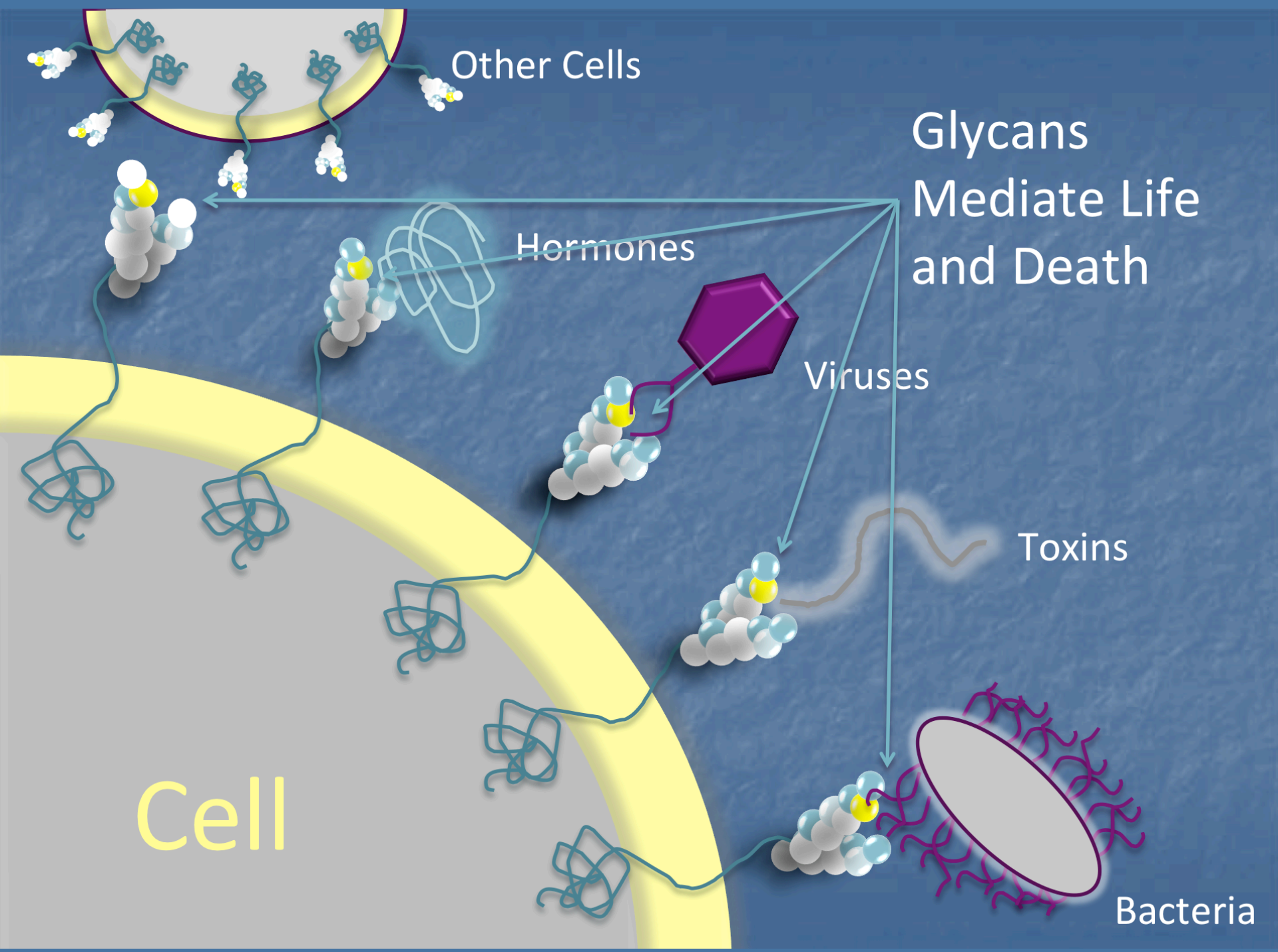
3D Building & Displaying Complex Carbohydrates

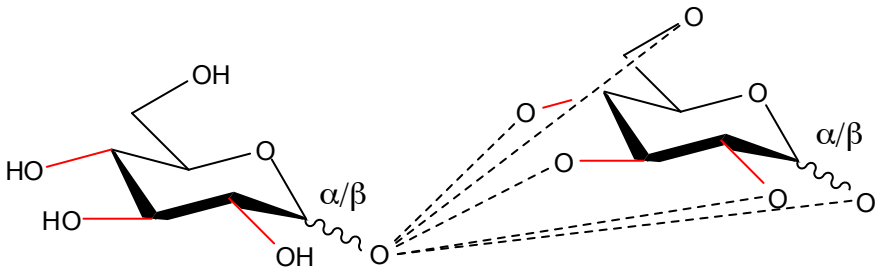
Robert J. Woods

Complex Carbohydrate Research Center
University of Georgia

www.glycam.org

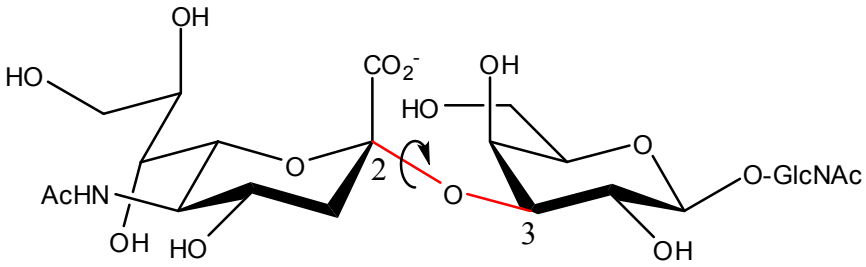

Other Cells
Glycans Mediate Life and Death
Hormones
Viruses
Toxins
Cell
Bacteria

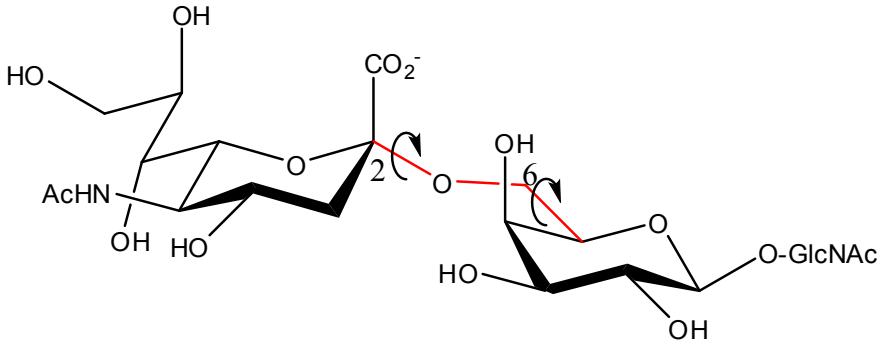
Carbohydrates form Nonlinear (Branched) Oligomers



Glc-α-(1-4)-Glc (Starch)
Glc-β-(1-4)-Glc (Cellulose)

The same two amino acids → 1 possible peptide
The same two monosaccharides → 20 possible disaccharides



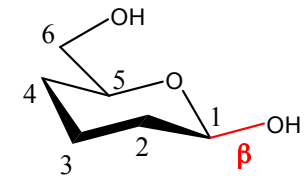


Avian Influenza A Receptor
Neu5Acα-(2-3)-Gal

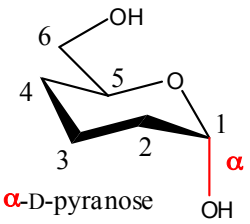
versus

Human Influenza A Receptor
Neu5Acα-(2-6)-Gal

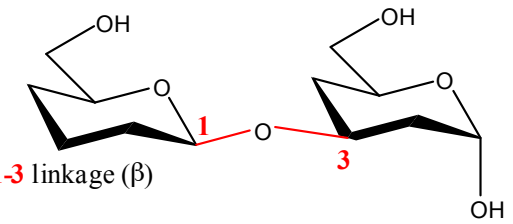
Monosaccharide Nomenclature…



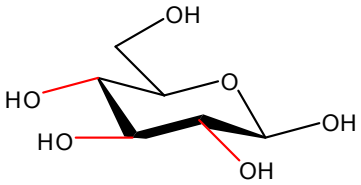
β-D-pyranose



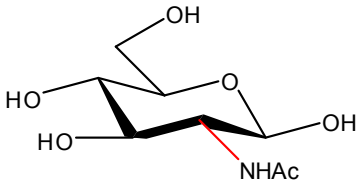
α-D-pyranose



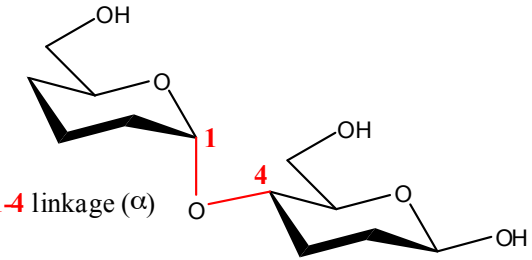
1-3 linkage (β)



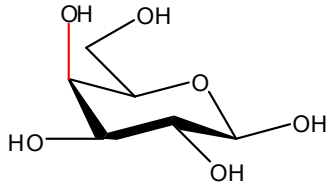
β-D-**gluco**pyranose, β-D-Glc



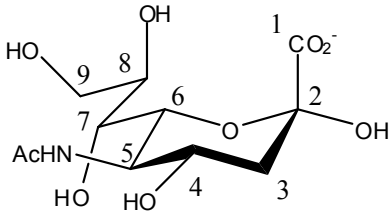
2-*N*-acetyl-β-D-glucopyranose, β-D-GlcNAc



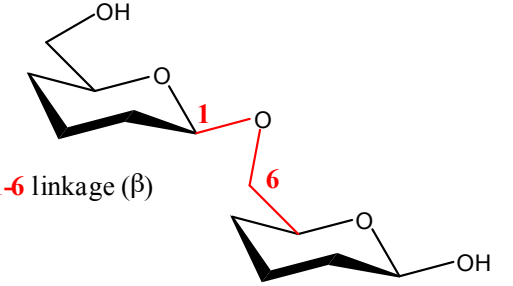
1-4 linkage (α)



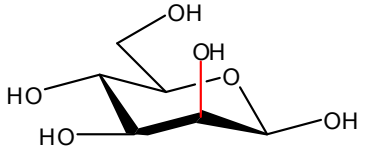
β-D-**galacto**pyranose, β-D-Gal



5–*N*-acetyl-α-neuraminic acid, α-Neu5Ac

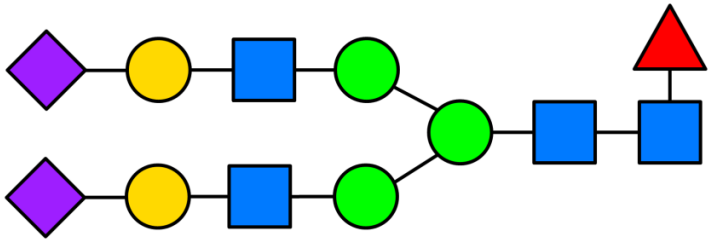


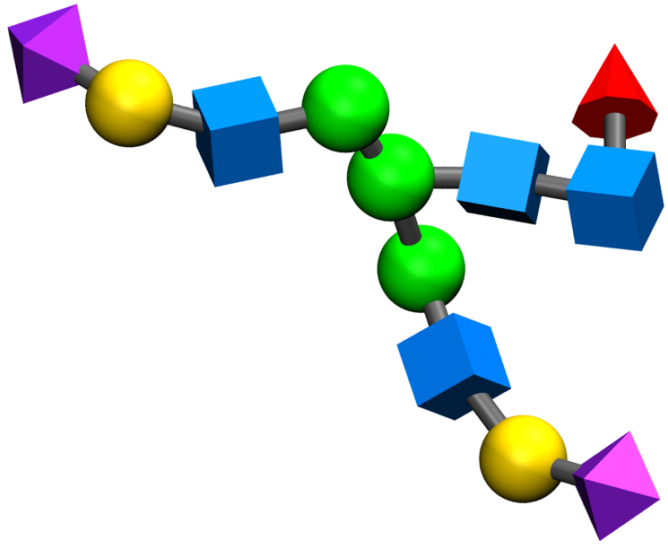
1-6 linkage (β)

