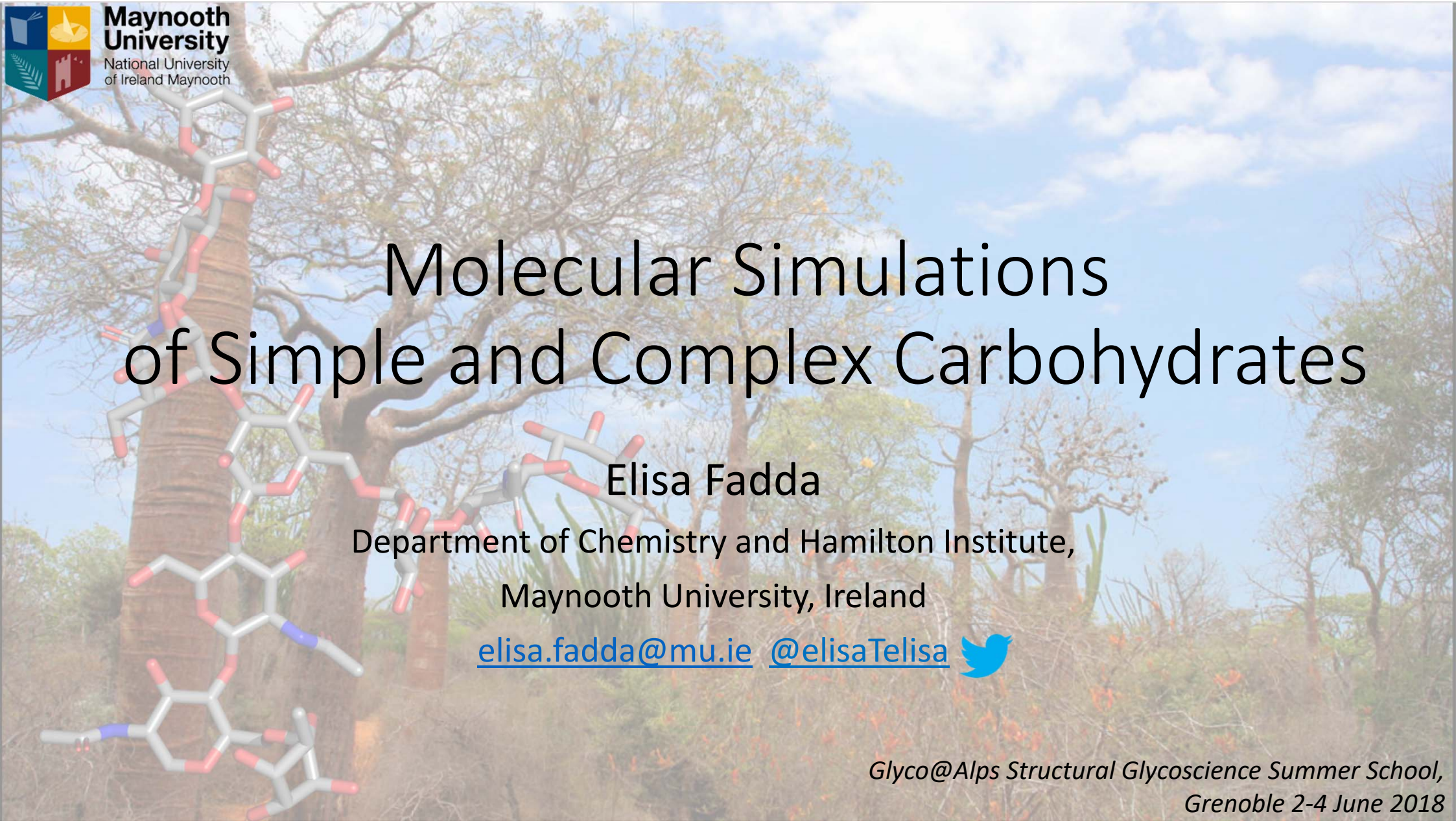


Molecular Simulations of Simple and Complex Carbohydrates

Elisa Fadda

Department of Chemistry and Hamilton Institute,

Maynooth University, Ireland

elisa.fadda@mu.ie @elisaTelisa

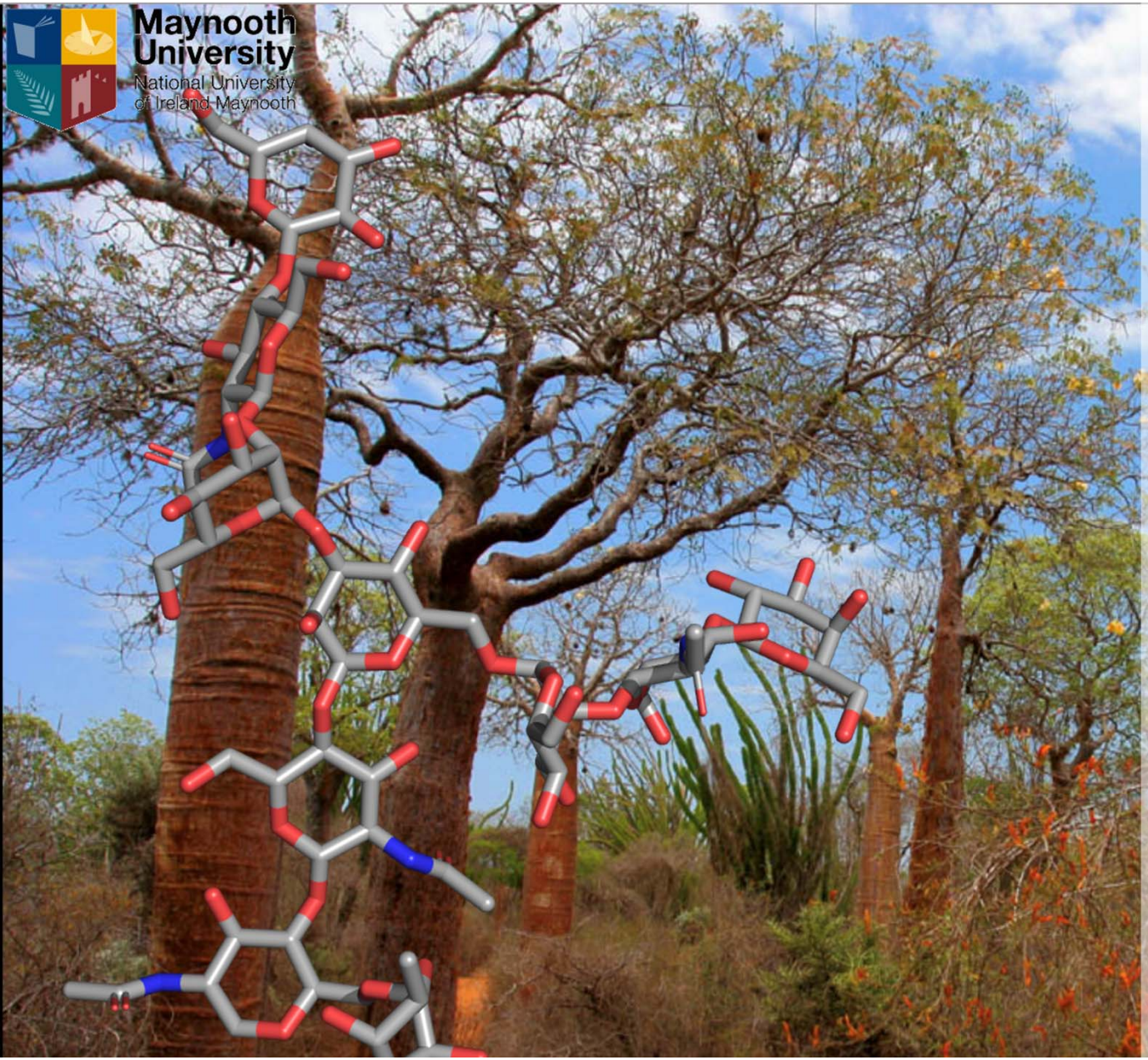
Glyco@Alps Structural Glycoscience Summer School,
Grenoble 2-4 June 2018

Maynooth University
National University of Ireland Maynooth

Chris Hadfield ✓
@Cmdr_Hadfield

Following

Tá Éire fíorálainn! Land of green hills and dark beer. With capital Dublin glowing in the Irish night.

