia* seeds	Yes	Ca^{++}, Mn^{++}	No	B>>A1	αGal, αGalNAc	Gal/GalNAc

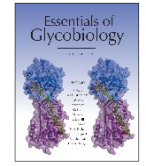
Objectives

- Learn the diversity and families of glycan binding proteins
- Learn the molecular/structural strategies used in protein-glycan binding and recognition
- Learn the methods and resources for determining glycan binding protein specificities
- Learn about plant lectins as tools in glycobiology research
- Learn about glycan binding strategies used by pathogens

Glycan recognition by pathogens

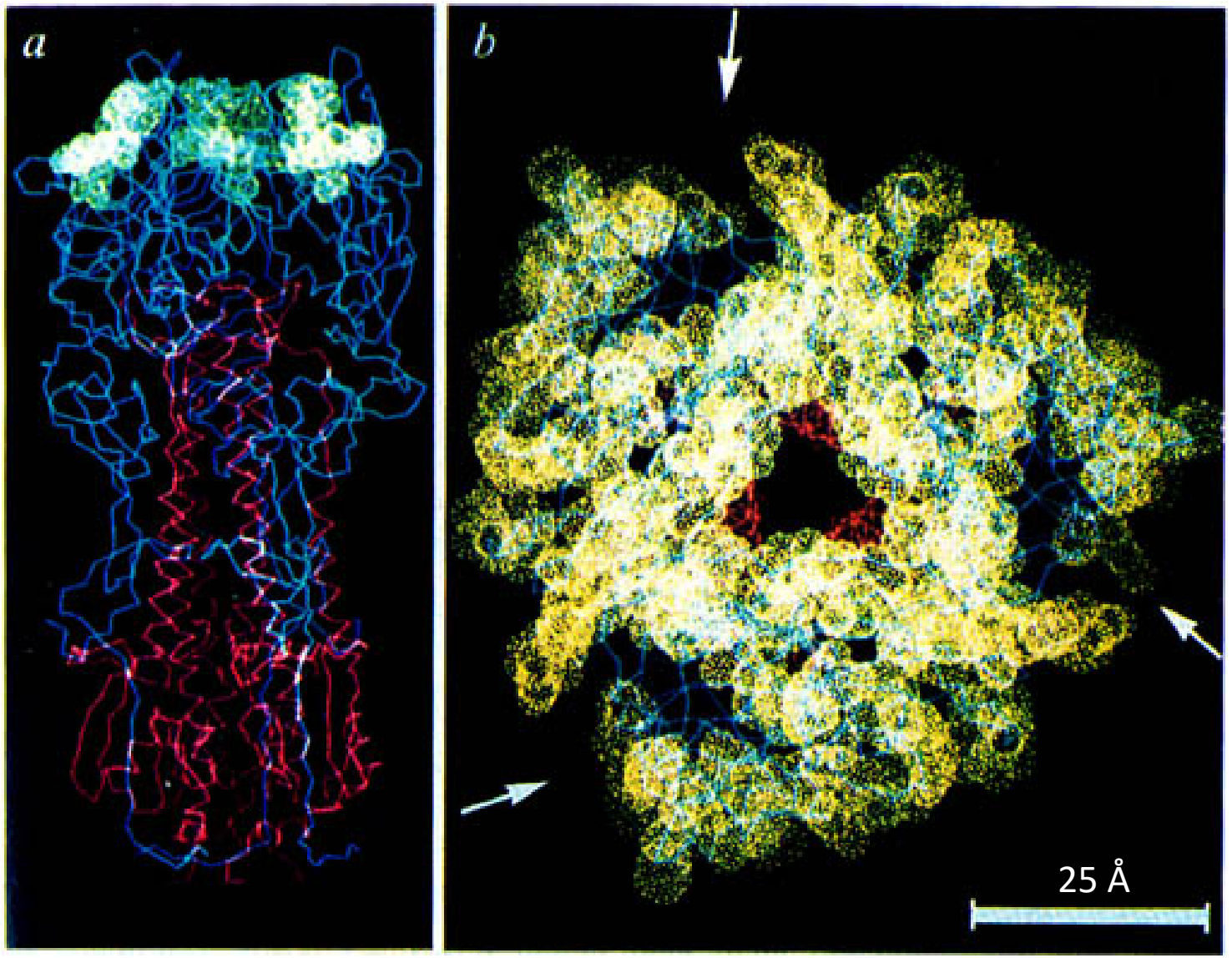
Examples of viral lectins and hemagglutinins

Virus	Lectin	Glycan-receptor specificity	Site of infection
Myxoviruses			
Influenza A and B (human, ferret, and porcine)	hemagglutinin	Neu5Acα2-6Gal-	upper respiratory tract mucosa (tracheal epithelial cells)
Influenza A and B (avian and porcine)	hemagglutinin	Neu5Acα2-3Gal-	intestinal mucosa
Influenza C	hemagglutinin-esterase	9-*O*-acetyl-Siaα-	unknown
Newcastle disease	hemagglutinin-neuraminidase	Neu5Acα2-3Gal-	unknown
Sendai	hemagglutinin-neuraminidase	Neu5Acα2-8Neu5Ac-	upper respiratory tract mucosa
Polyomaviruses			
Polyoma	capsid protein VP1	Neu5Acα2-3Gal-, Neu5Acα2-3Galβ1-3 (Neu5Acα2-6)GalNAc- on gangliosides such as GM1 and GT1b/GD1a	kidney and brain glial cells
Herpesviruses			
Herpes simplex	glycoproteins gB, gC, and gD	3-O-sulfated heparan sulfate	mucosal surfaces of the mouth, eyes, genital, and respiratory tracts
Picornaviruses			
Foot-and-mouth disease (enterovirus)	caspid proteins	heparan sulfate	gastrointestinal and upper respiratory tracts
Retroviruses			
HIV	gp120 V3 loop	heparan sulfate	CD4 lymphocytes
Flaviviruses			
Dengue	envelope protein	heparan sulfate	macrophages?
Calciviruses			
Norovirus	capsid proteins	fucose, GalNAc, or Gal on A and B blood group antigens	secretory cells of the intestinal epithelium



Influenza virus hemagglutinin

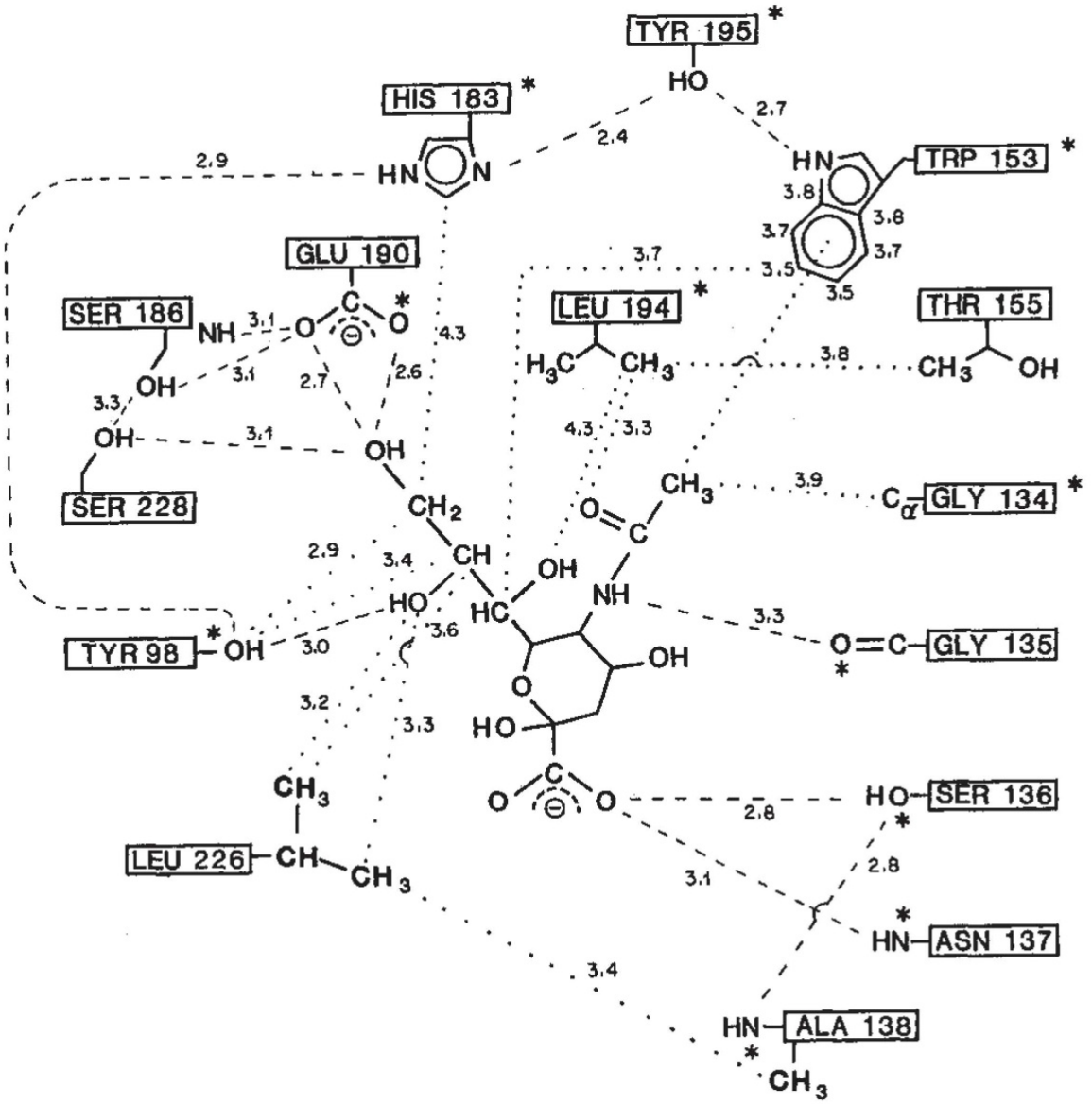
Side (left) and top (right) views; arrows, sialic acid binding sites