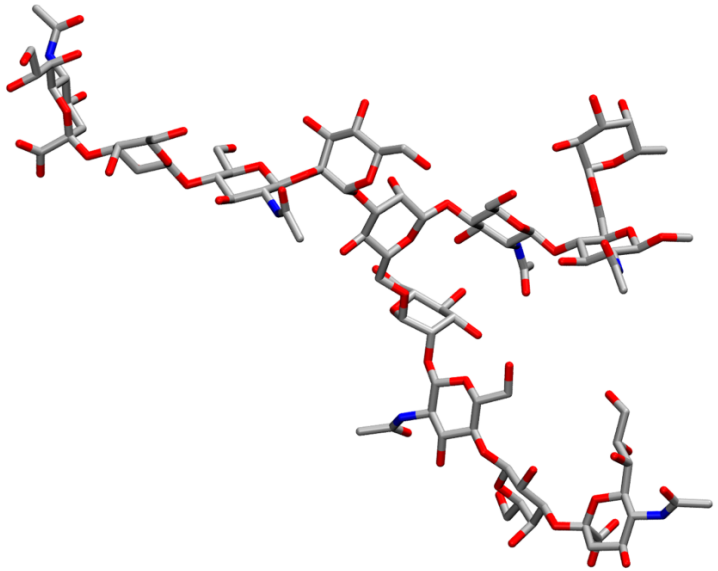


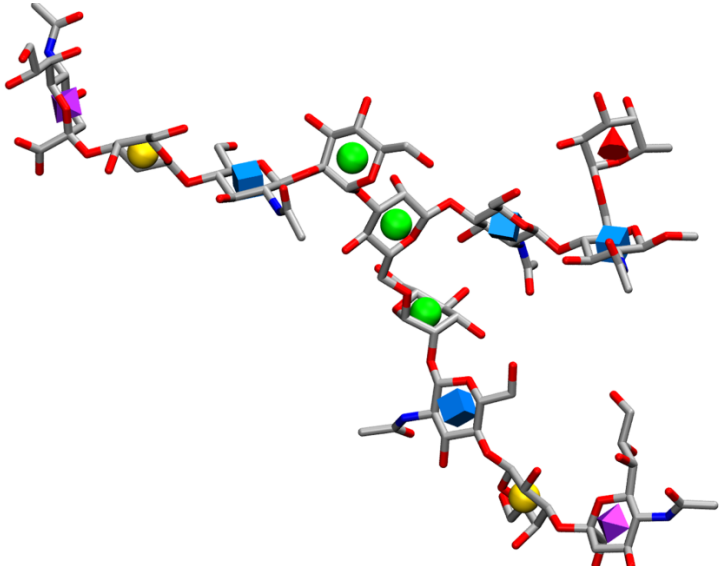
β-D-**manno**pyranose, β-D-Man

Symbol Representations for Carbohydrates



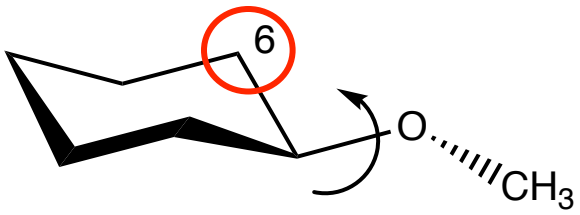


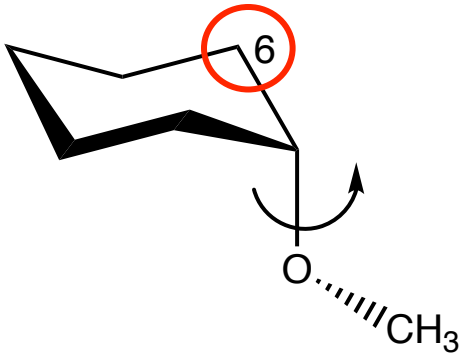


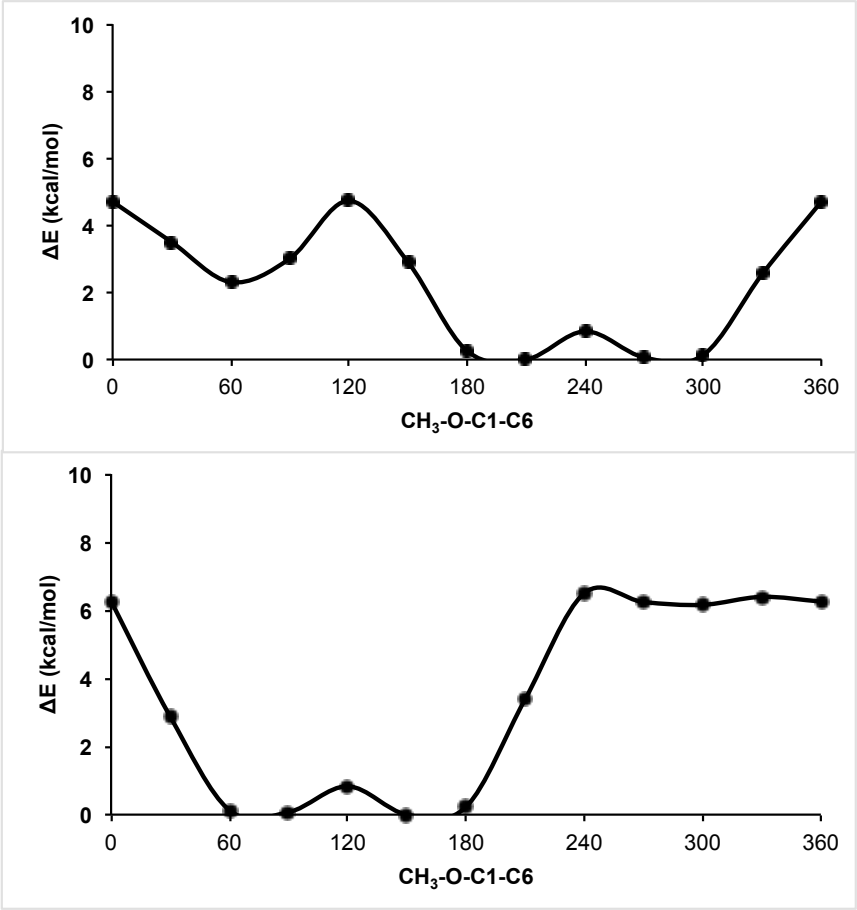


Thieker, et al. Glycobiology (2016) 26: 786-787

Glycosidic Linkages Have Unique Properties

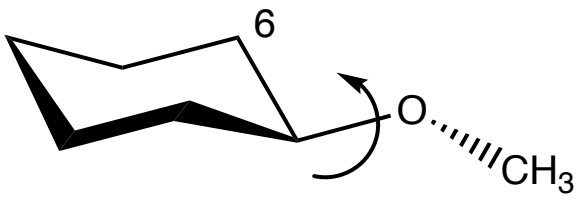


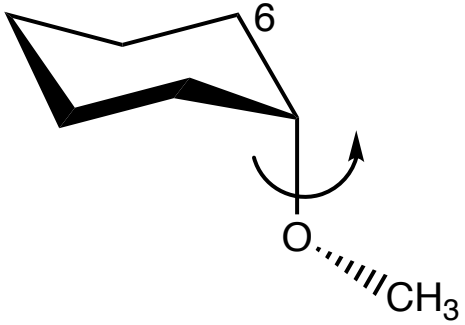


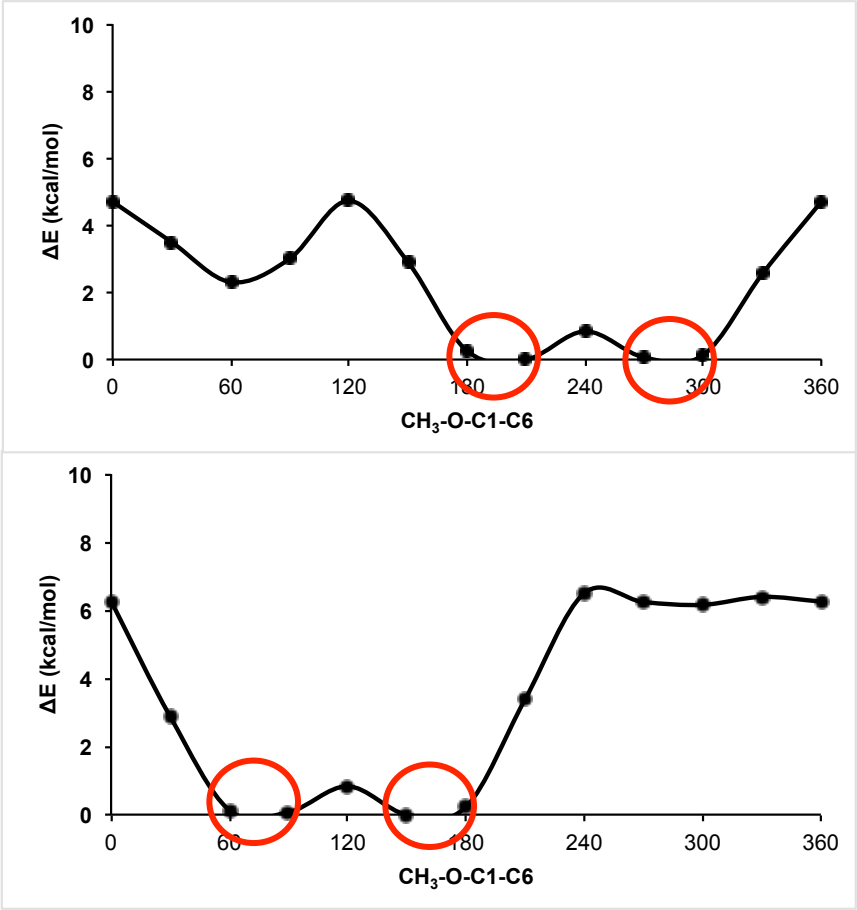


Woods, Chem. Rev., 2018, In Press

Glycosidic Linkages Have Unique Properties

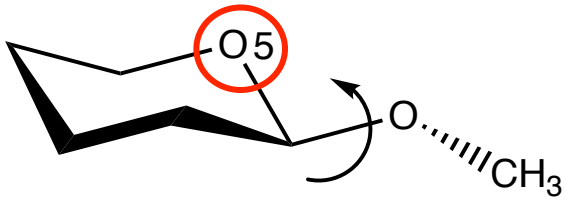


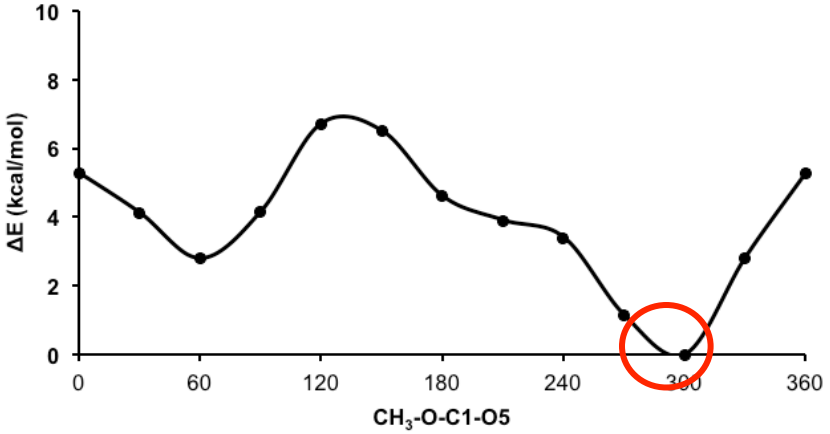


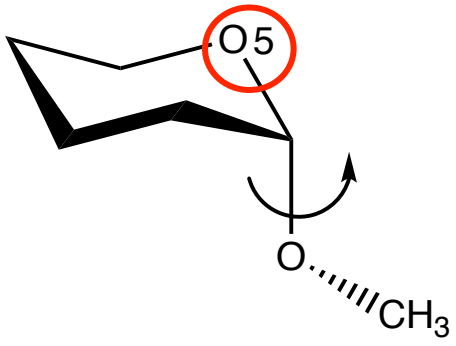


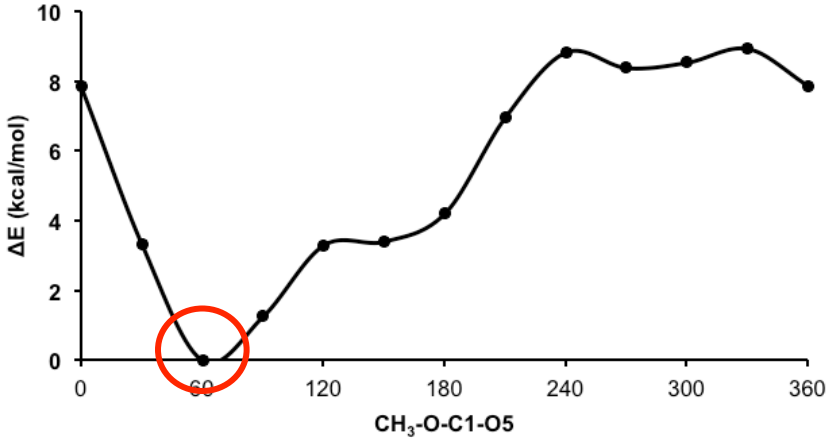
Woods, Chem. Rev., 2018, In Press

Glycosidic Linkages Have Unique Properties



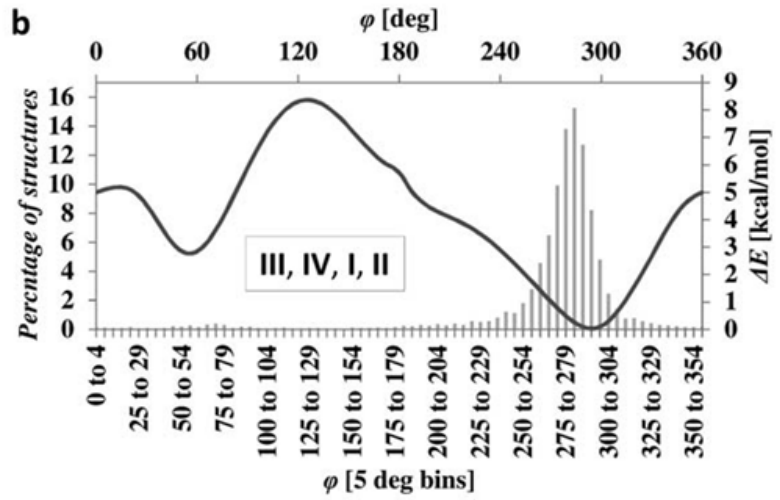


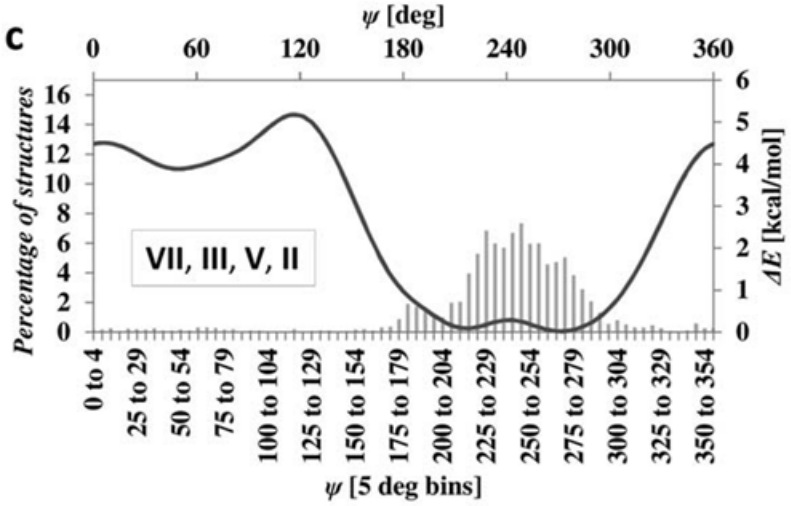




Woods, Chem. Rev., 2018, In Press

Theoretical Energies Compared to PDB Values



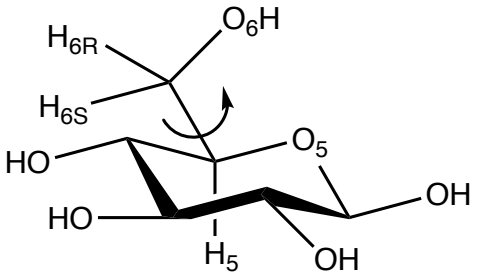


- Calculated energy curve (heavy line)
- PDB values as histograms
- Glycosidic linkages typically occupy one rotamer

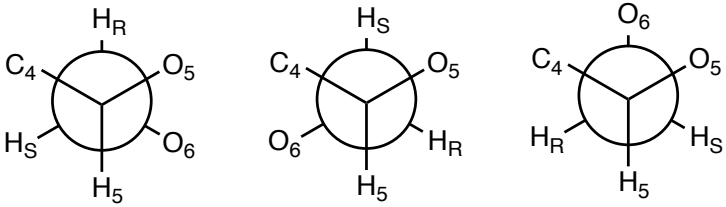
- If glycosidic linkages generally adopt only one conformation, why are oligosaccharides frequently said to be highly flexible?