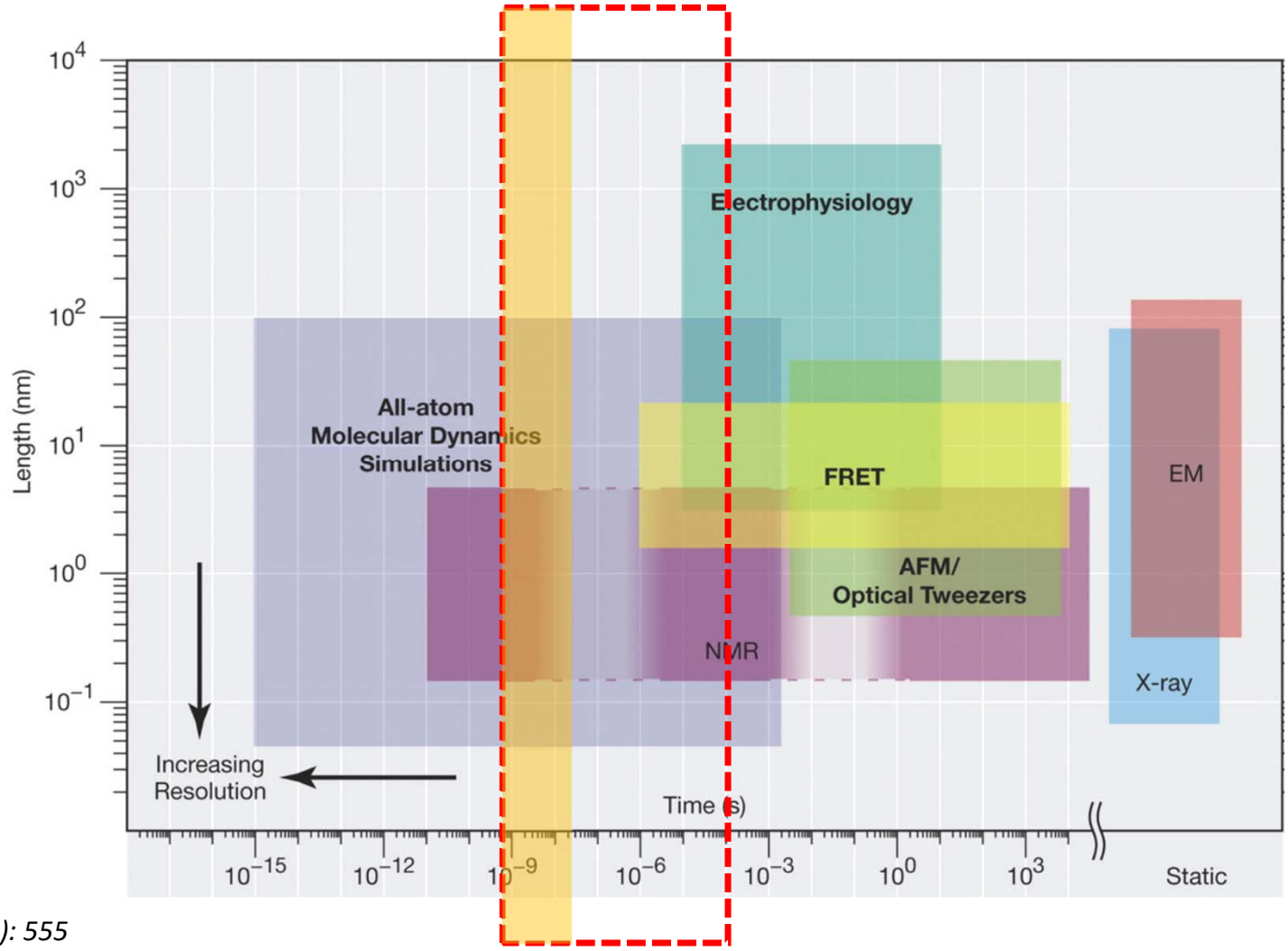


Complex carbohydrates (glycans) decorate the surface of proteins and lipids

- Recognition
- Binding
- Cell migration
- Cell-cell interaction
- Protein stability
- Protein conformation
- Protein function

Maynooth University
National University of Ireland Maynooth