Weis et al (1988) *Nature* 333:426

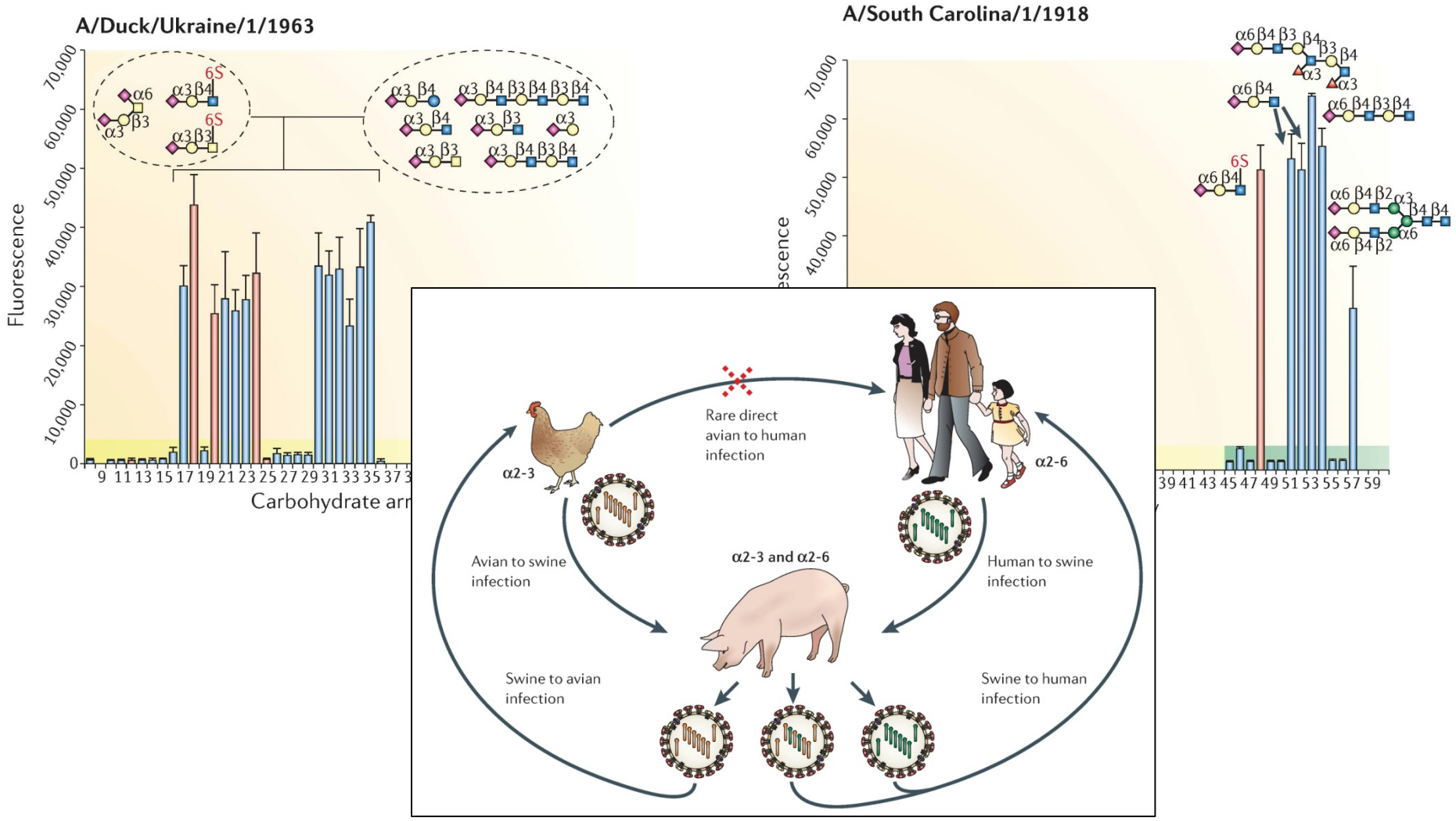
Sialic acid binding to influenza virus hemagglutinin

- hydrogen bonds
- hydrophobic stacking
- polyvalent binding



Weis et al (1988) *Nature* 333:426

Decoding influenza hemagglutinin binding specificity



Stevens, et al. (2006) *Nat Rev Microbiol.* 4:857-64

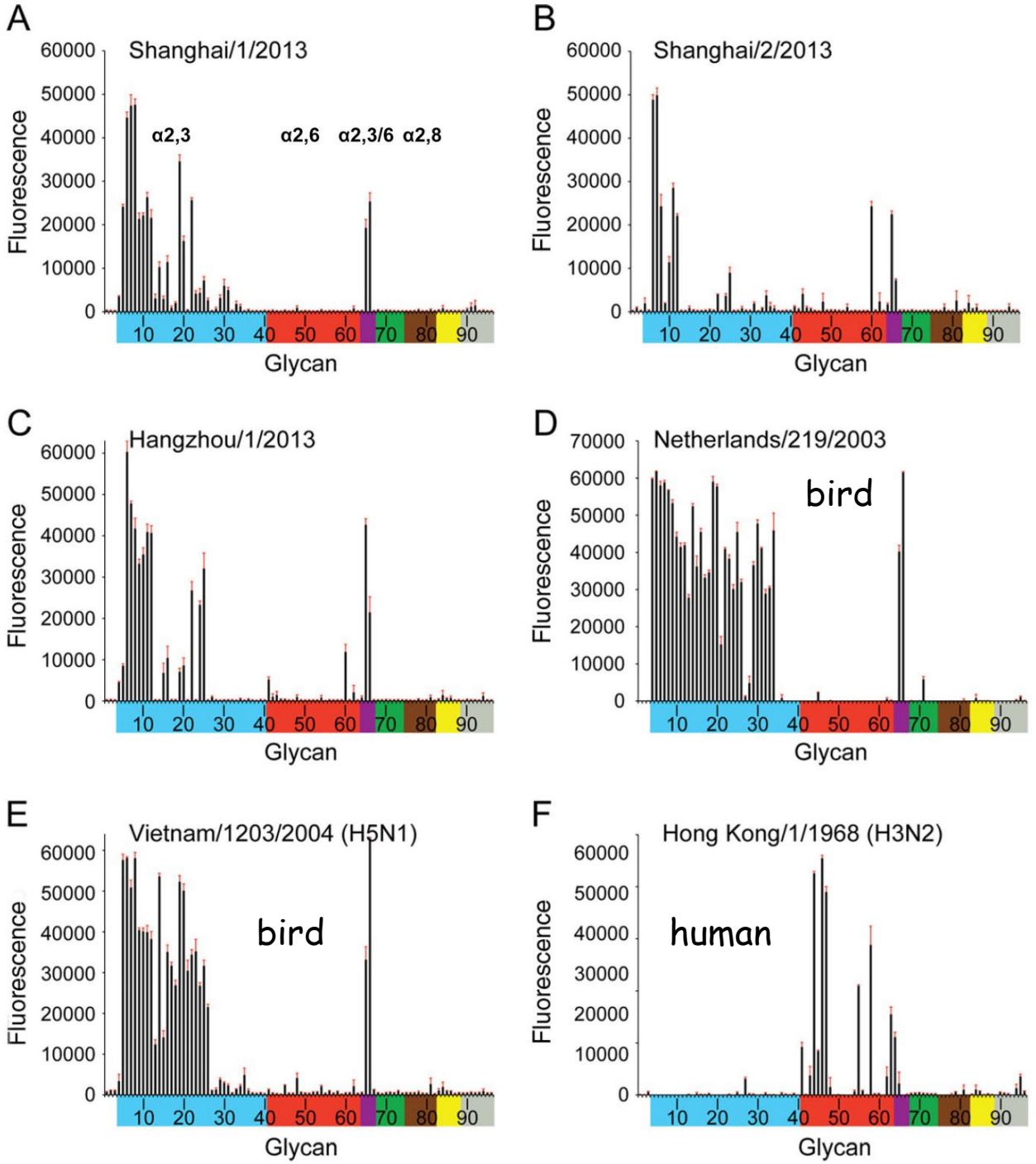
Monitoring for bird-human hemagglutinin modificaitons

H7N9 influenza (by mid 2013):