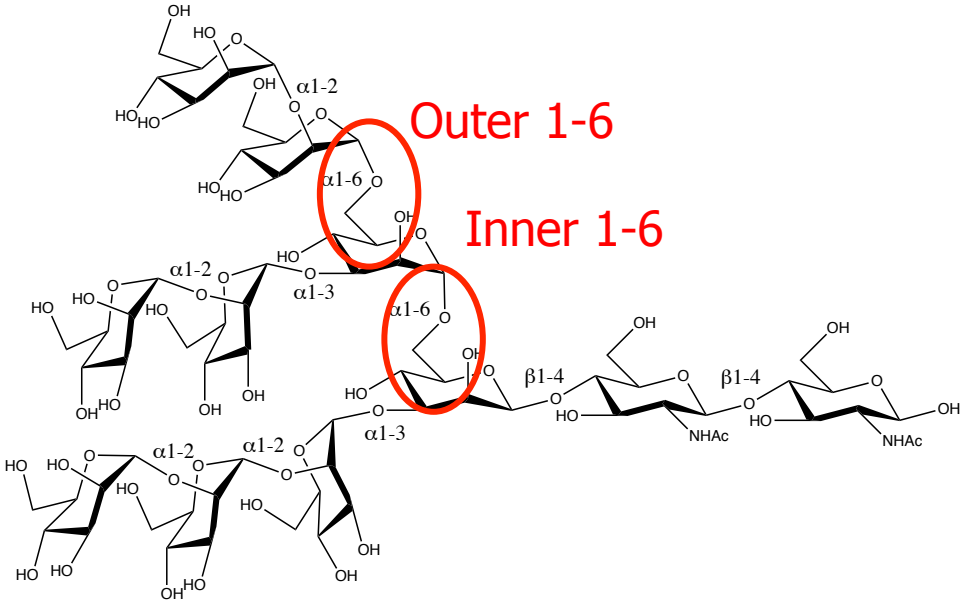
Nivedha et al. J. Comput. Chem. (2014) *35*, 526-539

3-Bond Linkages Generate Multiple Rotamers (Flexibility)

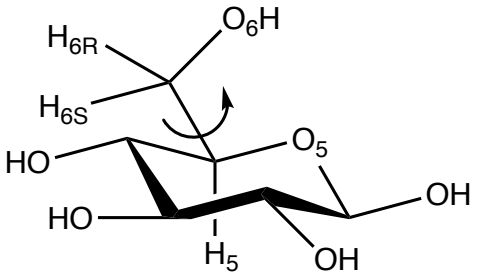


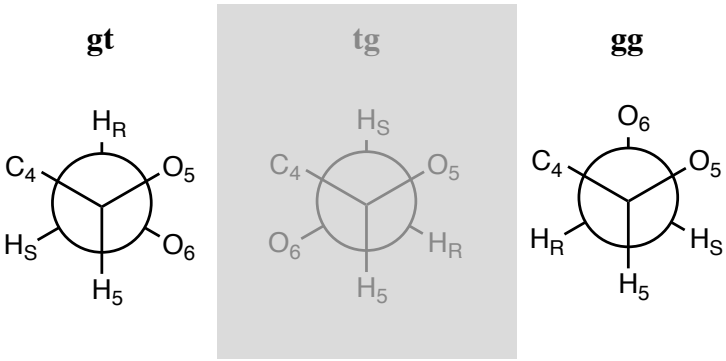
gt **tg** **gg**

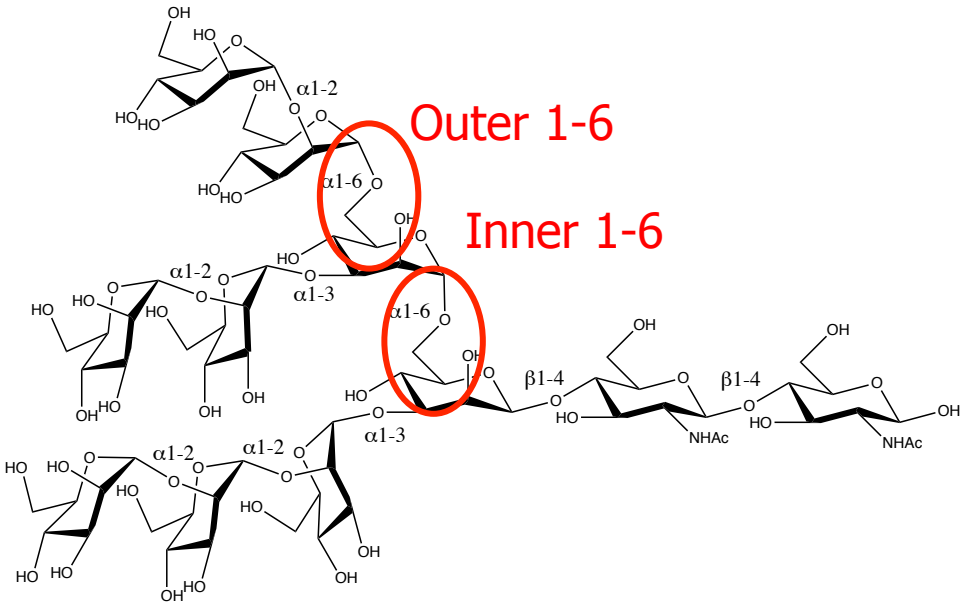




3-Bond Linkages Generate Multiple Rotamers (Flexibility)

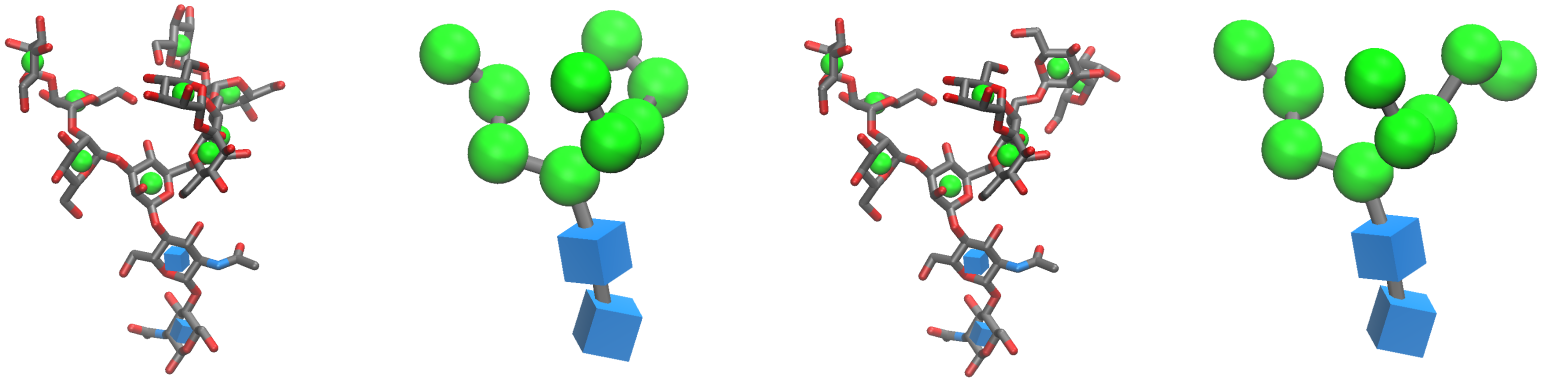




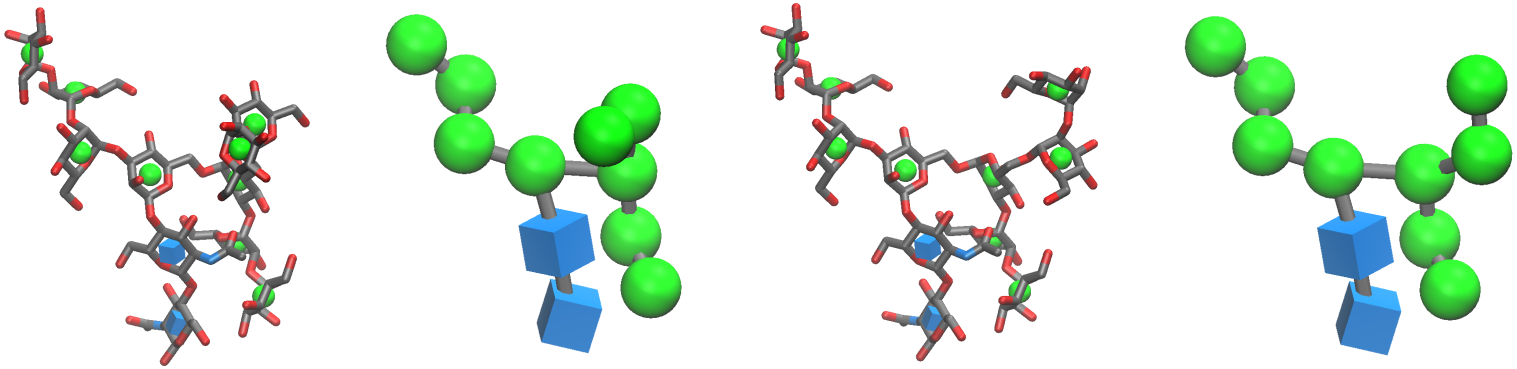


High-Mannose Glycan
Only 4 rotamers (2x2)

3-Bond Linkages Generate Multiple Rotamers (Flexibility)

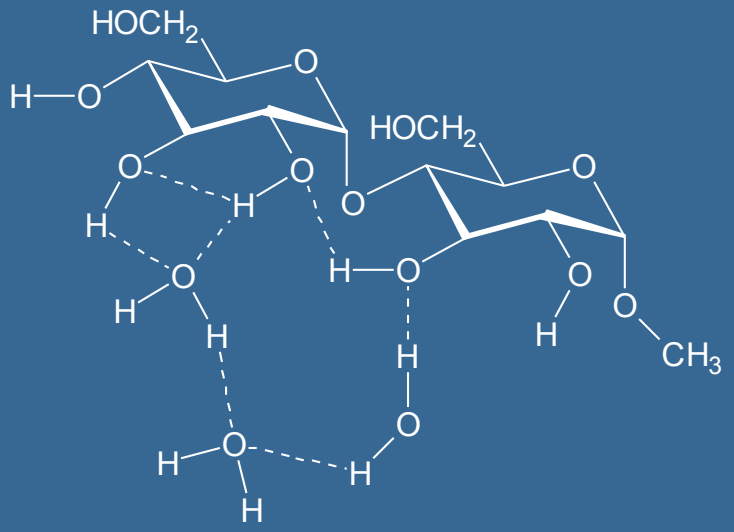


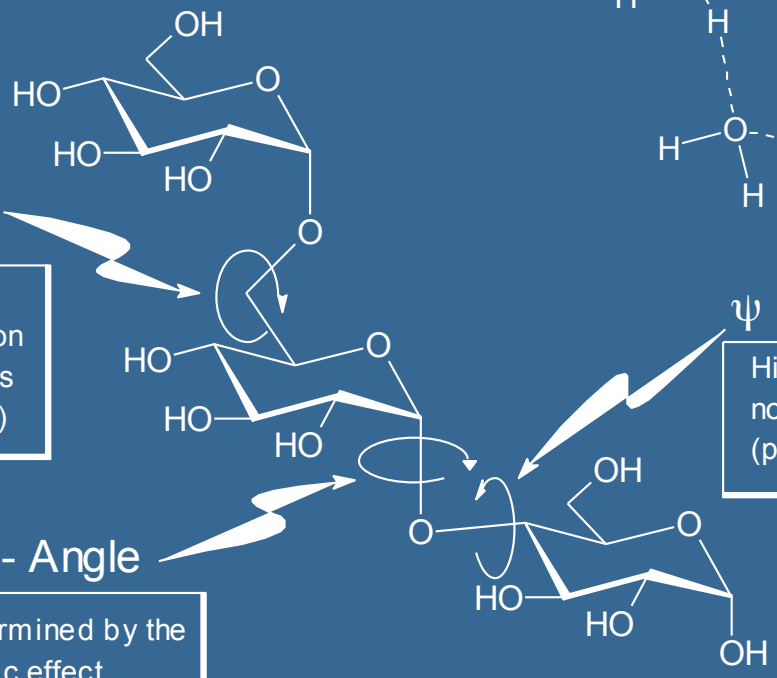
www.glycam.org



Factors Affecting Oligosaccharide Conformation

Balance of forces between water-carbohydrate interactions and internal forces





ω - Angle

Gauche preference determined by solvation (partial atomic charges and hydrogen bonding)