Timeframe of Molecular Motions



Dror et al, JGP (2010) 135 (6): 555

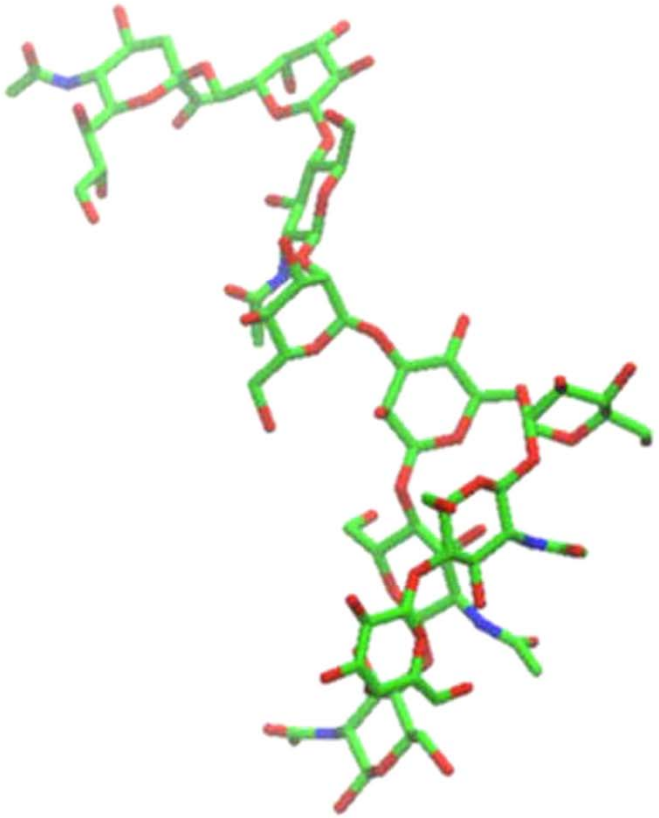
Maynooth University
National University of Ireland Maynooth

Structural/Conformational Disorder

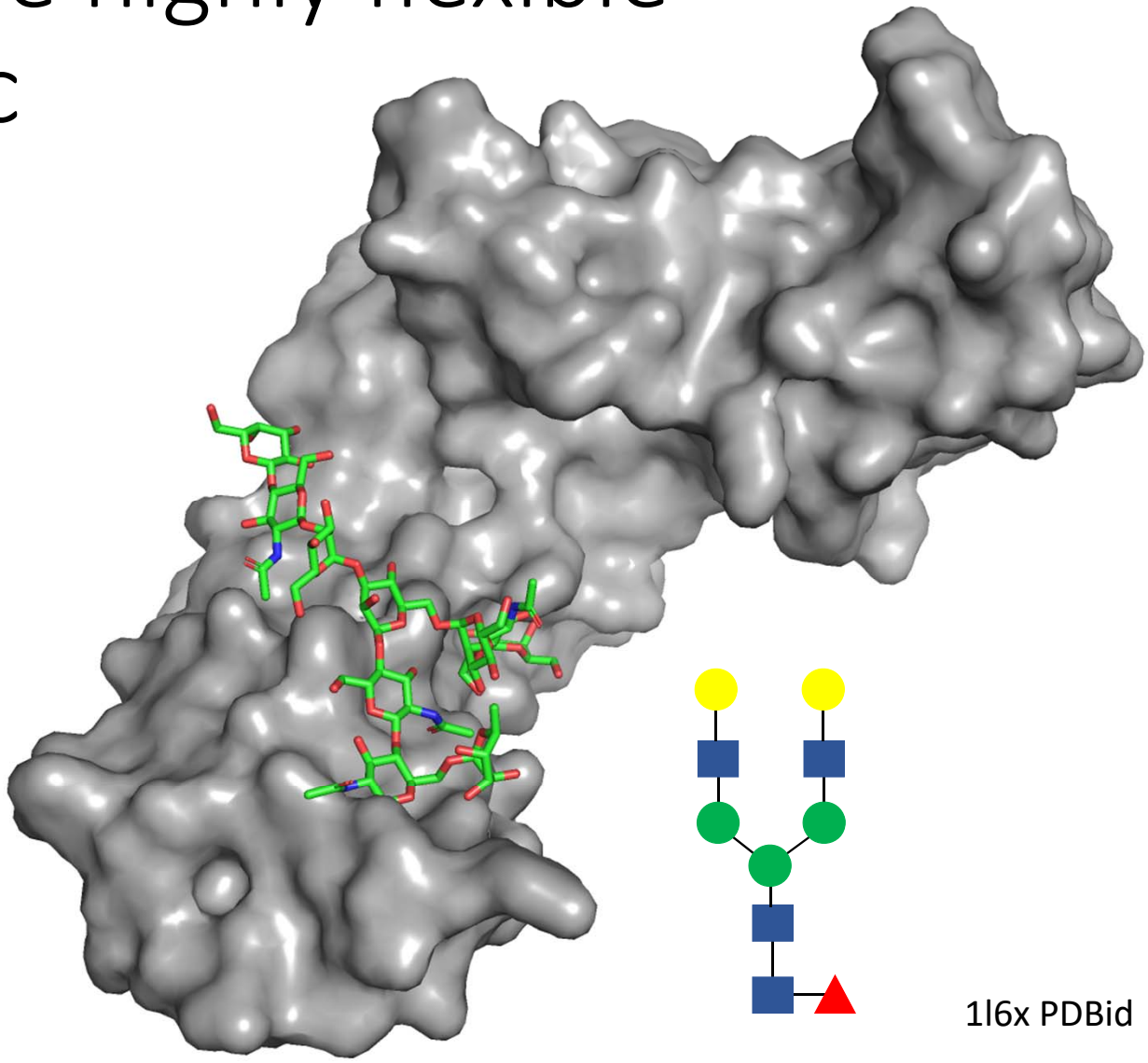
High conformational flexibility => “intrinsic disorder”