- 134 cases
- 127 hospitalizations
- 43 fatalities

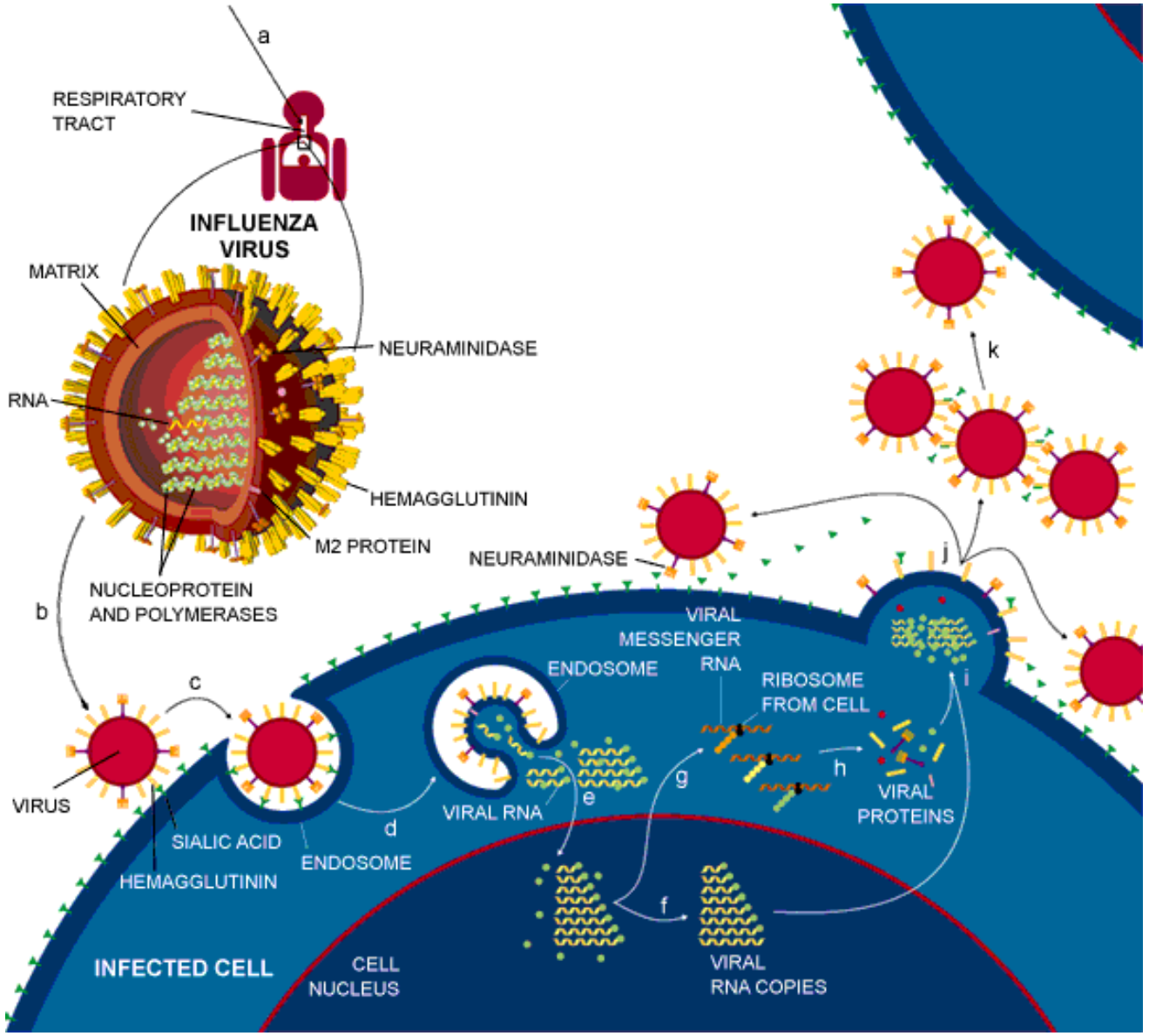
H7N9 influenza to date:

- 1,223 cases
- >450 fatalities

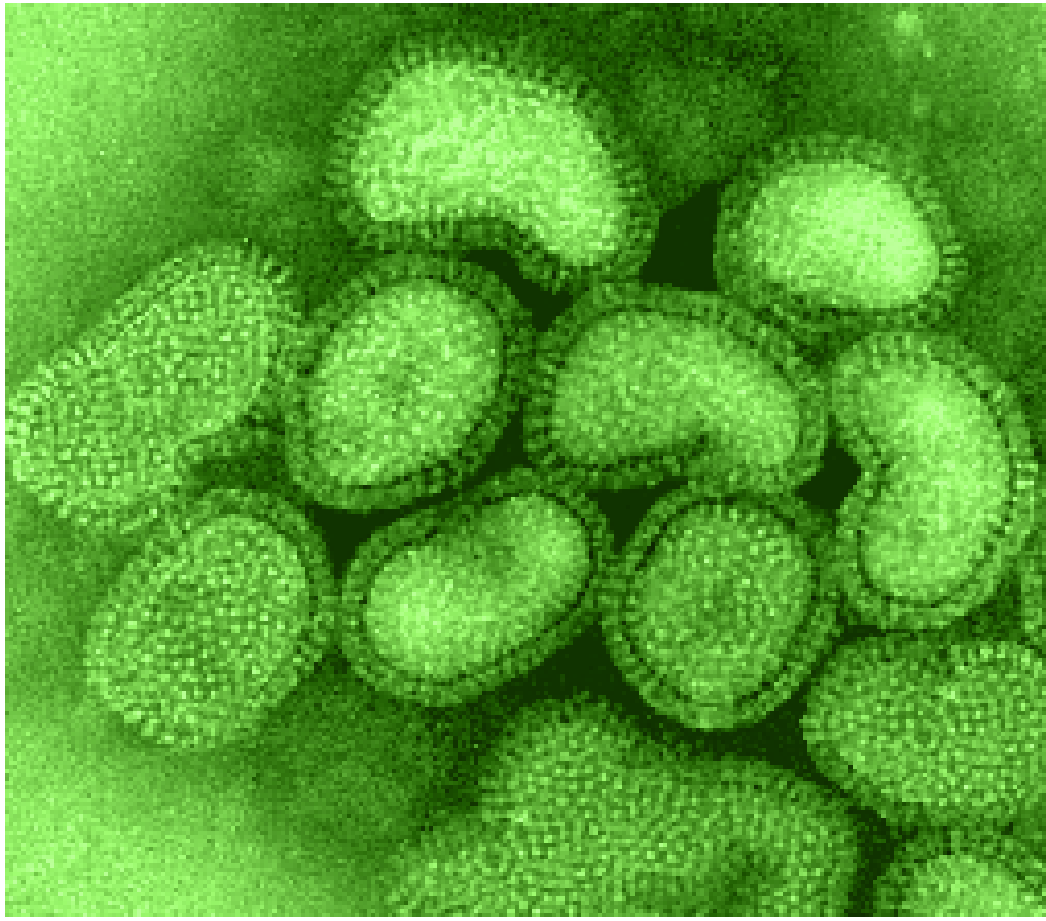


Yang, et al. (2013) *J Virol* 87, 12433-46

Influenza virus lifecycle

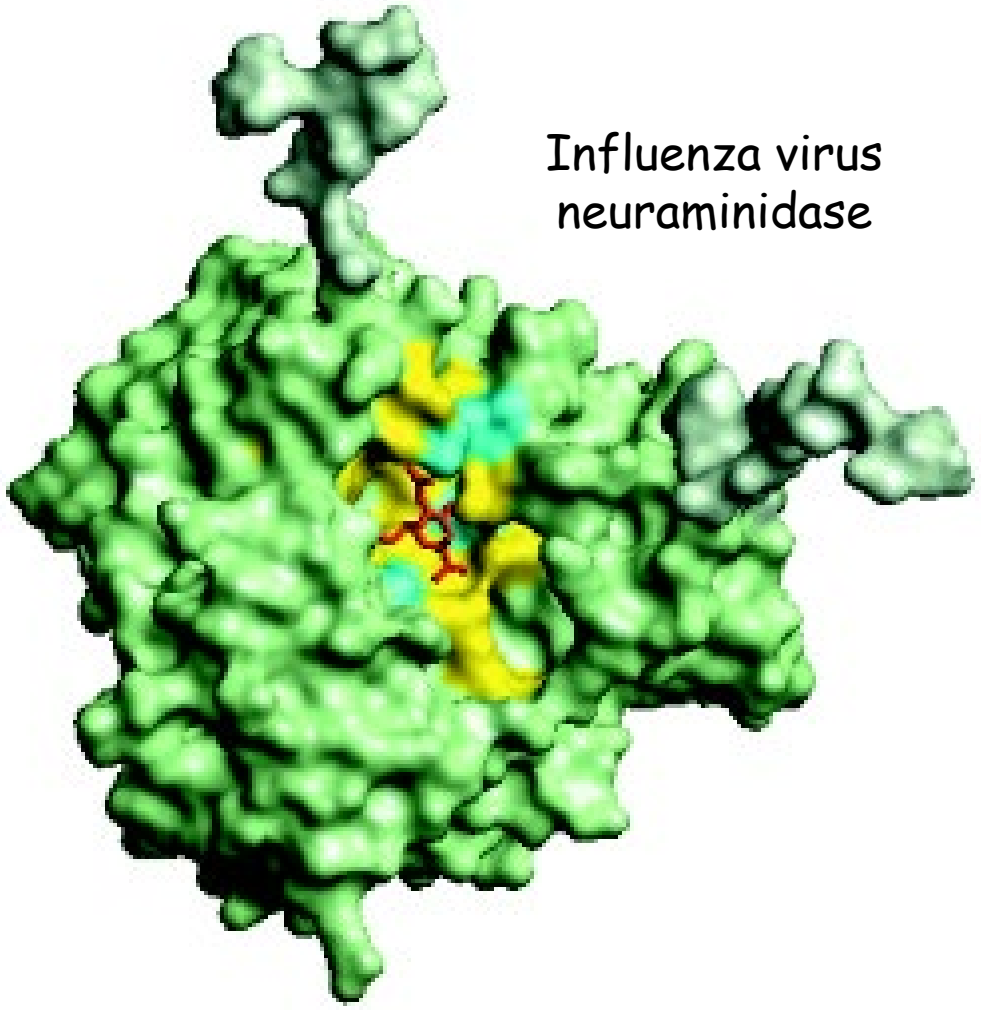


Laver et al (1999) *Scientific American*, Jan issue, pp78-87



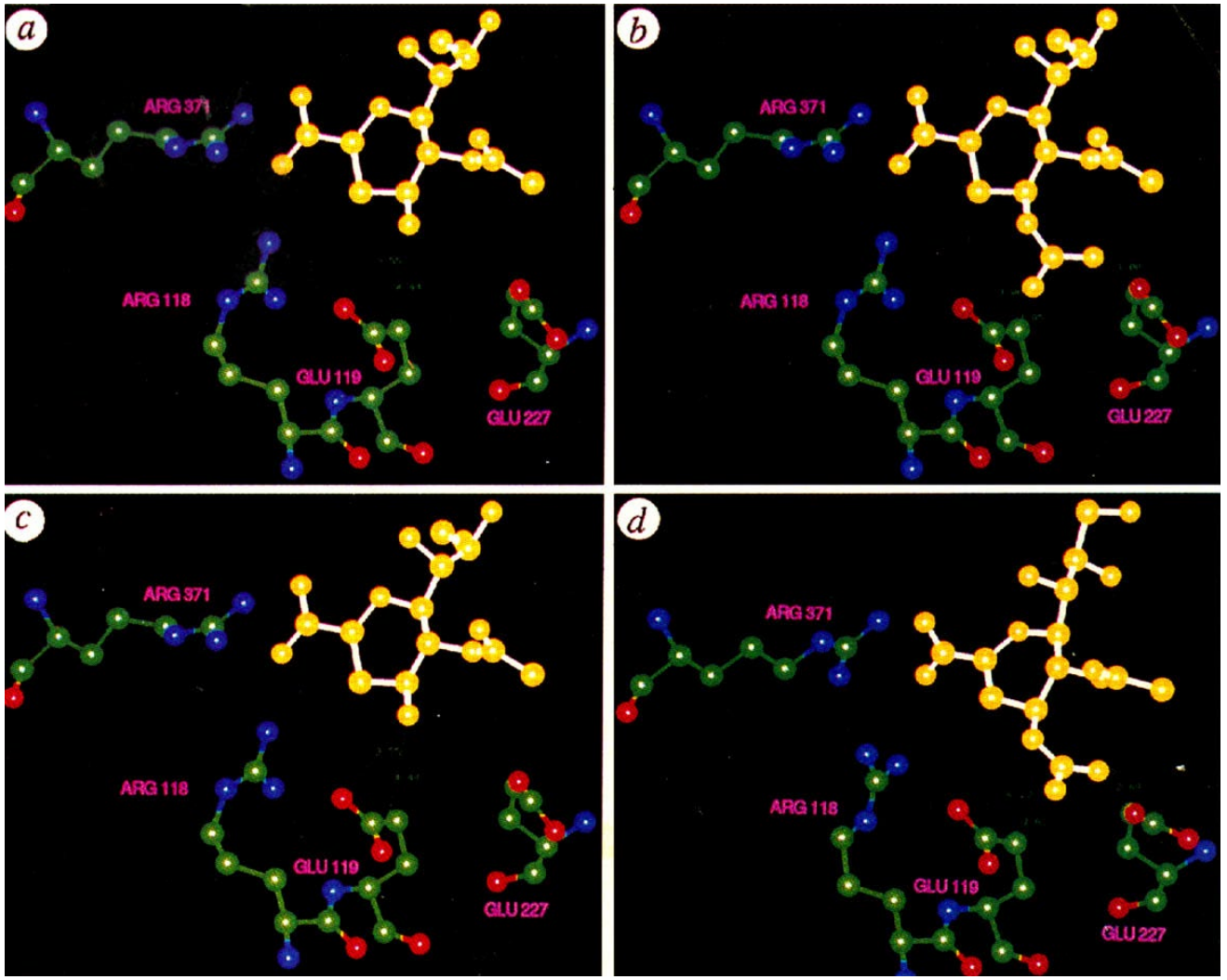
Linda Stannard:
http://web.uct.ac.za/depts/mmi/stannard/fluvirus.html