ψ - Angle

Highly influenced by non-bonded interactions (partial atomic charges)

ϕ - Angle

Largely determined by the *exo*-anomeric effect (torsional terms)

J. Phys. Chem. A (2001) 105: 4150–4155
Proc. Natl. Acad. Sci., USA (2001) 98, 10541-10545

GLYCAM/AMBER Classical Force Field

$$V_{total} = \sum_{bonds} K_r (r - r_{eq})^2 + \sum_{angles} K_\theta (\theta - \theta_{eq})^2 + \sum_{dihedrals} \sum_n \frac{V_n}{2} \left[1 + \cos(n\phi - \gamma_n)\right]$$

bond stretching + angle bending + non-classical internal rotations

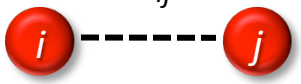
$$+ \sum_{\substack{non-bonded \\ i<j}} \left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{C_{ij}}{R_{ij}^{6}} + \frac{q_i q_j}{\varepsilon R_{ij}} \right]$$

+ van der Waals + electrostatics

GLYCAM: K_r K_θ V q derived from ab initio quantum data (cc-pVTZ)
A C derived empirically by fitting to bulk liquid properties ΔH_{vap}, ρ
r_{eq} θ_{eq} from neutron diffraction or QM data

K. Kirschner, A. Yongye, R.J. Woods. **2007**, *J. Comput. Chem., 33*, 622
K. Kirschner, R.J. Woods. **2001**. *Proc. Natl. Acad. Sci. USA 98*: 10541
R.J. Woods, R.A. Dwek, C.J. Edge, B. Fraser-Reid. **1995**. *J. Phys. Chem. 99*: 3832

Force Field Potential Functions

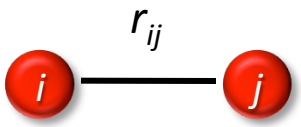


$$V_{vanderWaals} = 4\varepsilon\left[\left(\frac{\sigma_{ij}}{R_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{R_{ij}}\right)^{6}\right]$$

(*John Lennard-Jones – **1931)***

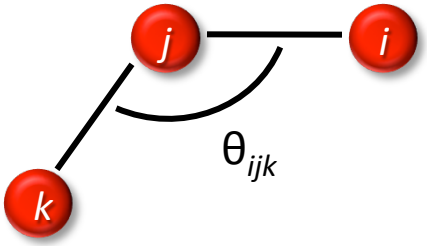
$$V_{Electrostatic} = \frac{q_i q_j}{4\pi\varepsilon R_{ij}}$$

(*Charles Augustin de Coulomb -**1785)***

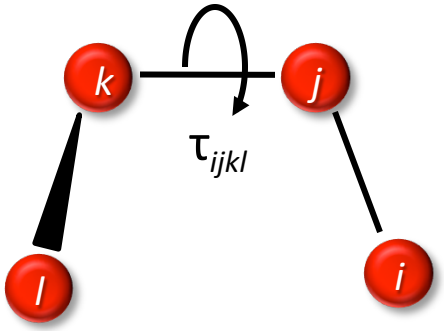


$$V_{bonds} = \frac{1}{2}k_r^{ij}\left(r_{ij} - r_{ij}^0\right)^2$$

(*Robert Hooke - **1660)***



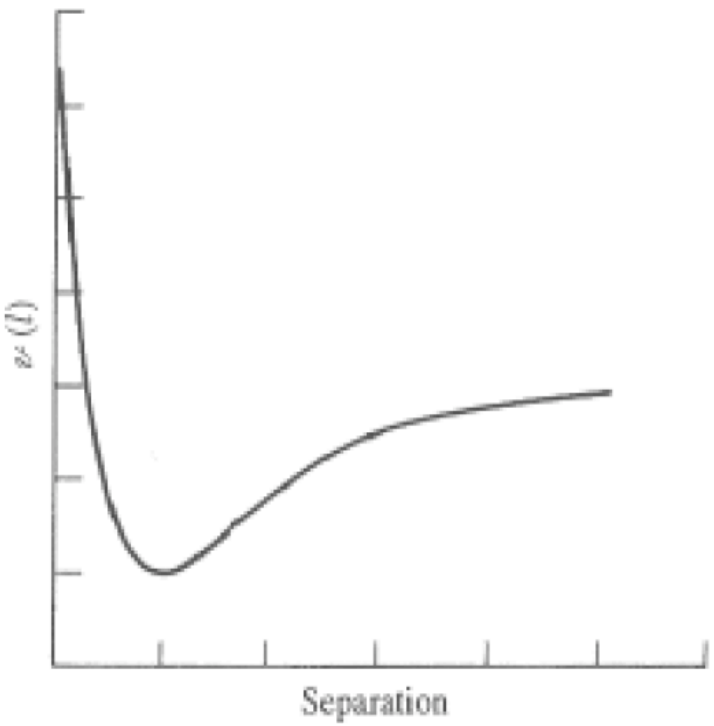
$$V_{angles} = \frac{1}{2}k_\theta^{ijk}\left(\theta_{ij} - \theta_{ij}^0\right)^2$$

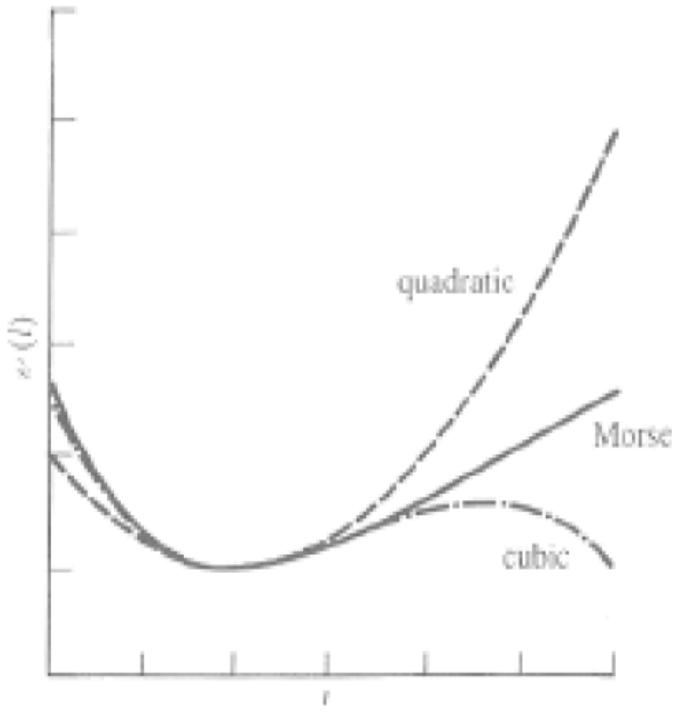


$$V_{torsions} = \frac{1}{2}\sum_n k_n^{ijkl}\left(1 - \cos(n\tau)\right)$$

(*Jean Baptiste Joseph Fourier – **1822)***

The Models are NOT Reality





In reality, bonds can break if stretched, but they cannot break when using a harmonic model

This is the essence of Molecular Mechanics; no bonds are broken or formed, therefore molecular mechanics is rarely used for modeling chemical reactions

Classical MD Simulation

Given the position of a particle at time, its new position after time Δt is described by the familiar Taylor expansion:

$$x(t+\Delta t) = x(t) + v(t)\Delta t + \frac{1}{2}a(t)\Delta t^2 + \ldots$$

If we knew the forces on the atoms (F) we could compute the atomic accelerations (*a*) from Newton's second law:

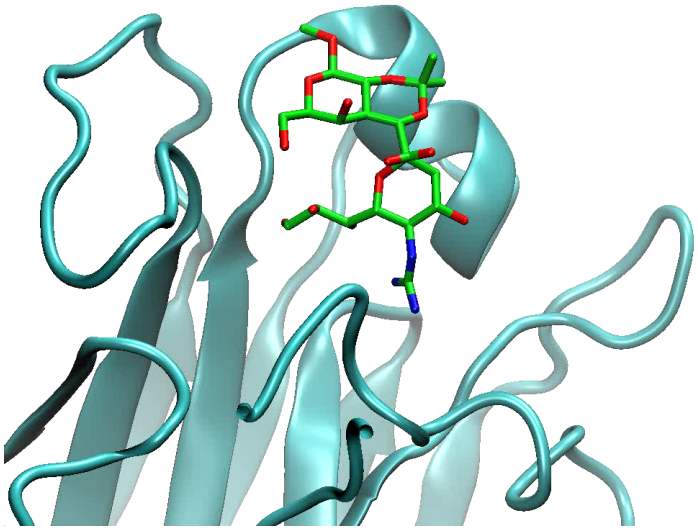
$$F_i = -\frac{\partial V}{\partial x_i} = m_i a_i$$

For which we need the derivative of the potential energy (V), as defined by the force field:

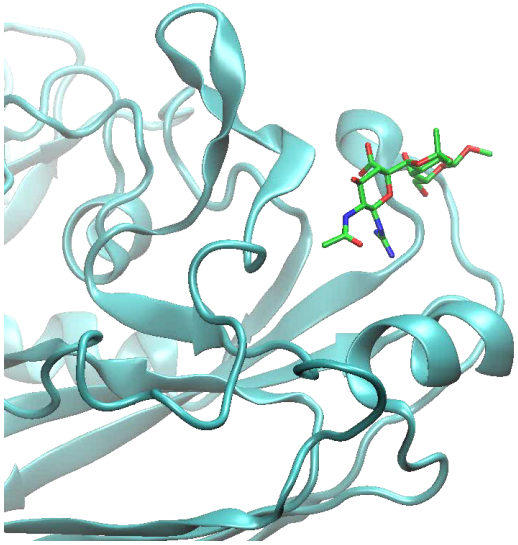
$$V_{total} = \sum_{bonds} K_r (r - r_{eq})^2 + \sum_{angles} K_\theta (\theta - \theta_{eq})^2 + \sum_{dihedrals} \sum_n \frac{V_n}{2}\left[1 + \cos(n\phi - \gamma_n)\right]$$

$$+ \sum_{\substack{non-bonded \\ i<j}} \left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{C_{ij}}{R_{ij}^{6}} + \frac{q_i q_j}{\varepsilon R_{ij}} \right]$$