Maynooth University
National University of Ireland Maynooth

Sugars can be highly flexible and dynamic



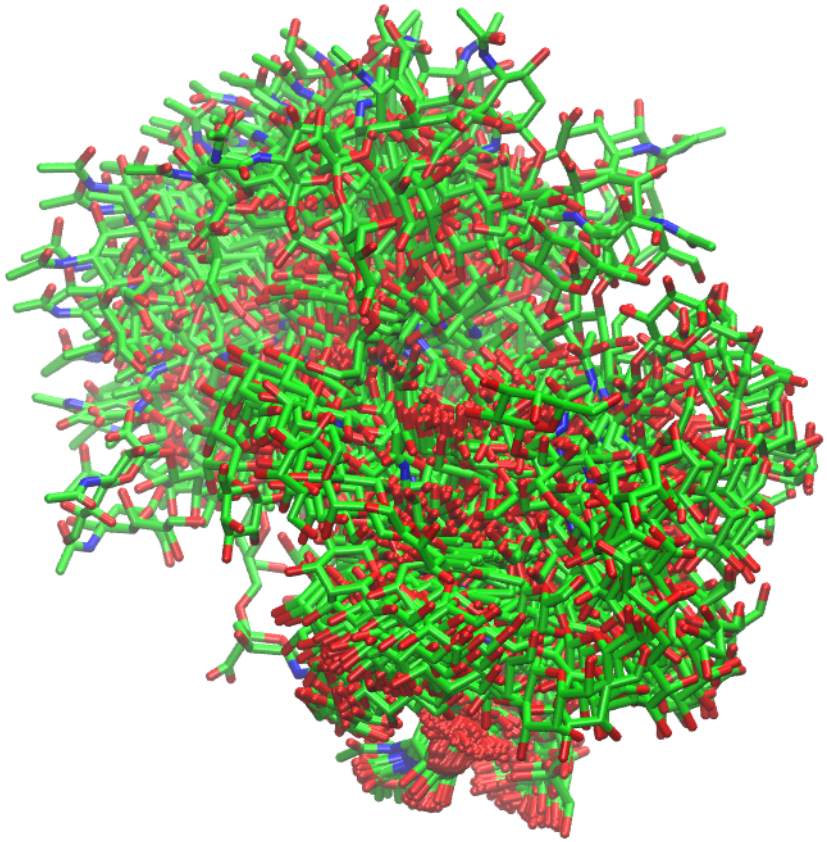
250 ns single trajectory



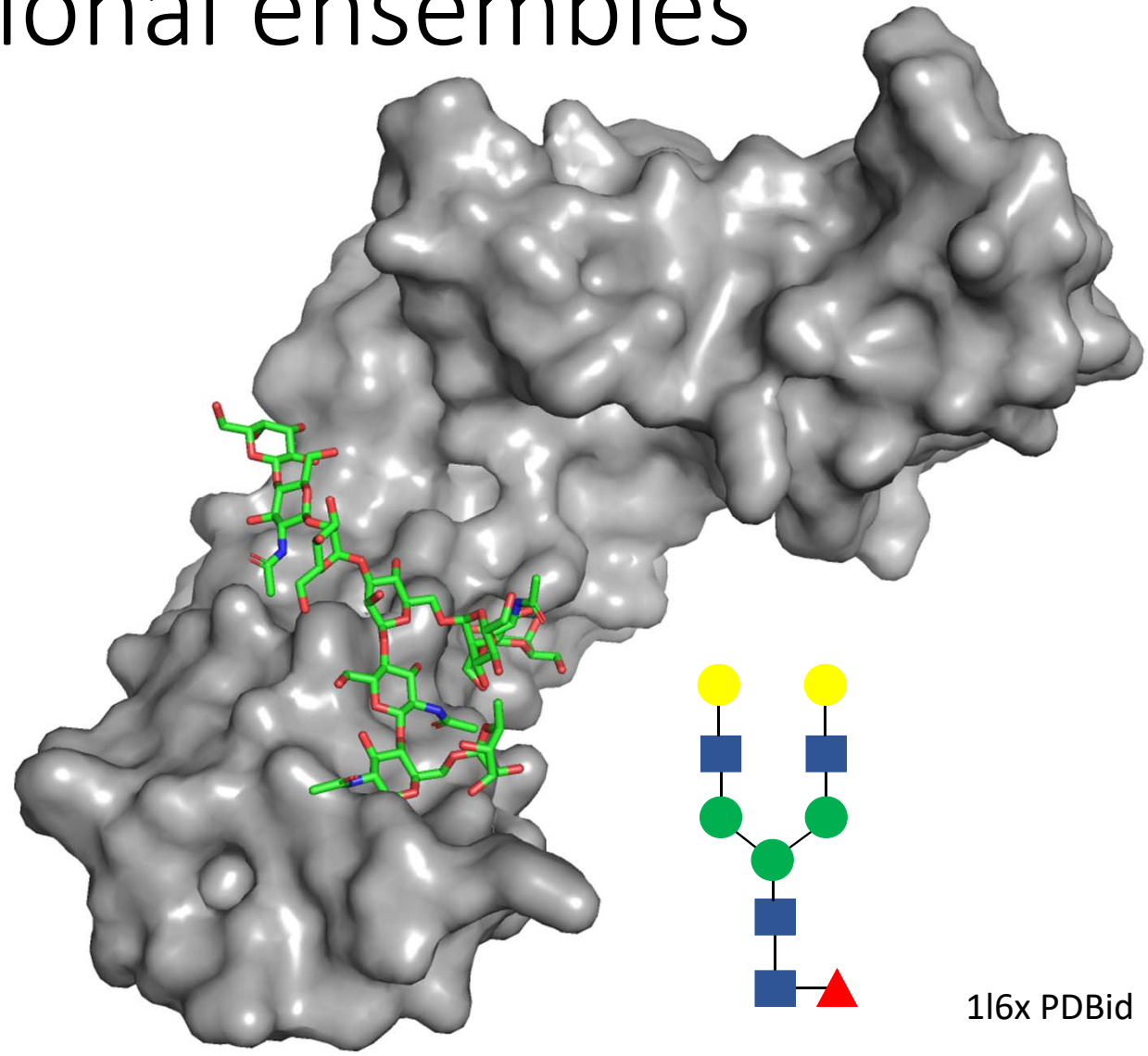
1l6x PDBid

Maynooth University
National University of Ireland Maynooth