Rational design of an anti-influenza drug



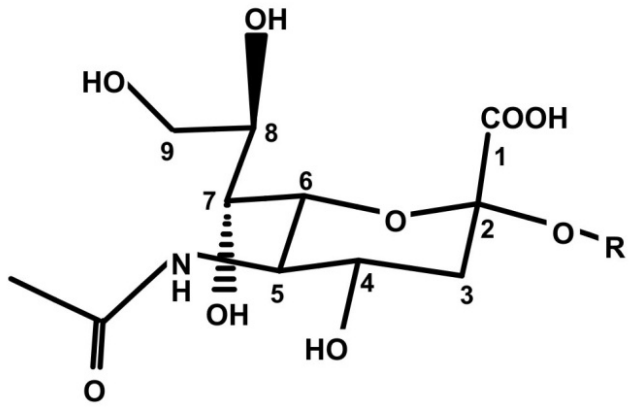
Influenza virus
neuraminidase

Gubareva et al (2000) Lancet 355:827

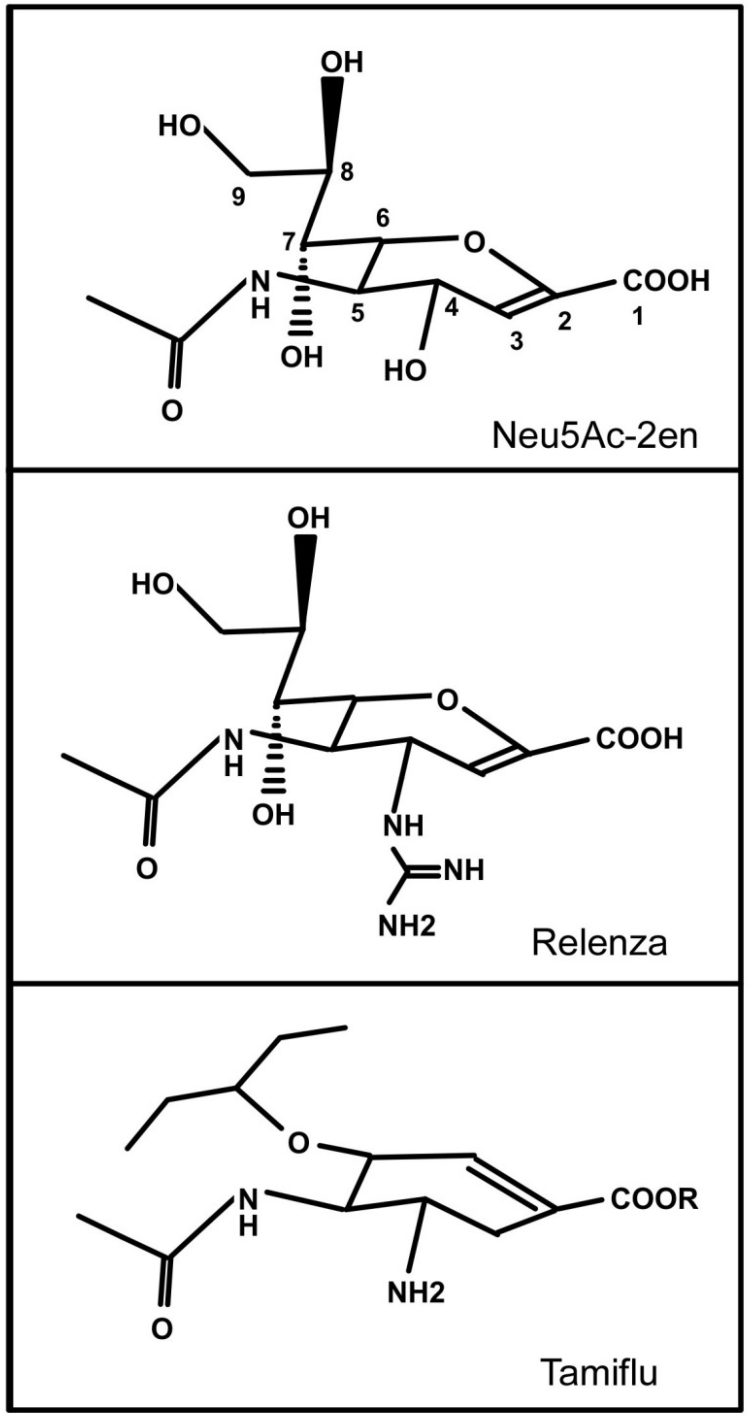


von Itzstein et al (1993) Nature 363:418

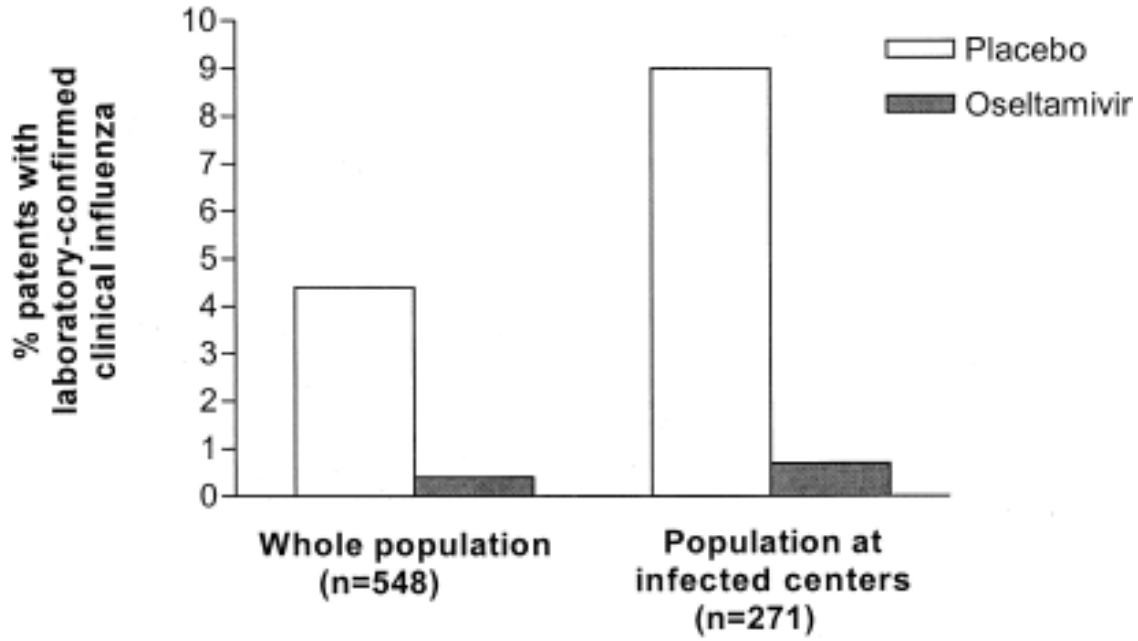
Sialic acid-based influenza neuraminidase inhibitors