MD Removes Investigator Bias

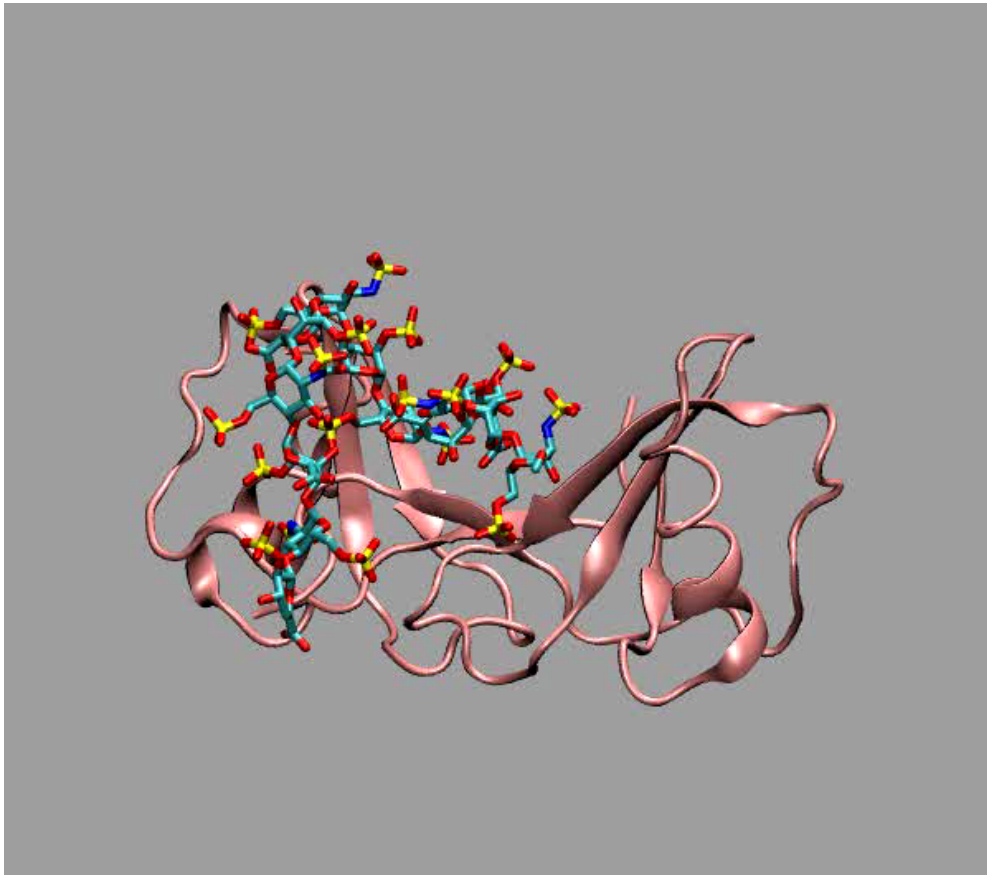


Putative Inhibitor (5-GAN) Complex 1

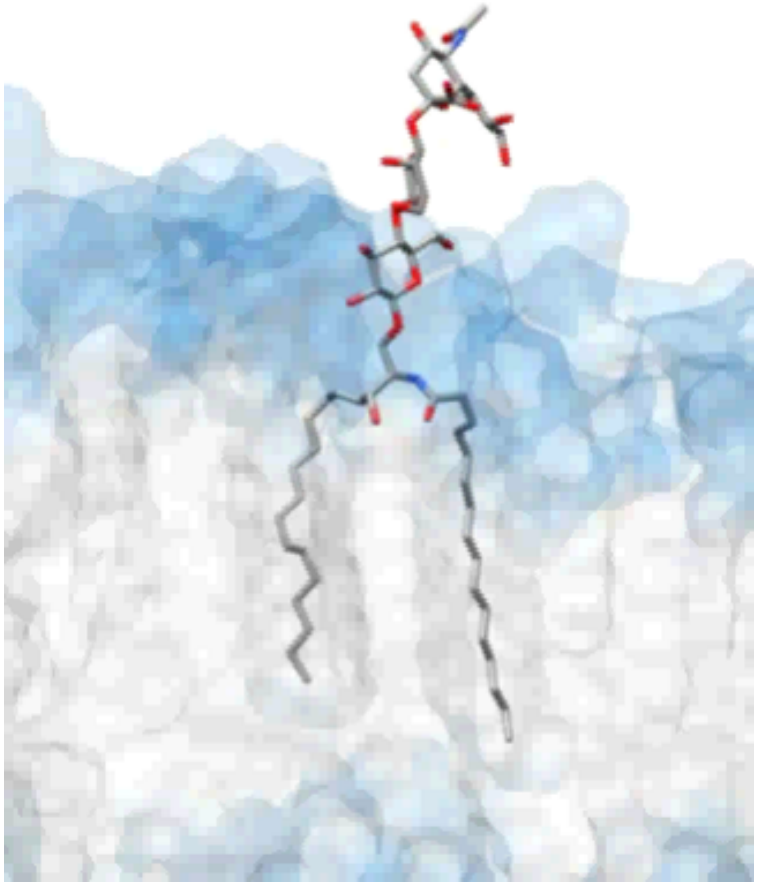


Putative Inhibitor (6-GAN) Complex 2

Complex Systems can be Simulated



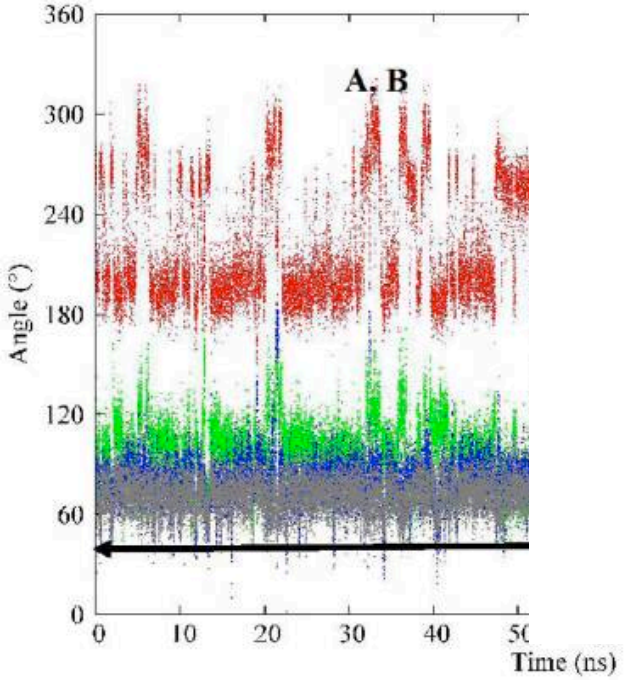
Protein-heparin complex

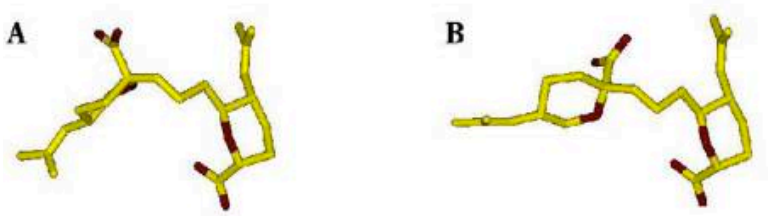


Glycolipid-Membrane Complex

How Long is Long Enough?

Look For Statistical Convergence in Key Properties

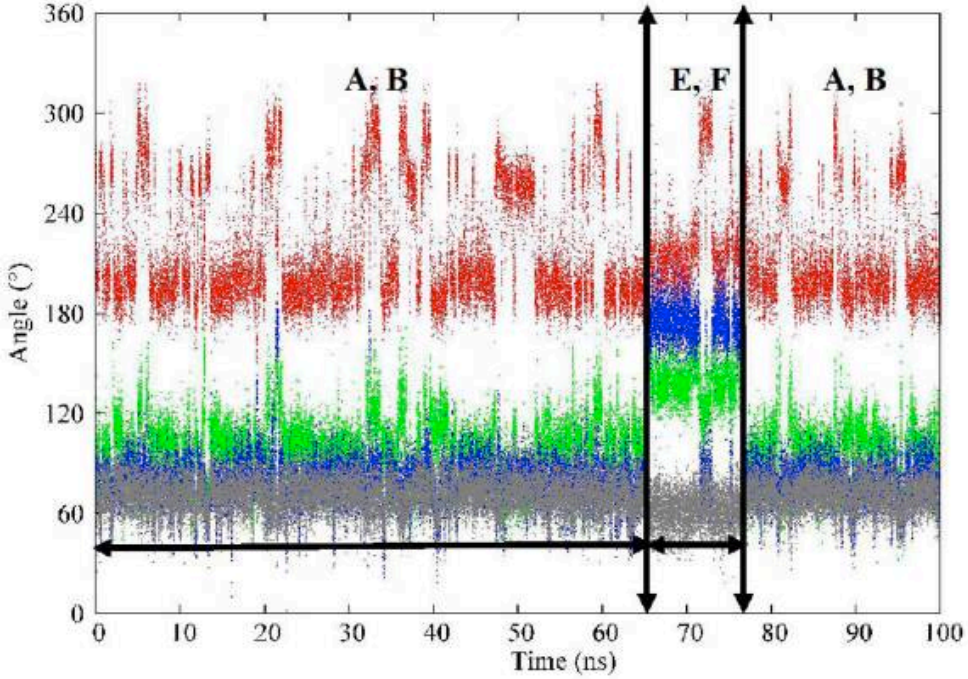


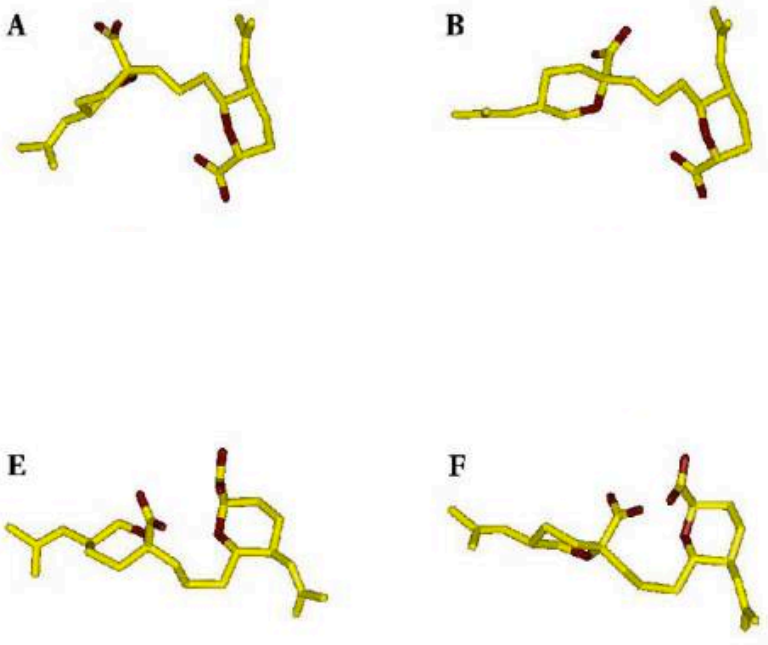


Yongye, A. B., et al. *Biochemistry* (2008) **47**: 12493–12514.

How Long is Long Enough?

Look For Statistical Convergence in Key Properties





Yongye, A. B., et al. *Biochemistry* (2008) **47**: 12493–12514.

Modeling Carbohydrate Binding: 3 Examples

1: Defining the conformational epitope of a "rigid" bacterial polysaccharide: group B *Streptococcus*, type III (GBSIII)

Kadirvelraj, R., et al. *PNAS* (2006) 103: 8149-8154.

2: Defining the conformational epitope of a "flexible" bacterial polysaccharide: Neisseria B (NmB)

Yongye, A. B., et al. *Biochemistry* (2008) 47: 12493–12514.

3: Quantifying key carbohydrate/protein residues responsible for antigenicity (mAb CS-35 – mycobacterial polysaccharide)

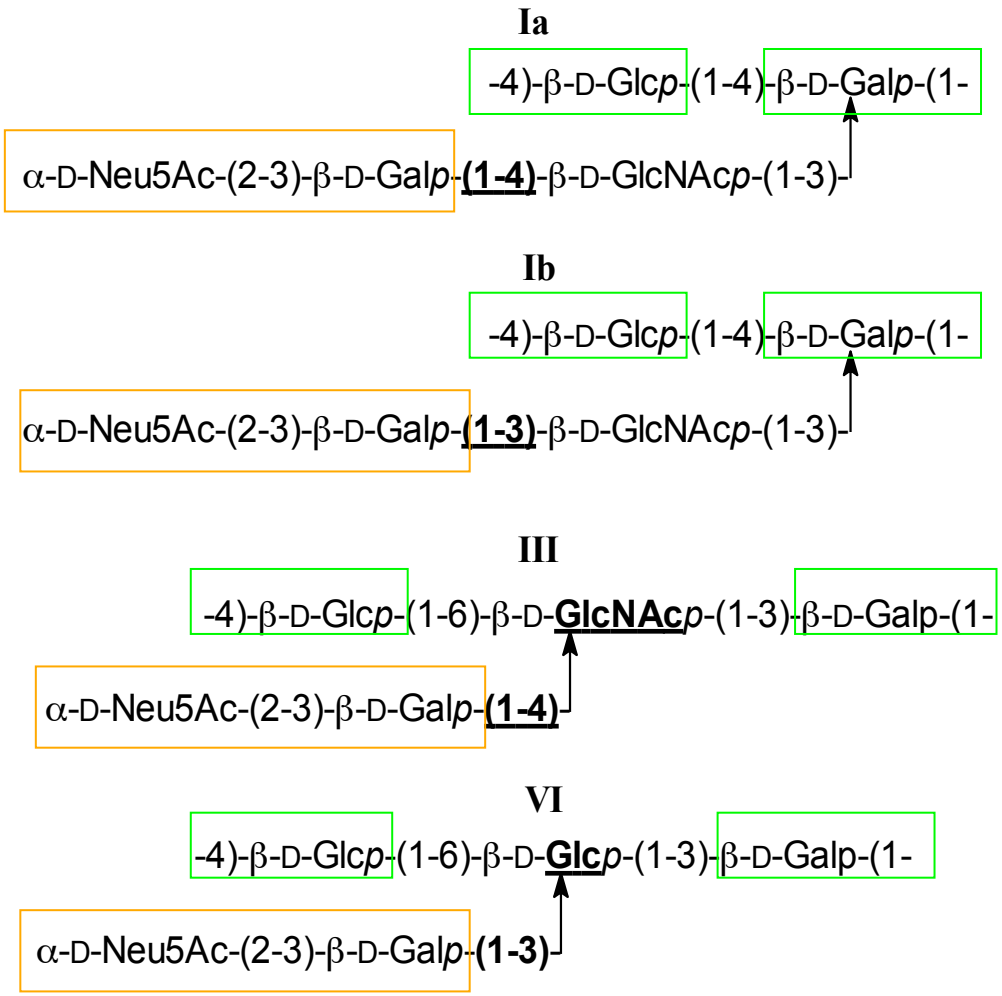
Lak et al. *Chem. Eur. J.* (2015) 21: 1138-1148

Example 1: Group B *Streptococcus* (GBS)

Gram-positive bacterium responsible for most cases of bacterial sepsis and meningitis in newborns and infants

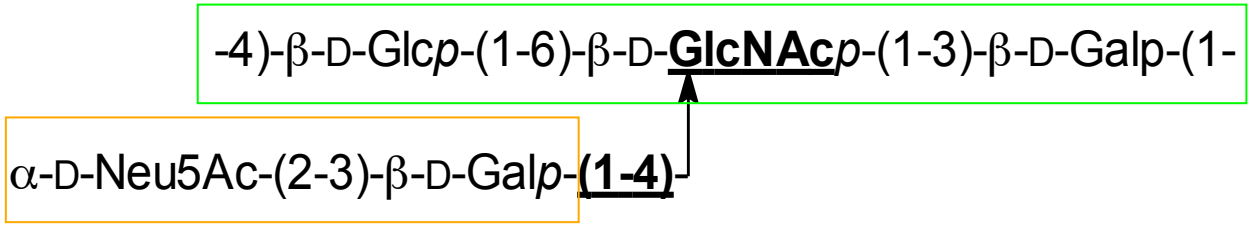
GBS *serotypes are defined by the carbohydrate sequence* and linkages

i.e. Despite sequence similarities, antibodies don't cross-react

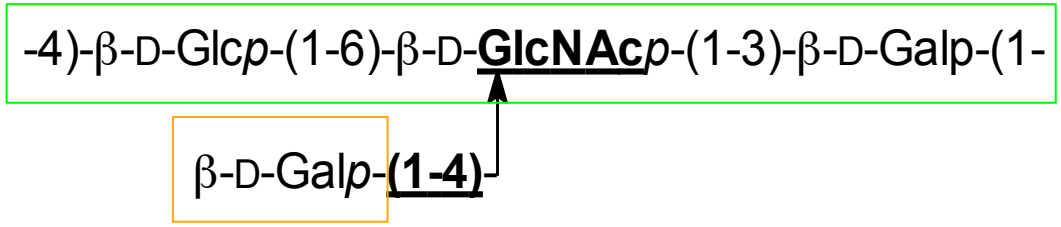


Similar Composition ≠ Similar 3D Shape

GBS III Polysaccharide



S. pneumoniae (Pn) 14 Polysaccharide



Loss of the Neu5Ac residues removes affinity for antibody

Each oligosaccharide antigen has unique shapes and properties

The shape of GBS III is not a random polymer

- A typical mAb combining site can fit only up to approximately 6 monosaccharides.
- To inhibit intact CPS from binding to mAb 1B1 requires an oligosaccharide fragment containing 3-7 repeating units (15-35 residues).
    - *Therefore GBS III has a conformational epitope*

Zou et al. J. Immunol. (1999)

- No experimental 3-D structure is available for either the CPS or the mAb (very typical situation)
- Can we use *Computational Methods* to develop a general approach to examine CPS-mAb interactions?