Conformational ensembles



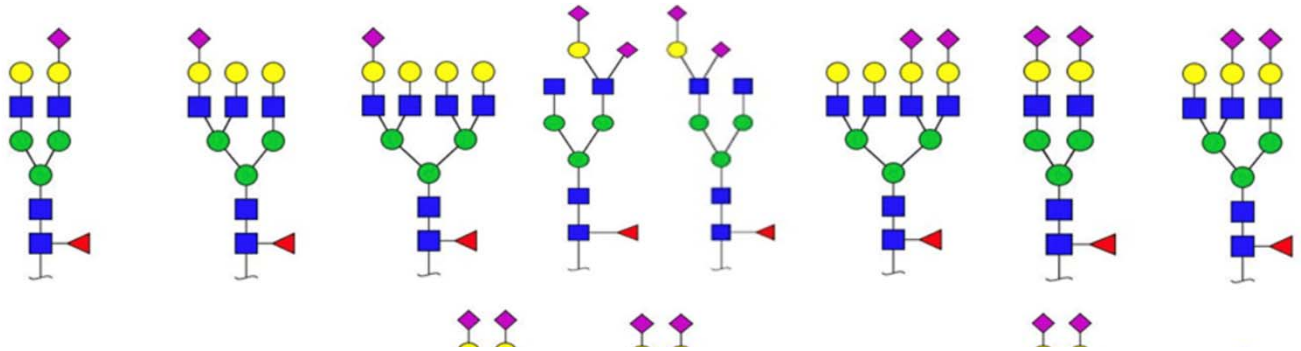
250 ns single trajectory overlay



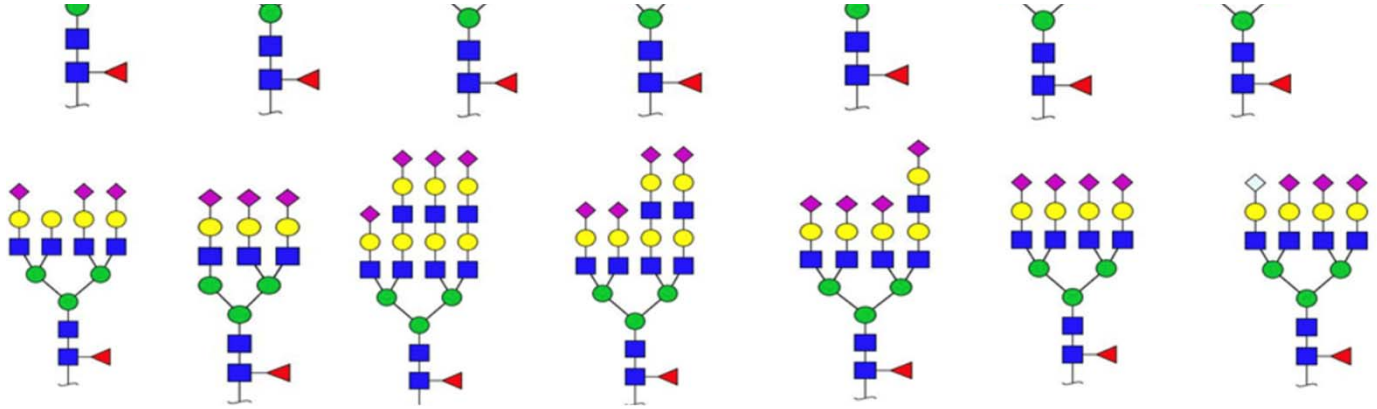
1l6x PDBid

Maynooth University
National University of Ireland Maynooth

Not every N-glycan acts the same!