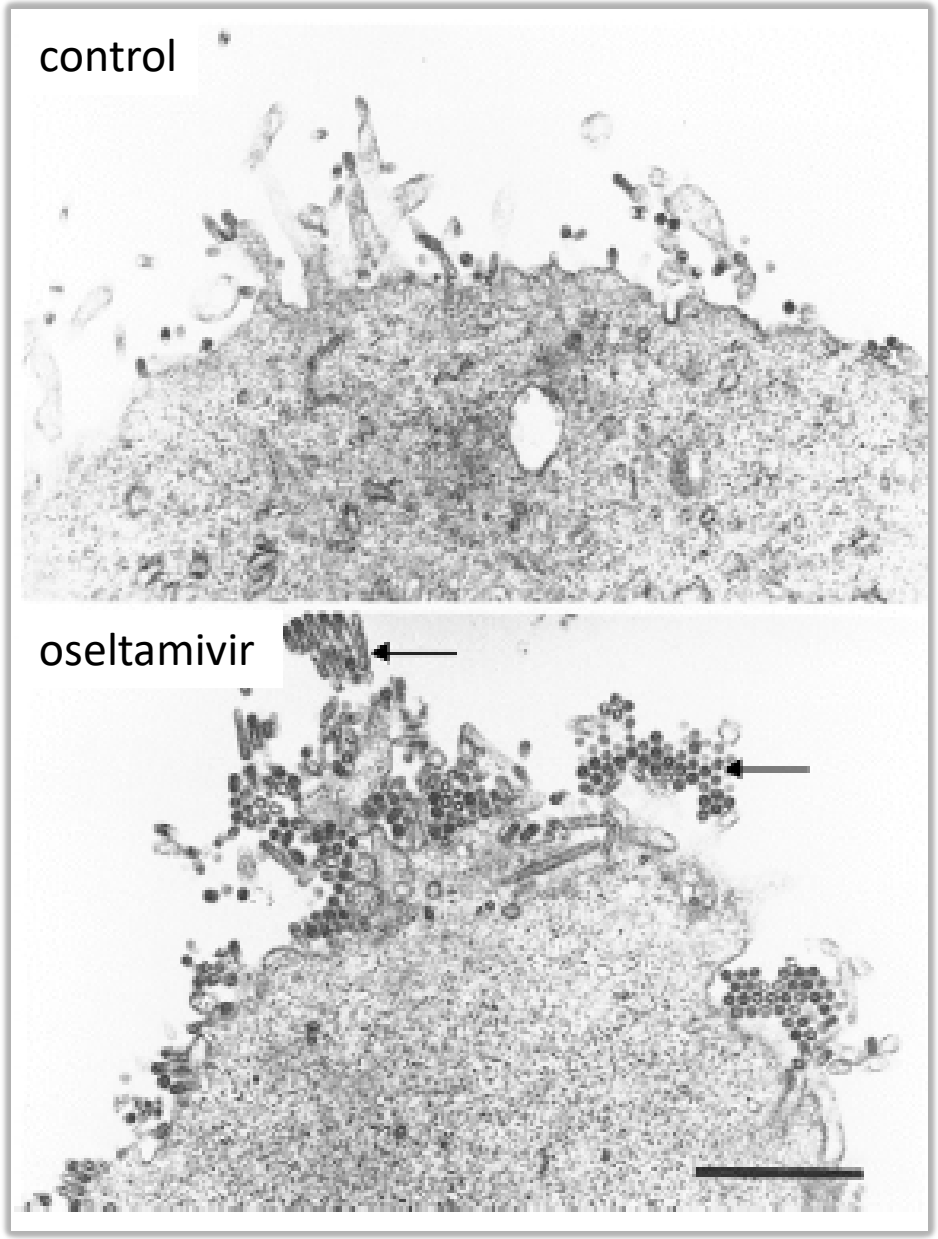
Sialic Acid



Tamiflu mechanism of action: blocks viral dissemination



Peters, P.H., et al. (2001) *J Am Geriatr Soc* 49, 1025

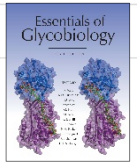


Gubareva et al (2000) *Lancet* 355:827

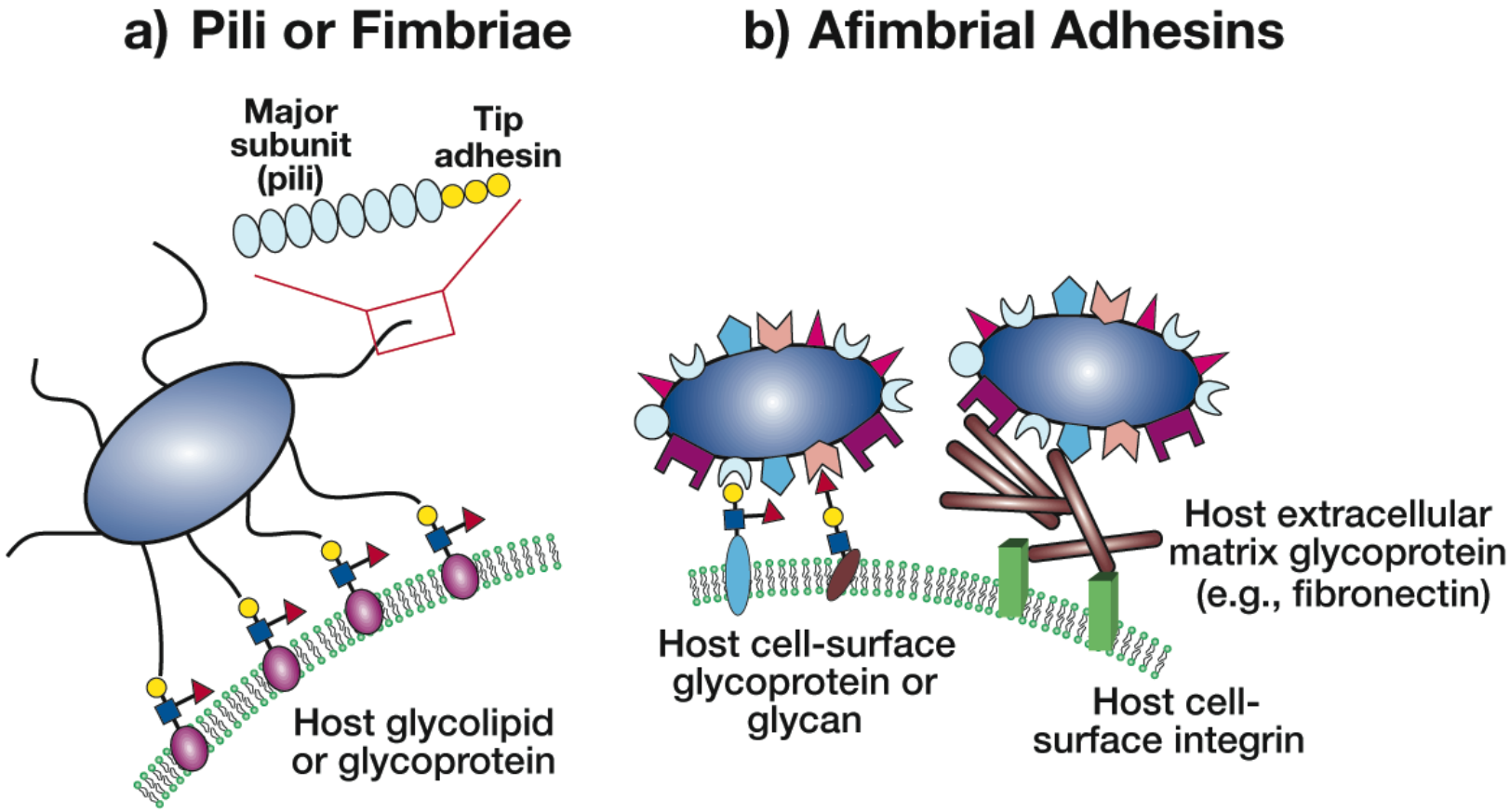
Glycan recognition by pathogens

Examples of interactions of **bacterial adhesins** with glycans

Microorganism	Adhesin	Glycan-receptor specificity	Site of infection
Actinomyces naeslundii	fimbriae	Galβ1–3GalNAcβ-	oral
Bordetella pertussis	filamentous hemagglutinin (FHA)	sulfated glycolipids, heparin	ciliated epithelium in respiratory tract
Borrelia burgdorferi	ErpG protein	heparan sulfate	endothelium, epithelium, and extracellular matrix
Campylobacter jejuni	flagella, LPS	Fucα1-2Galβ1-4GlcNAcβ- (H-antigen)	intestinal cells
Escherichia coli	P fimbriae	Galα1-4Galβ-	urinary tract
	S fimbriae	gangliosides GM3, GM2	neural
	type-1 fimbriae	Manα1-3(Manα6Manα1-6) Man	urinary tract
	K99 fimbriae	gangliosides GM3, Neu5Gcα2-3Galβ1-4Glc	intestinal cells
Haemophilus influenzae	HMW1 adhesin	Neu5Acα2-3Galβ1-4 GlcNAcβ-, heparan sulfate	respiratory epithelium
Helicobacter pylori	BabA	sialyl Lewis x	stomach
Helicobacter pylori	SabA	lewis b	stomach and stomach duodenum
Mycobacterium tuberculosis	heparin-binding hemagglutinin adhesin (HBHA)	heparan sulfate	respiratory epithelium
Neisseria gonorrhoeae	Opa proteins protein	LacCer; Neu5Acα2-3Galβ1-4 GlcNAcβ-, syndecans, heparan sulfate	genital tract
Pseudomonas aeruginosa	type IV pili	asialo GM1 and GM2	respiratory tract
Staphylococcus aureus	signal peptide of panton valentine leukocidin	heparan sulfate	connective tissues and endothelial cells
Streptococcus agalactiae	αC protein	heparan sulfate	brain endothelial cells
Streptococcus pneumoniae	carbohydrate-binding modules of β-galactosidase, BgaA	lactose or *N*-acetyl-lactosamine	respiratory tract