Computational Approach:

1. Generate experimentally-consistent 3D "structures" of the antigen (CPS) and mAb antibody (Fv fragment)

2. Generate a model for the immune complex (dock the CPS to the antibody Fv)

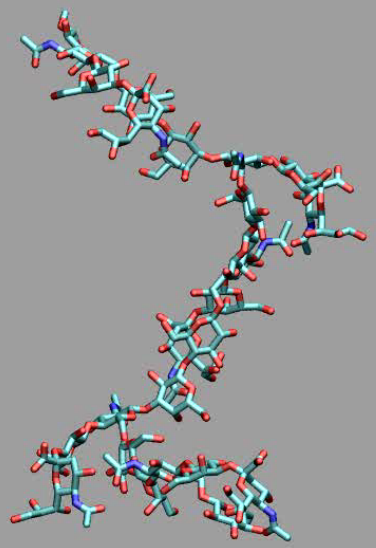
3. Simulate the dynamics of the Fv-CPS complex with *MD simulations*

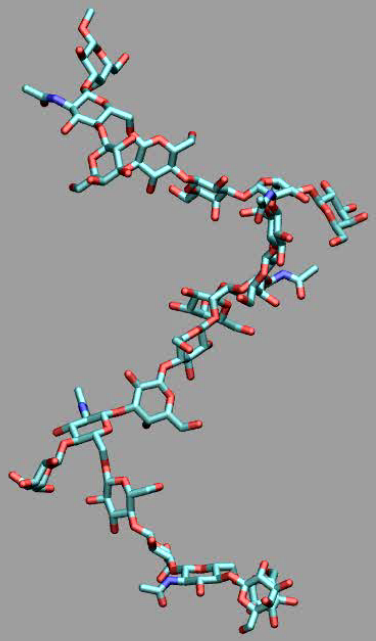
4. Examine the antigenicities in the presence and absence of the sialic acid residues (GBS III vs Pn 14)

MD Simulations of Native and Desialylated GBSIII

Native GBS III CPS

Desialylated CPS (Pn 14)





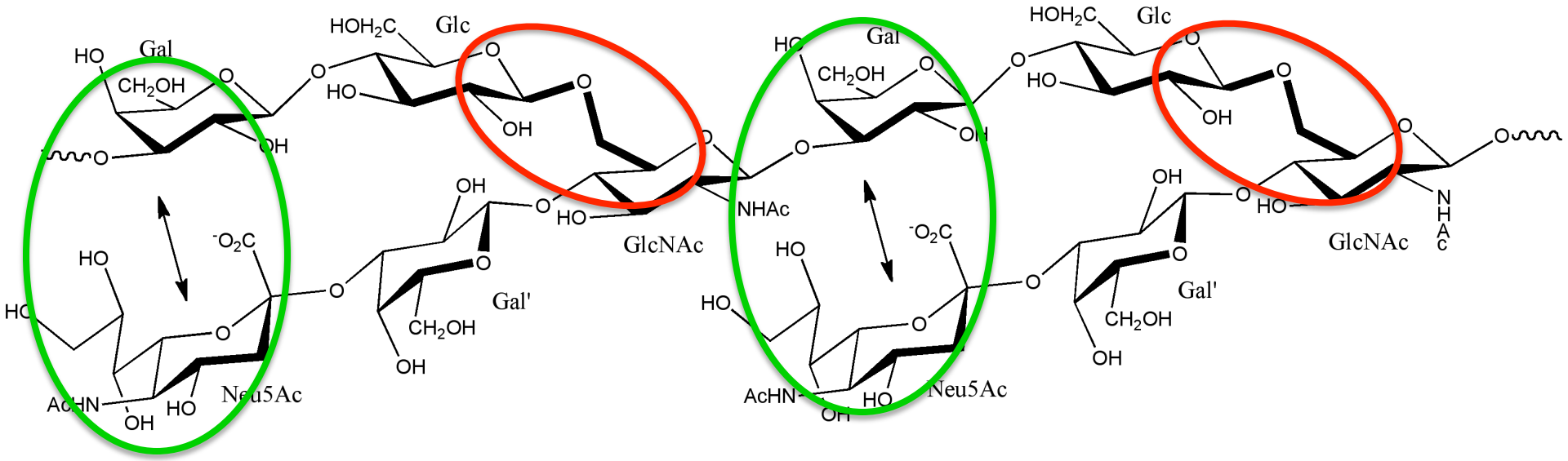
50 ns
5 – pentamer repeat units
6000 waters

Validation of MD: NMR $^3J_{CH}$ Coupling Constants

Linkage	Torsion Angle	^3J from NMR	^3J computed[a] from 50 ns MD
Gal(1-4)Glc	Φ	2.2	2.9 (1.2)
	Ψ	-	6.4 (0.5)
GlcNAc(1-3)Gal	Φ	4.6	4.0 (1.2)
	Ψ	3.7	5.5 (0.9)
Glc(1-6)GlcNAc	Φ	4.3	3.4 (1.2)
	Ψ	4.2	4.2 (1.7)
Gal(1-4)GlcNAc	Φ	4.0	3.8 (1.0)
	Ψ	5.3	6.4 (0.3)

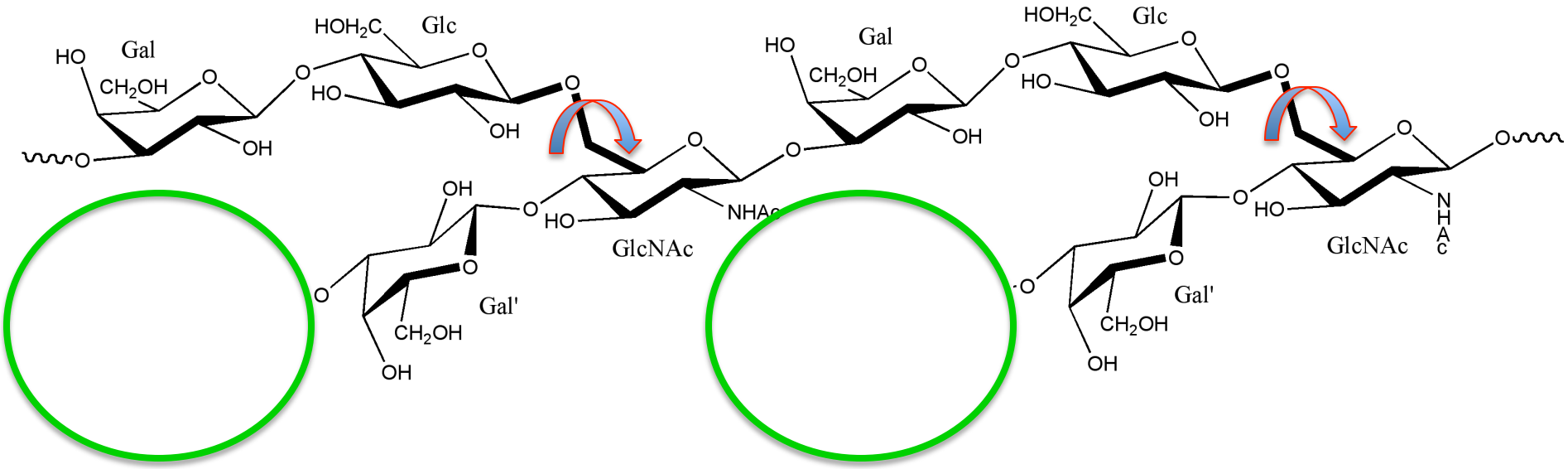
Brisson et al., *Biochemistry* (1997)
Gonzalez-Outeiriño et al., *Carbohydr Res.* (2005)

Neu5Ac Interactions Stabilize 1-6 Linkages



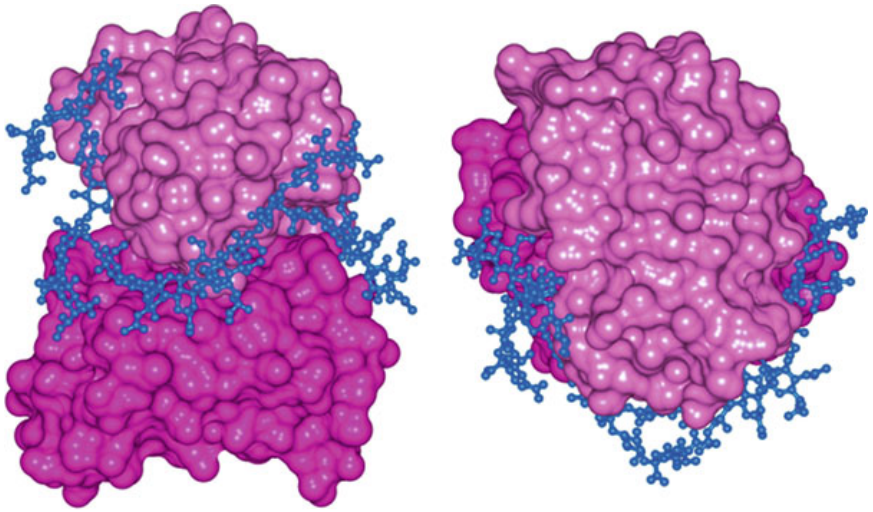
Two repeating units of GBS III polysaccharide indicating the side chain – back bone interactions that stabilize the 1,6-linkages

Internal 1-6 Flexibility Destroys Conformational Epitope

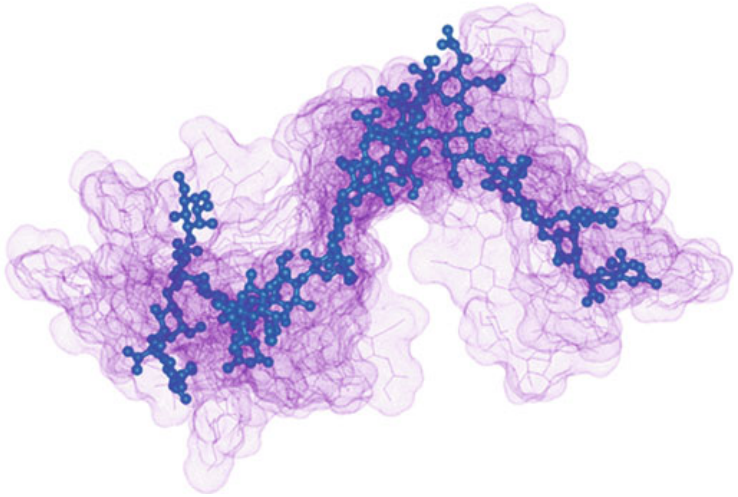


Removal of the Neu5Ac residues permits backbone rotation
Resulting in loss of 3D stability