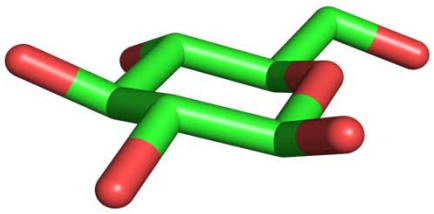


… so where is the difference?

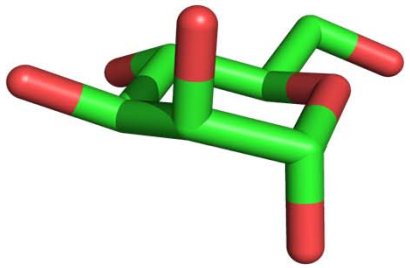


Sequence-to-structure relationships

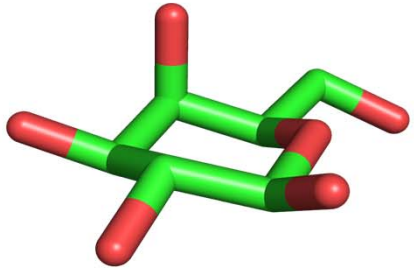
This conformational propensity/degree of flexibility (or lack-there-of) depends on the glycan's **sequence** and linkages



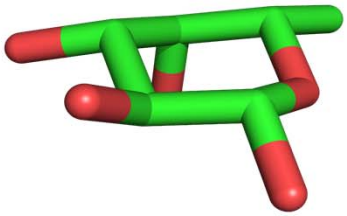
Glu



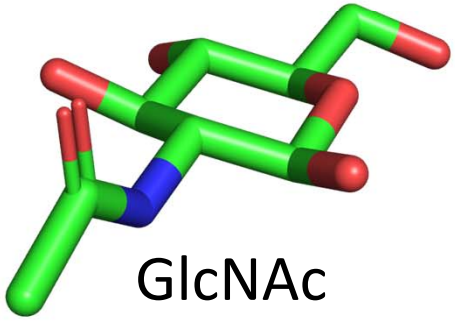
Man



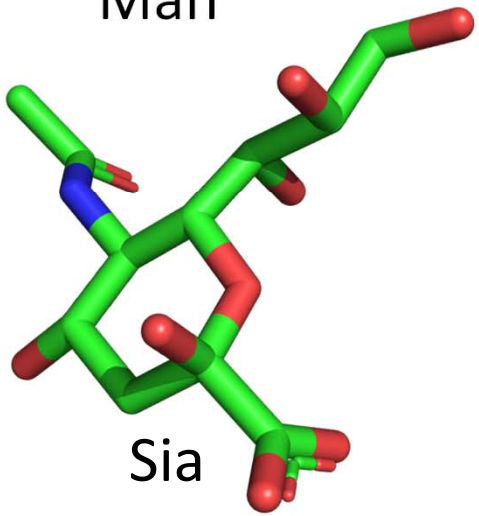
Gal



Fuc



GlcNAc



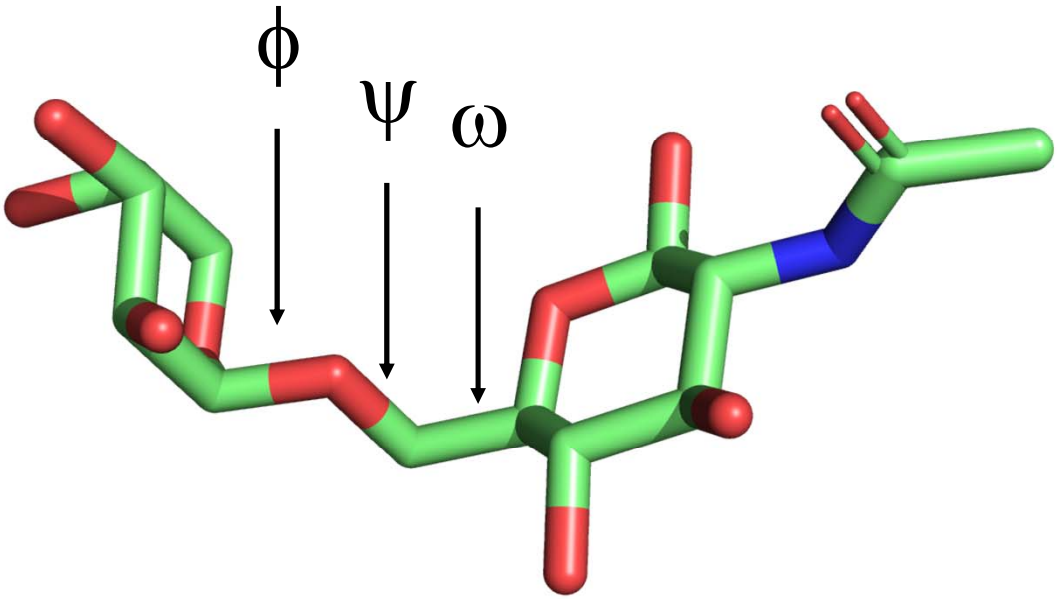
Sia

<u>One trillion</u> possible combinations!!