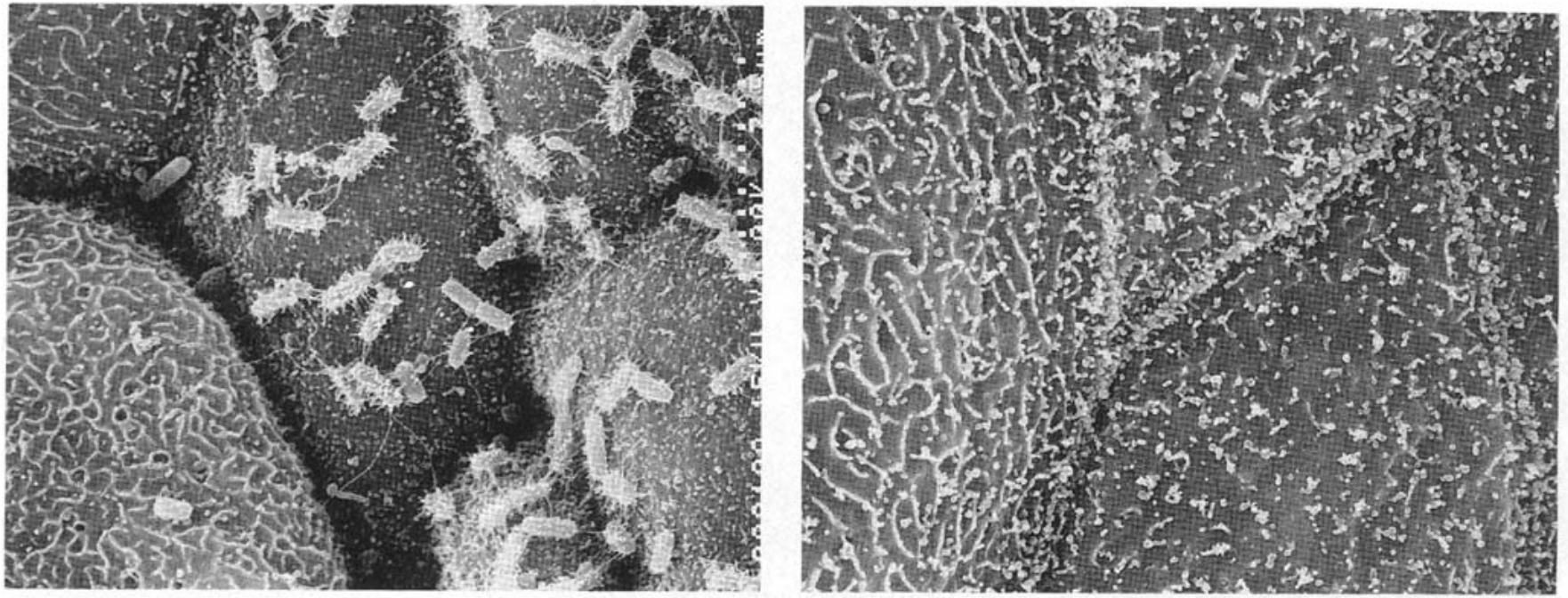
Bacterial adhesins



E. coli binding to urinary epithelium (in vitro)

control

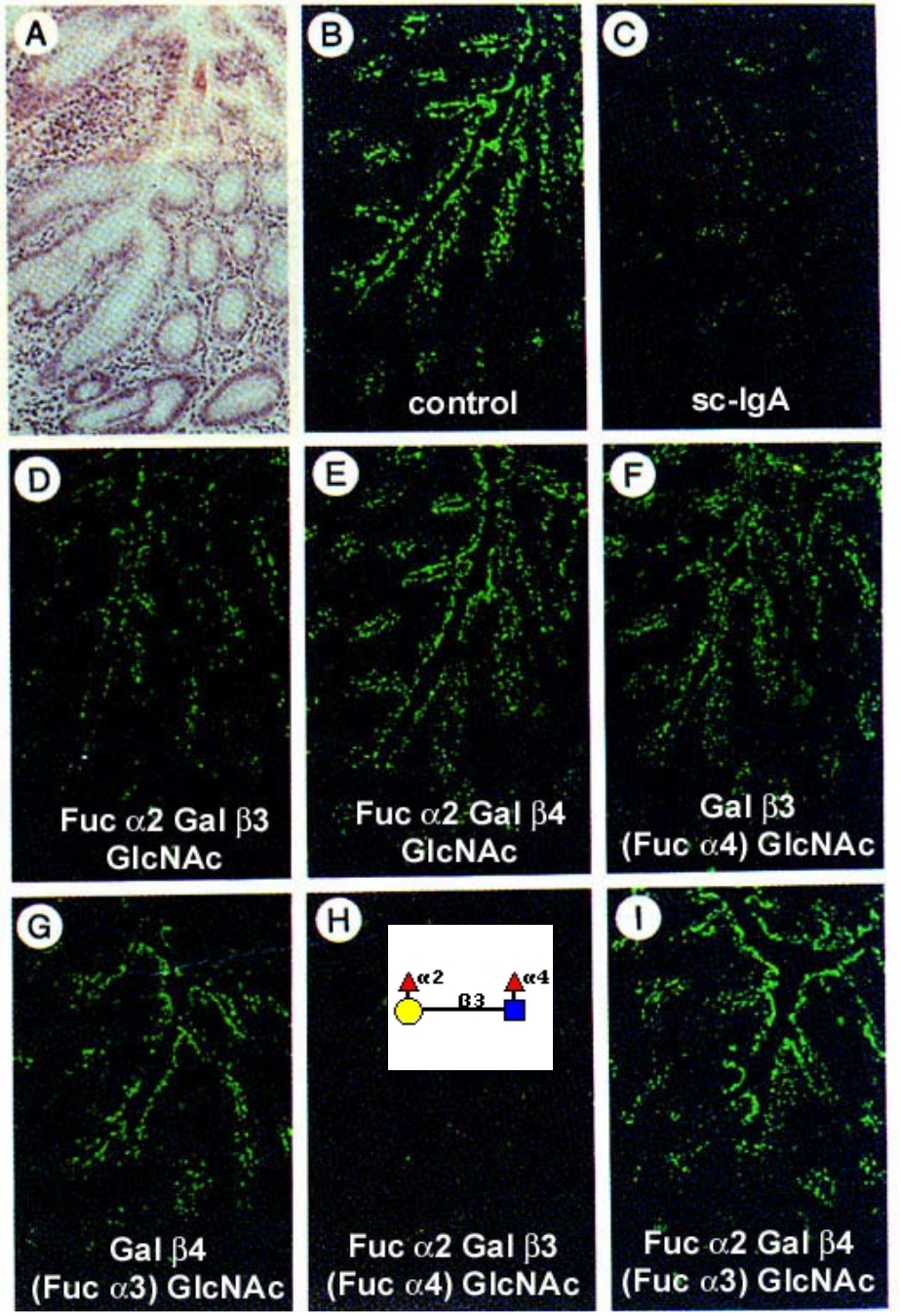
+ Gal α1-4 Gal



Sharon and Lis (1993) *Scientific American* Jan:82

Helicobacter pylori epithelial adherence

Inhibition of *H. pylori* binding (green) to human gastric epithelium by glycans

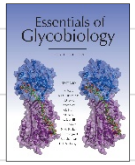


Boren et al (1993) *Science* 262:1892

Glycan recognition by pathogens

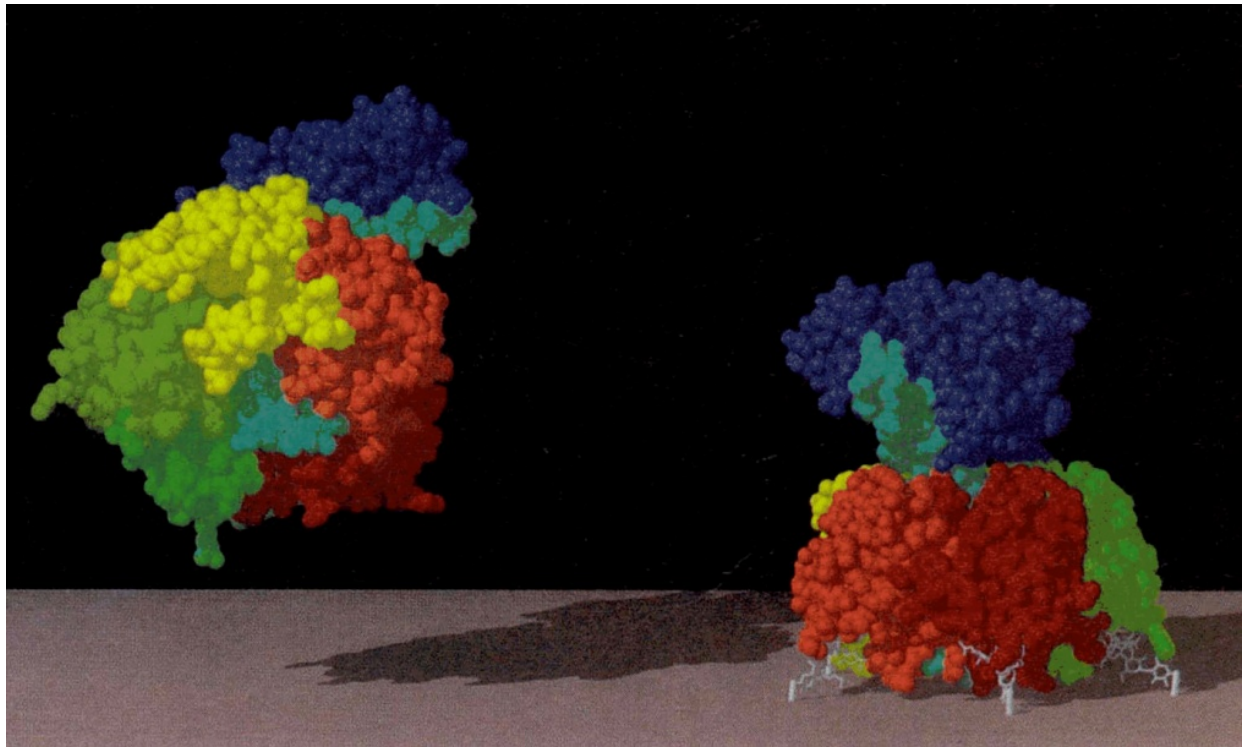
Examples of glycan receptors for **bacterial** toxins

Microorganism	Toxin	Glycan-receptor specificity	Site of infection
Bacillus thuringiensis	crystal toxins	Galβ1-3/6Galα/β1-3(±Glcβ1-6) GalNAcβ GlcNAcβ1-3Manβ1-4 GlcβCer	intestinal epithelium of insects/nematodes
Clostridium botulinum	botulinum toxins (A–E)	gangliosides GT1b and GQ1b	nerve membrane
Clostridium difficile	toxin A	GalNAcβ1-3Galβ1-4GlcNAcβ1-3 Galβ1-4GlcβCer	large intestine
Clostridium tetani	tetanus toxin	ganglioside GT1b	nerve membrane
Escherichia coli	heat-labile toxin	GM1	intestine
Shigella dysenteriae	Shiga toxin	Galα1-4GalβCer, ;Galα1-4Galβ1-4 GlcβCer	large intestine
Vibrio cholerae	cholera toxin	GM1	small intestine

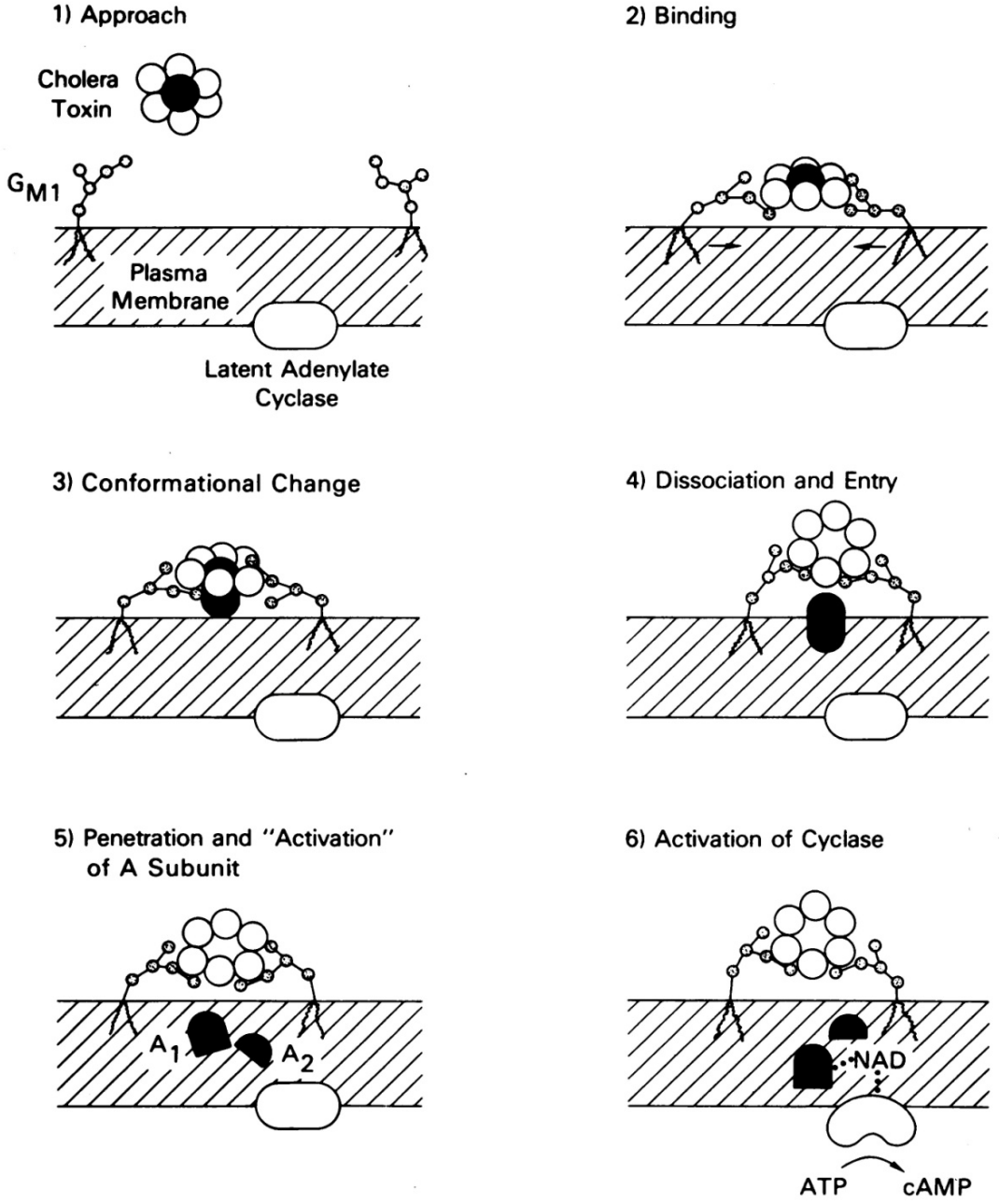


Bacterial AB5 toxins

- Cholera toxin (CTB)
- *E coli* enterotoxin
- Pertussis toxin
- Shiga toxin
- Shiga-like verotoxins