Antigenic Conformation = Dominant Solution Conformation



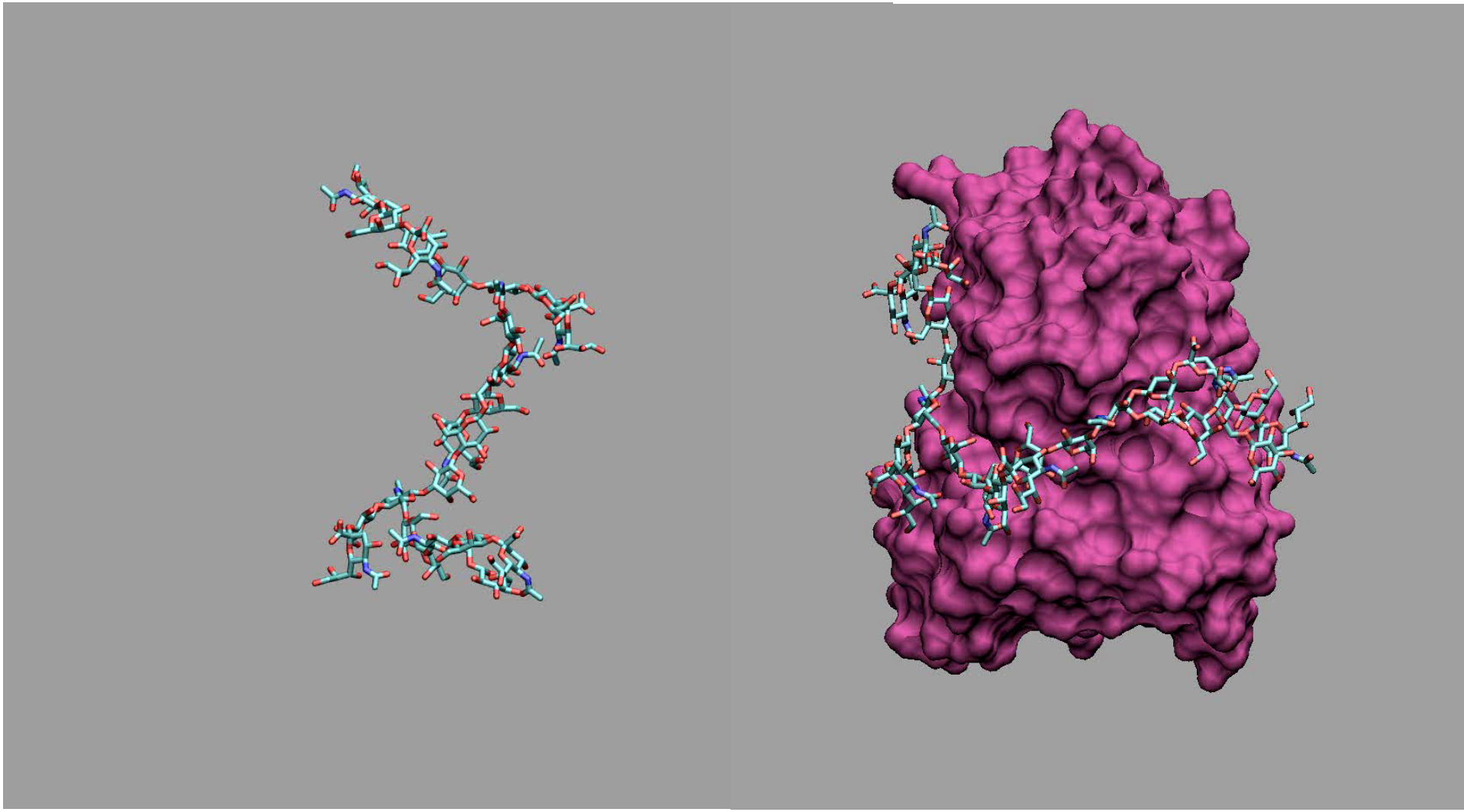
Predicted structure of the GBSIII polysaccharide bound to mAb 1B1



Bound conformation of GBS III overlayed with solution conformations taken from a 25 ns MD simulation

Dominant Solution Conformation = Immunogenic Conformation

MD Simulations of Free and Bound GBSIII CPS



Example 2: A Highly Flexible Bacterial Polysaccharide: *Neisseria meningitidis* serogroup B (NmB)

Table 15.1 CPS structures for the five most virulent *N. meningitidis* serotypes

Serotype	Oligosaccharide repeat units (anionic residues in bold)
A	→6)α-D-ManNAc(**1→OPO$_3$→**
B	→8)**αNeu5Ac**(2→
C	→9)**αNeu5Ac**(2→
E. coli K92	→8)α**Neu5Ac**(2→9)α**Neu5Ac**(2→
W135	→6)α-D-Gal(1→4)**αNeu5Ac**(2→
Y	→6)α-D-Glc(1→4)**αNeu5Ac**(2→

It is difficult to stimulate an immune response to NmB polysaccharide, possibly because it is the same as that in human neural cell adhesion glycoproteins (NCAM)

Can we generate a better vaccine if we chemically modify NmB?

Chemical modification of NmB enhances immunogenicity, but…

A protective immune response was elicited from a synthetic derivative of the B-conjugate vaccine, made by replacing the *N*-acetyl groups in the CPS with *N*–propionyl groups

B	→8)**α**Neu**5Ac**(2→

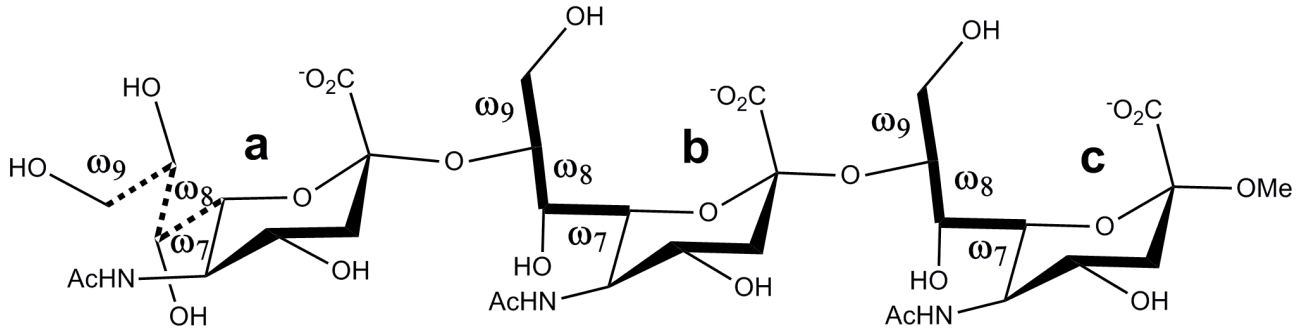
However, IgG class antibodies produced by the synthetic analog were unable to bind to the native antigen!

There is a trade-off between stimulating the immune system and compromising the 3D structure of the antigen

Defining the Conformational Flexibility of NmB

NMR J- and NOE data are not interpretable without a 3D model, but how can we be confident in the 3D models for such a complex highly flexible system?

Start by validating performance on well-defined di-/tri-saccharide fragments



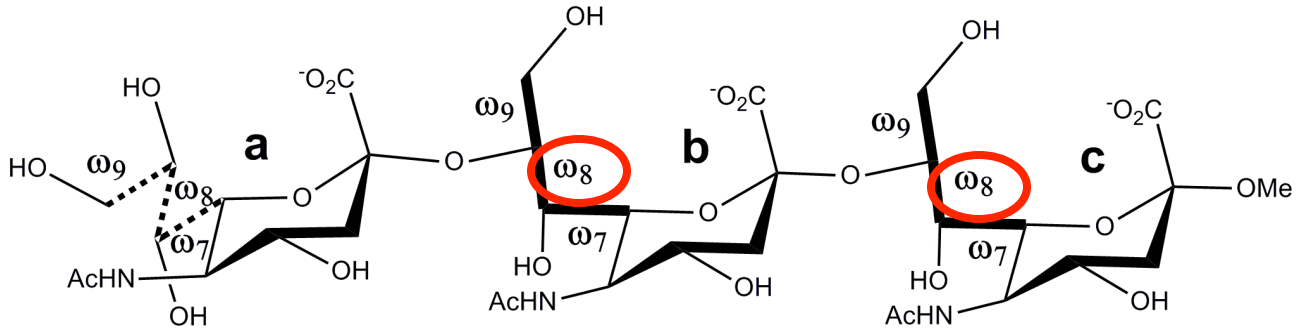
NMR and computed $^{3}J_{HH}$ coupling constants (Hz) for inter-residue torsion angles in *Nm*B sialotrioside.

Angle	Linkage	Spins	NMR	MD (100ns)
ω_7	Terminal	aH6-aH7	1.5 ± 0.2	1.0 ± 0.8
	Internal	bH6-bH7	< 1.0	0.9 ± 0.8
	Internal	cH6-cH7	< 1.0	1.1 ± 0.9
ω_8	Terminal	aH7-aH8	9.6 ± 1.0	7.7 ± 0.6
	Internal	bH7-bH8	< 4.0	3.6 ± 0.9
	Internal	cH7-cH8	< 4.0	2.1 ± 0.7

Defining the Conformational Flexibility of NmB

NMR J- and NOE data are not interpretable without a 3D model, but how can we be confident in the 3D models for such a complex highly flexible system?

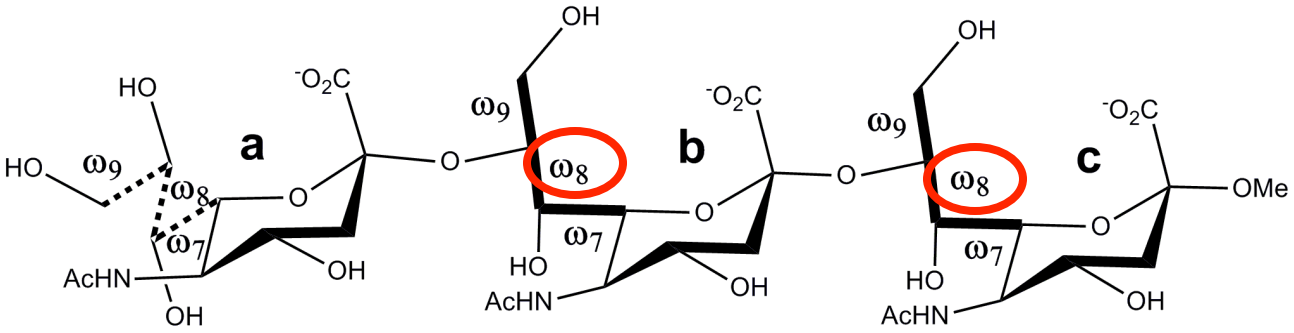
Start by validating performance on well-defined di-/tri-saccharide fragments



NMR and computed $^{3}J_{HH}$ coupling constants (Hz) for inter-residue torsion angles in *Nm*B sialotrioside.

Angle	Linkage	Spins	NMR	MD (100ns)
ω_7	Terminal	aH6-aH7	1.5 ± 0.2	1.0 ± 0.8
	Internal	bH6-bH7	< 1.0	0.9 ± 0.8
	Internal	cH6-cH7	< 1.0	1.1 ± 0.9
ω_8	Terminal	aH7-aH8	9.6 ± 1.0	7.7 ± 0.6
	Internal	bH7-bH8	< 4.0	3.6 ± 0.9
	Internal	cH7-cH8	< 4.0	2.1 ± 0.7

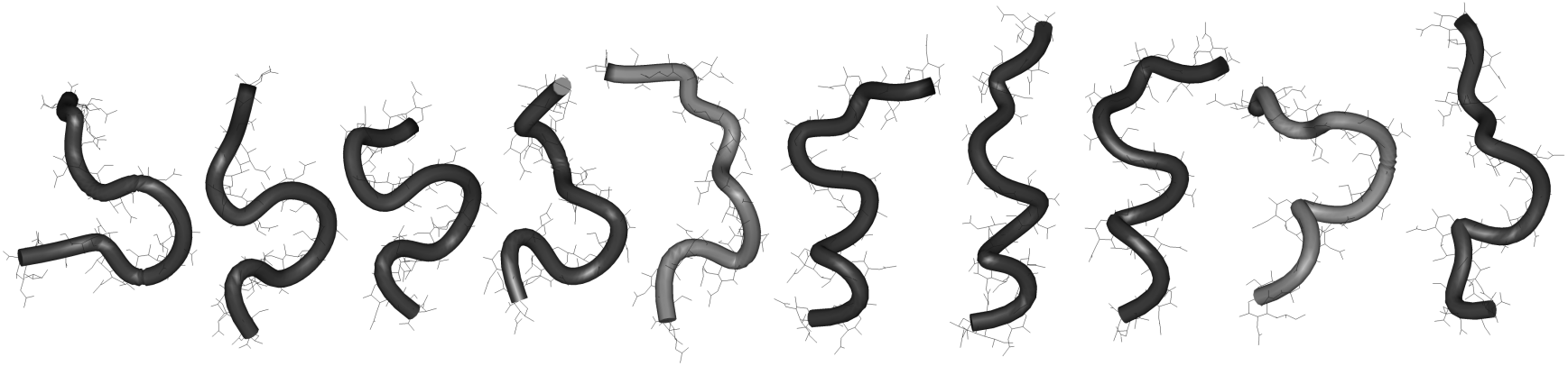
Defining Conformational Flexibility NMR/Simulation



Angle	Linkage	NMR[a, b] *gauche/trans/-gauche*	MD (100 ns) *g/t/-g*
ω_7	Terminal a	100/0/0	100/0/0
ω_7	Internal b	100/0/0	100/0/0
ω_7	Internal c	100/0/0	100/0/0
ω_8	Terminal a	0/100[c]/0	15/74/7
ω_8	Internal b	60 ± 9 / 25 ± 3 / 0	70 / 30 / 0
ω_8	Internal c	66 ± 5 / 31 ± 9 / 0	89 / 11 / 0

Internal linkage flexibility leads to conformational chaos

In contrast to the 'rigid' structure of GBSIII, NmB has no conformational epitope and is essentially 'over cooked spaghetti'

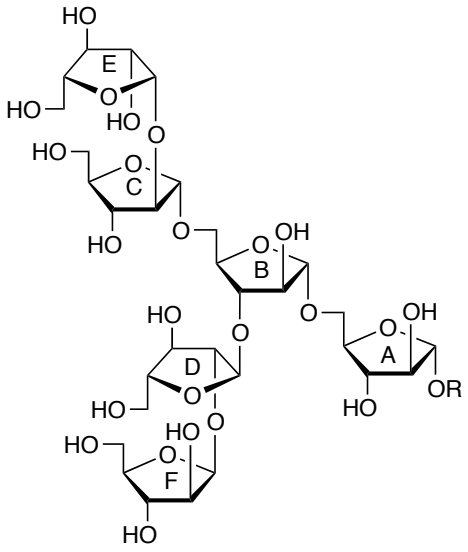


Snapshots from a 25 ns MD simulation of a 12-mer of *Nm*B selected at 2.5 ns intervals illustrate the plasticity of this polysaccharide.