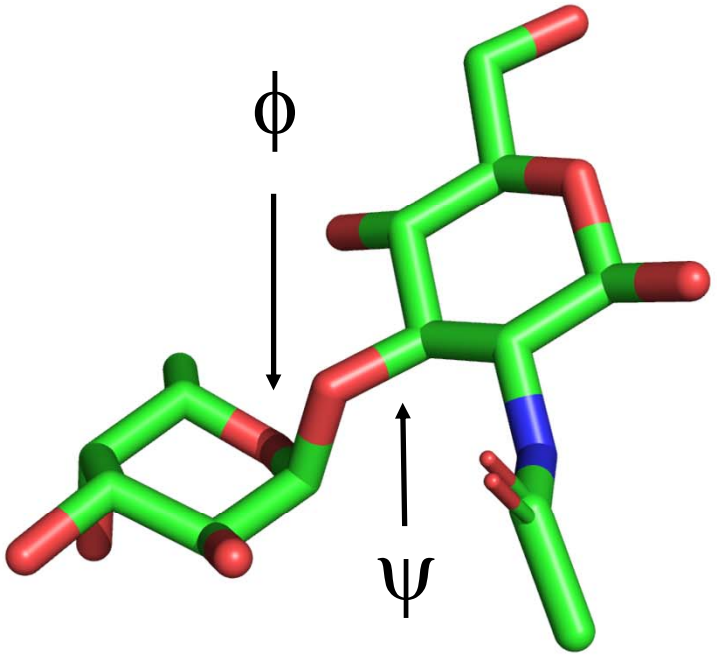
… but only a few of those are actually found

Sequence-to-structure relationship

This conformational propensity/degree of flexibility (or lack-there-of) depends on the glycan's sequence and **linkages**