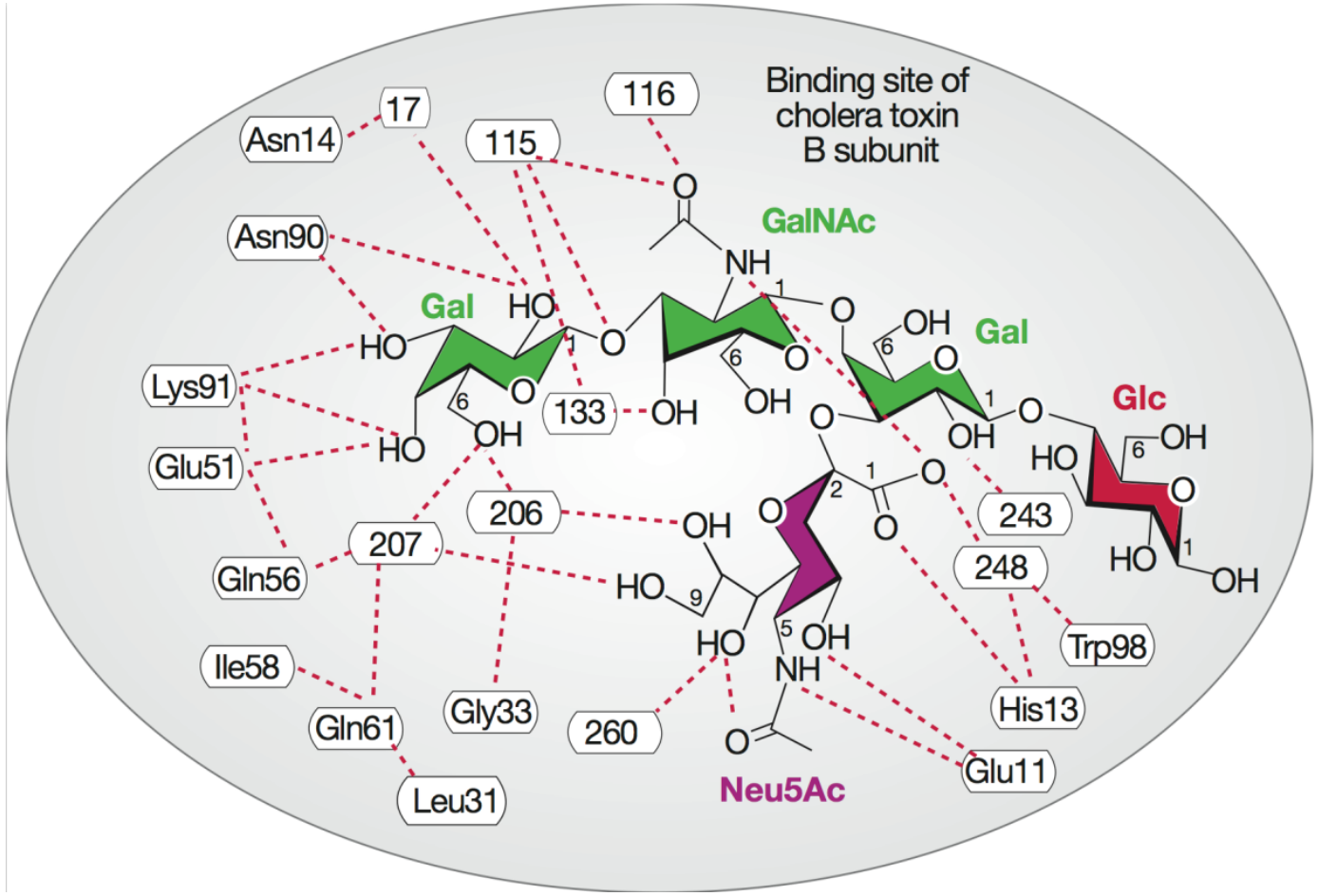


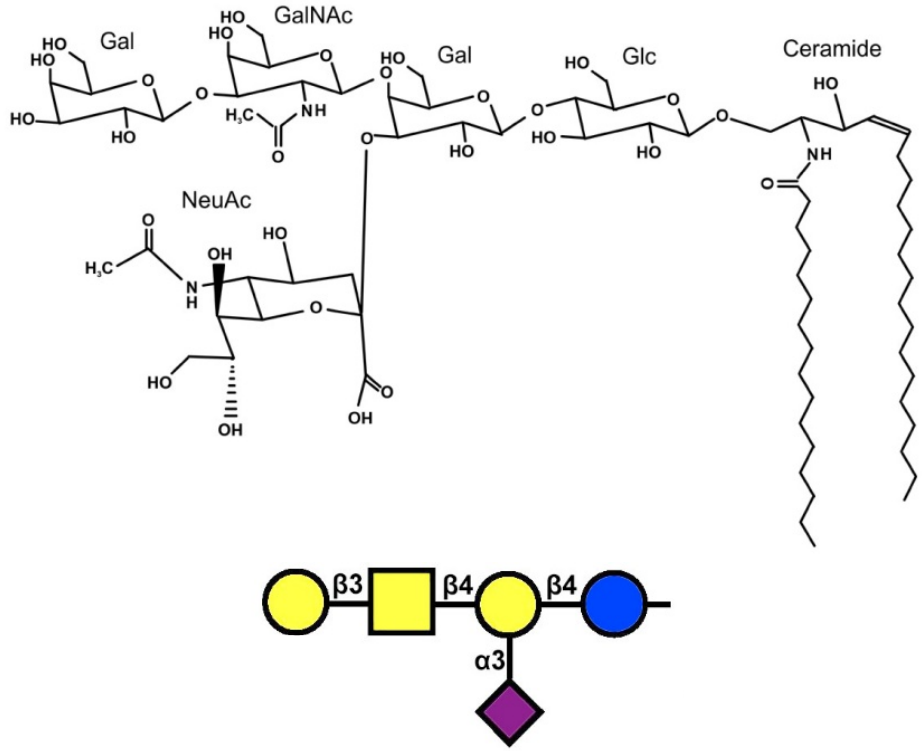
Merritt et al (1994) *Protein Science* 3:166



Fishman and Brady (1976) *Science* 194:906

CTB binding: a lattice of hydrogen bonds

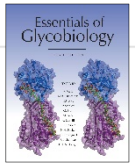




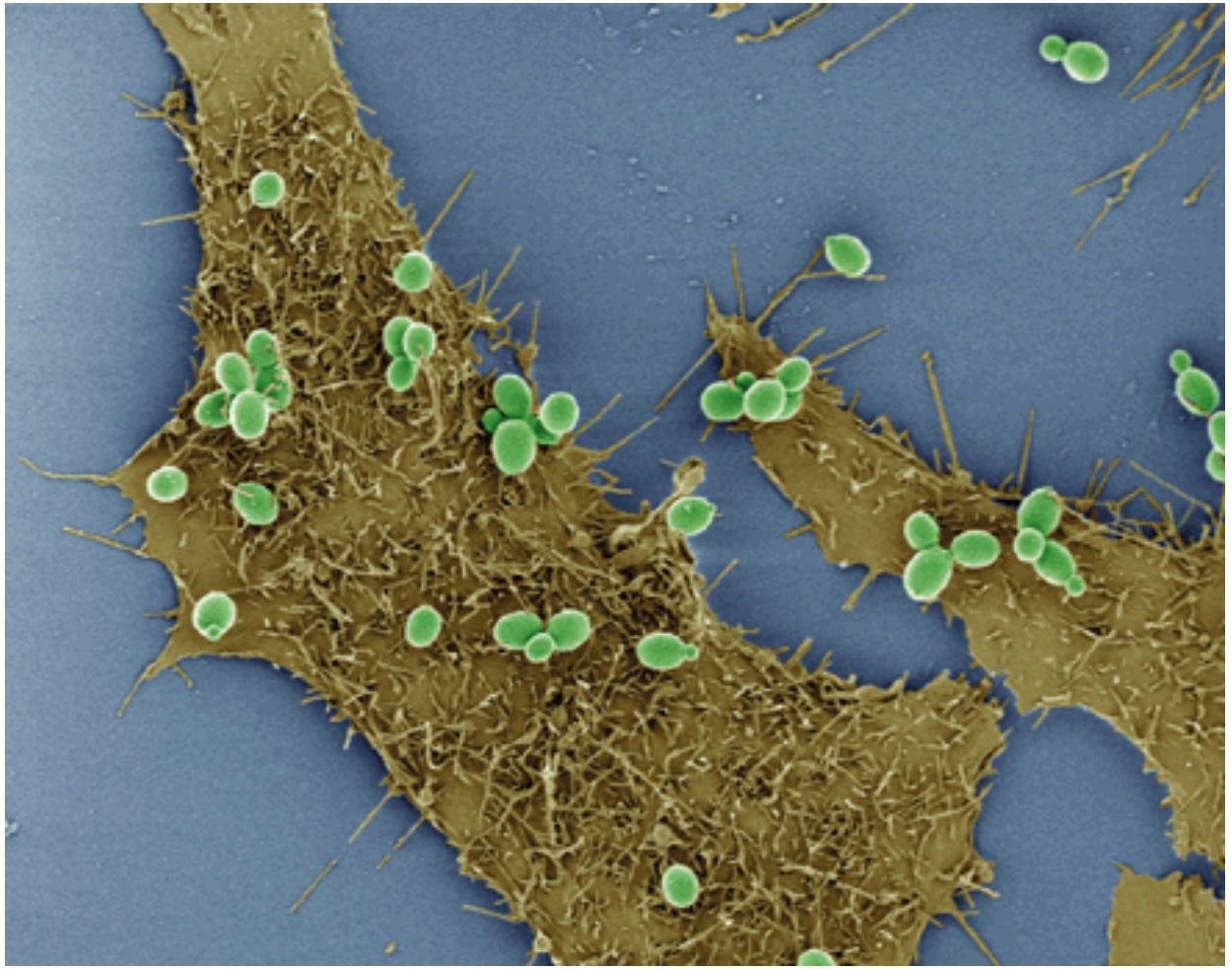
Glycan recognition by pathogens

Examples of glycan receptors for parasites

Parasite	Adhesin	Glycan-receptor specificity	Site of infection
Entamoeba histolytica	260-kDa surface-anchored lectin on trophozoites	terminal Gal/GalNAc residues	mucosa of human colon
Plasmodium falciparum	EBA-175; circumsporozoite (CS) protein	sialic acid–containing glycans (Neu5Acα2-3Galβ-) on glycophorins; heparan sulfate proteoglycans	erythrocytes (infected cells bind to placental vasculature) and hepatocytes
Trypanosoma cruzi	surface "mucins"	sialic acid–containing glycans and heparan sulfate	multiple cell types
Leishmania amazonensis	unknown	heparan sulfate	macrophages, fibroblasts, and epithelium
Cryptosporidium parum	lectin p30	terminal Gal-GalNAc	intestinal epithelium
Giardia lamblia	unknown	mannose-terminated oligosaccharides	duodenum and small intestine
Toxoplasma gondii	microneme protein 1 (TgMIC1)	α2-3-linked sialyl-*N*-acetyllactosamine sequences	intestinal epithelium



Binding of *Candida glabrata* to human epithelial cells



Cormack et al (1999) *Science* 285:578

Glycoconjugate	[50% inhibition]
Glucose	>200 mM
Galactose*	10 mM
Fucose	>200 mM
Xylose	>200 mM
Mannose	>200 mM
Lactose *	1.5 mM
LacNAc *	1.25 mM
Sialyl-LacNAc	>10 mM
NANA	>200 mM
GlcNAc	>200 mM
GalNAc	>200 mM
Methyl-D pyranosic	>200 mM
Mannosamine	>200 mM
Dextran	>1 mg/ml
Dextran-SO_4	>1 mg/ml
Mannan	>1 mg/ml
Hyaluronic acid	>1 mg/ml
Heparin	>1 mg/ml
Albumin	>1 mg/ml
Fucoidan	>1 mg/ml
Fetuin	>1 mg/ml
Asialofetuin	>1 mg/ml
Chondroitin-SO_4-A	>1 mg/ml
Chondroitin-SO_4-B	>1 mg/ml
Chondroitin-SO_4-C	>1 mg/ml

EPA7 glycan array screening

Objectives

- Learn the diversity and families of glycan binding proteins
- Learn the molecular/structural strategies used in protein-glycan binding and recognition
- Learn the methods and resources for determining glycan binding protein specificities
- Learn about plant lectins as tools in glycobiology research
- Learn about glycan binding strategies used by pathogens