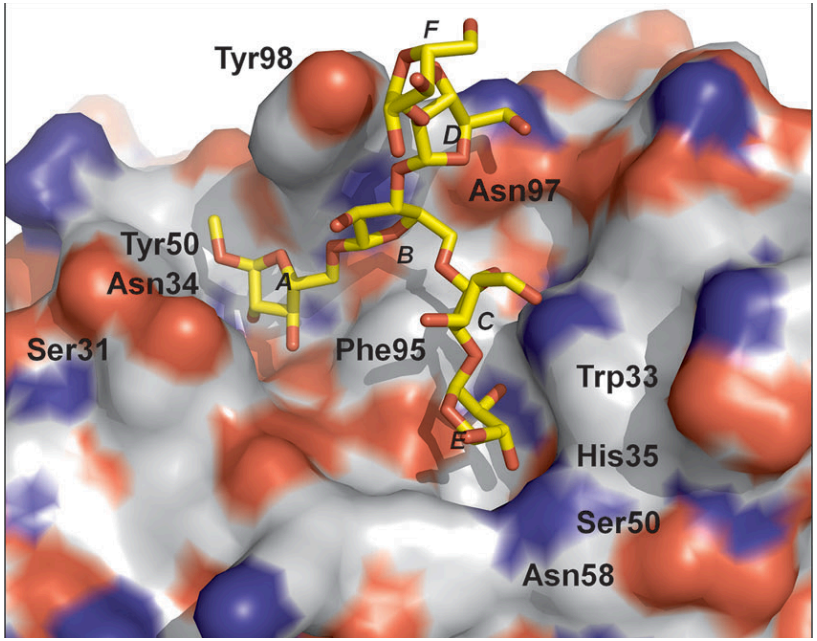
Multiple conformational states are present with regions of helical structure transitioning (grey) between more convoluted conformations.

The result is that only short sequences can be recognized by antibodies

Example 3: Quantifying Residues Responsible for Affinity



1, R = Lipoarabinomannan
2, R = OCH_3
3, R = $O(CH_2)_8NHAcyl$-CM5 Chip
4, R = $O(CH_2)_8NH_2$



Left: Structure of the terminal hexasaccharide motif in mycobacterial LAM that is recognized by CS-35, and three synthetic derivatives, 2–4. Right: Binding pocket of CS-35 Fab in complex with 2 (PDB ID: 3HNS). The highlighted residues were picked for mutation studies.

Murase et al. (2009) *Mol. Immunol.*

Interaction Energies Identify Key Residues

Residue	CDR[c]	H-Bond[d]	van der Waals	Electrostatic	Polar Desolvation	Non Polar Desolvation	Total
Protein							
Trp33	H1	Y	–4.2	–2.4	1.6	–0.5	–5.5
Tyr98	H3	Y	–4.6	–3.8	4.1	–0.8	–5.0
Asp91	L3	Y	0.6	–16.7	13.5	–0.1	–2.5
His35	H1	Y	–0.1	–3.4	1.2	0.0	–2.4
Asn97	H3	Y	–2.8	–2.6	4.2	–0.3	–1.5
Ser50	H2	Y	–0.1	–2.1	0.8	0.0	–1.5
Tyr96	L3	Y	–1.0	–0.8	1.0	–0.1	–0.9
Asn58	H2	Y	–0.6	–0.7	0.7	–0.1	–0.6
Phe95	H3	N	–3.7	–0.2	0.5	–0.4	–3.8
Val99	H3	N	–1.9	–0.3	0.7	–0.1	–1.5
Tyr50	L2	N	–1.9	0.4	0.3	–0.2	–1.5
Pro100	H3	N	–1.2	0.4	–0.3	0.0	–1.2
Pro94	L3	N	–1.0	–0.2	0.3	–0.1	–1.1
Tyr49	L2	N	–0.9	–0.2	0.2	0.0	–0.9
Tyr32	L1	N	–1.1	0.2	0.4	–0.1	–0.6
Tyr52	H2	N	–0.5	–0.3	0.4	–0.1	–0.5
Gly96	H3	N	–0.8	–0.6	0.9	0.0	–0.5
Subtotal			–25.8	–33.3	30.5	–2.9	–31.5
Ligand							
A	-	Y	–10.7	–29.7	27.5	–1.9	–14.9
B	-	N	–5.4	–11	11.9	–0.5	–5.1
E	-	Y	–7.4	–1.8	5.5	–1.4	–5.1
D	-	N	–2.8	–1.2	3.3	–0.6	–1.3
C	-	Y	–3.6	–0.6	3.4	–0.3	-1
F	-	N	–0.6	7.4	–6.5	–0.1	0.1
Subtotal			–30.6	–36.9	45.1	–4.8	–27.2

[a] In kcal/mol
[b] Residues that contribute greater than 0.5 kcal/mol to the total binding energy.
[c] Complementarity determining regions
[d] Intermolecular hydrogen bonds observed, based on a distance cut-off of 3.5 Å

The ability to partition the binding energy between the monosaccharide residues in the antigen is a unique strength of the computational analysis

The simulation indicated that approximately 93% of the affinity is provided by only three residues (A, 55%; B, 19%; E, 19%).

Lak et al. (2014) *Chem. Eur. J.*, **20**, 1-12.

Interaction Energies Identify Key Residues

Residue	CDR[c]	H-Bond[d]	van der Waals	Electrostatic	Polar Desolvation	Non Polar Desolvation	Total
Protein							
Trp33	H1	Y	–4.2	–2.4	1.6	–0.5	–5.5
Tyr98	H3	Y	–4.6	–3.8	4.1	–0.8	–5.0
Asp91	L3	Y	0.6	–16.7	13.5	–0.1	–2.5
His35	H1	Y	–0.1	–3.4	1.2	0.0	–2.4
Asn97	H3	Y	–2.8	–2.6	4.2	–0.3	–1.5
Ser50	H2	Y	–0.1	–2.1	0.8	0.0	–1.5
Tyr96	L3	Y	–1.0	–0.8	1.0	–0.1	–0.9
Asn58	H2	Y	–0.6	–0.7	0.7	–0.1	–0.6
Phe95	H3	N	–3.7	–0.2	0.5	–0.4	–3.8
Val99	H3	N	–1.9	–0.3	0.7	–0.1	–1.5
Tyr50	L2	N	–1.9	0.4	0.3	–0.2	–1.5
Pro100	H3	N	–1.2	0.4	–0.3	0.0	–1.2
Pro94	L3	N	–1.0	–0.2	0.3	–0.1	–1.1
Tyr49	L2	N	–0.9	–0.2	0.2	0.0	–0.9
Tyr32	L1	N	–1.1	0.2	0.4	–0.1	–0.6
Tyr52	H2	N	–0.5	–0.3	0.4	–0.1	–0.5
Gly96	H3	N	–0.8	–0.6	0.9	0.0	–0.5
Subtotal			–25.8	–33.3	30.5	–2.9	–31.5
Ligand							
A	-	Y	–10.7	–29.7	27.5	–1.9	–14.9
B	-	N	–5.4	–11	11.9	–0.5	–5.1
E	-	Y	–7.4	–1.8	5.5	–1.4	–5.1
D	-	N	–2.8	–1.2	3.3	–0.6	–1.3
C	-	Y	–3.6	–0.6	3.4	–0.3	-1
F	-	N	–0.6	7.4	–6.5	–0.1	0.1
Subtotal			–30.6	–36.9	45.1	–4.8	–27.2

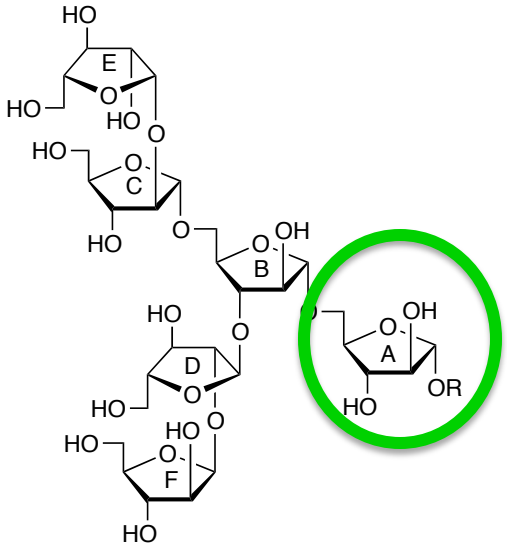
[a] In kcal/mol
[b] Residues that contribute greater than 0.5 kcal/mol to the total binding energy.
[c] Complementarity determining regions
[d] Intermolecular hydrogen bonds observed, based on a distance cut-off of 3.5 Å

The ability to partition the binding energy between the monosaccharide residues in the antigen is a unique strength of the computational analysis

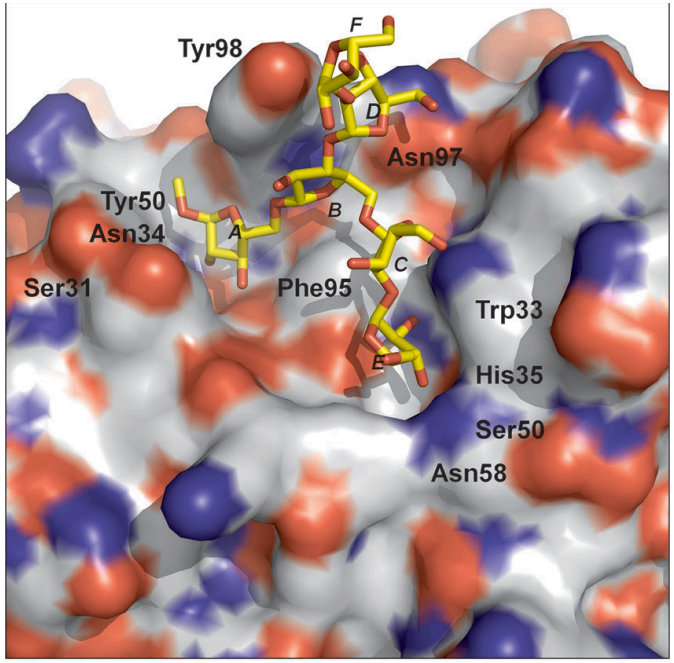
The simulation indicated that approximately 93% of the affinity is provided by only three residues (A, 55%; B, 19%; E, 19%).

Lak et al. (2014) *Chem. Eur. J.*, **20**, 1-12.

Per-residue Energy Decomposition Identifies Key Residues



1, R = Lipoarabinomannan
2, R = OCH_3
3, R = $O(CH_2)_8NHAcyl$-CM5 Chip
4, R = $O(CH_2)_8NH_2$

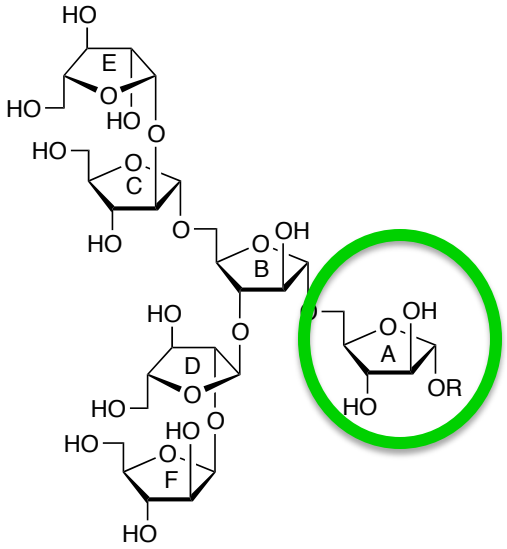


The energy analysis indicated that residue A is the immunodominant component of the antigen.

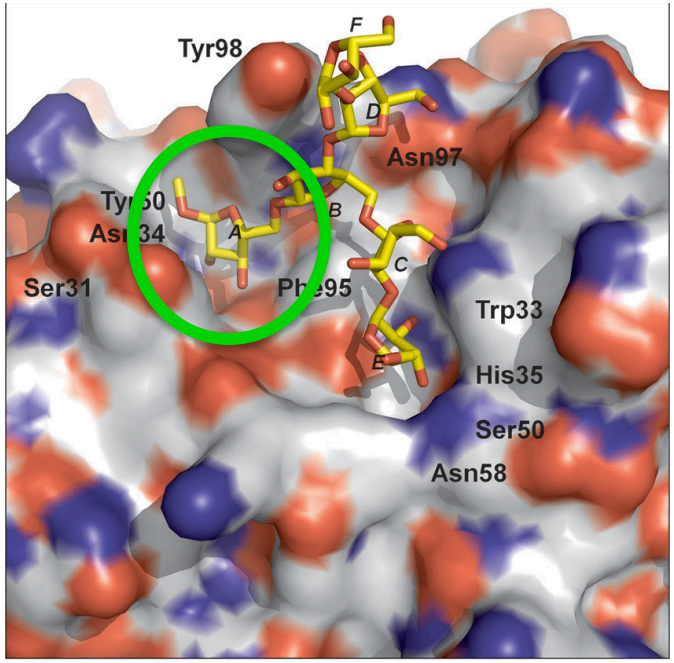
Murase et al. (2009) *Mol. Immunol.*

Lak et al. (2014) *Chem. Eur. J.*, **20**, 1-12.

Per-residue Energy Decomposition Identifies Key Residues



1, R = Lipoarabinomannan
2, R = OCH_3
3, R = $O(CH_2)_8NHAcyl$-CM5 Chip
4, R = $O(CH_2)_8NH_2$



Conclusion from Example 3: The per-residue energy analysis indicated that residue A is the immunodominant component of the carbohydrate.

Murase et al. (2009) *Mol. Immunol.*

Lak et al. (2014) *Chem. Eur. J.*, **20**, 1-12.

Summary

Computational methods provide a physical model for interpreting complex data such as NMR observables

MD simulations can:

- predict the size of the carbohydrate epitope
- predict the effect of chemical modifications
- predict the affinity contributions from individual residues in the protein
- predict the affinity contributions from individual residues in the oligosaccharide
- provide testable models for avidity effects

Therefore MD simulations can help in the design of carbohydrate-based therapeutics and vaccines

Acknowledgements

Oliver Grant

Ji (Mia) Ye

Jodi Hadden

David Thieker

James Paulson

Wenjie Peng

Corwin Nycholat

Robert DeVries

Harry Jennings

Todd Lowary