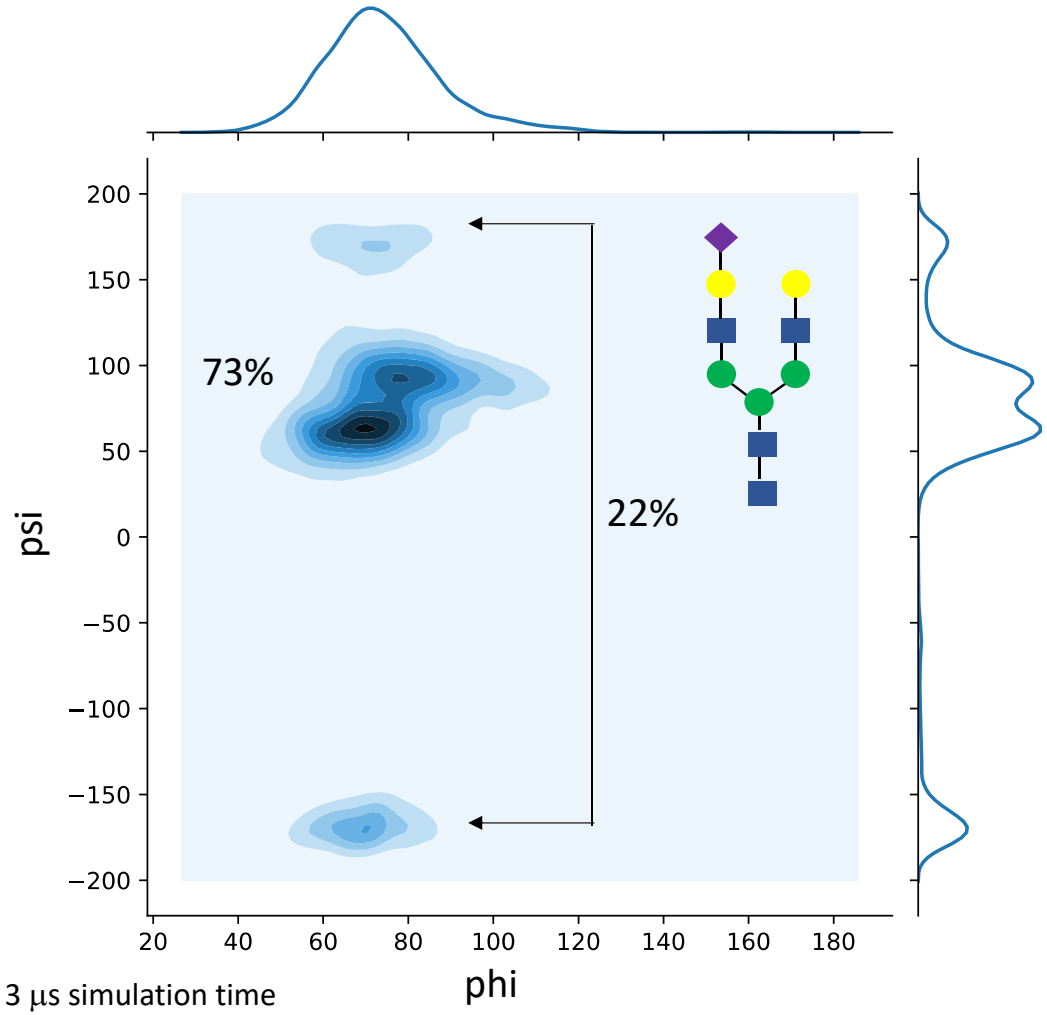
L-Fuc-α(1-6)-GlcNAc



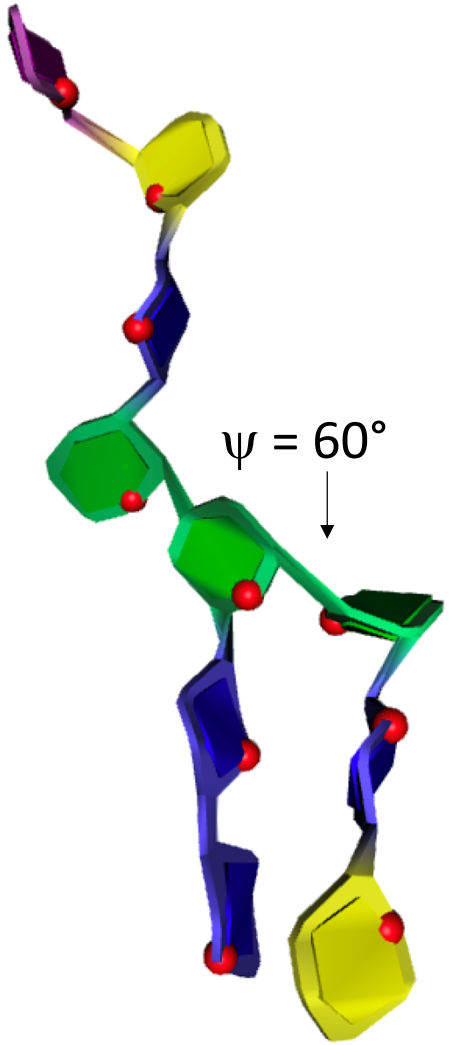
L-Fuc-α(1-3)-GlcNAc

Maynooth University
National University of Ireland Maynooth

Conformational analysis of MD data



3 μs simulation time

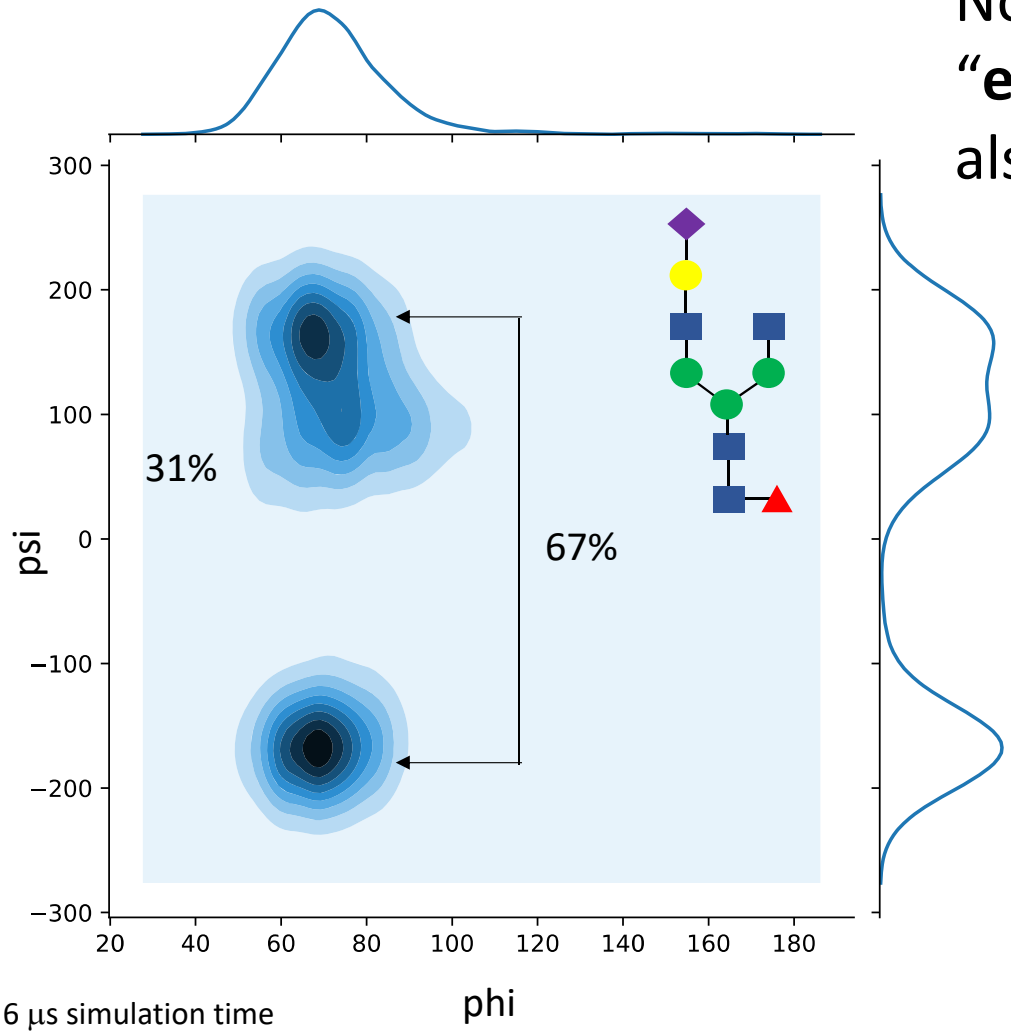


Conformational analysis (3 μs) shows that the Gal (1-6) arm is prevalently found in a "**folded-over**" conformation, while the (1-3) arm is mostly extended

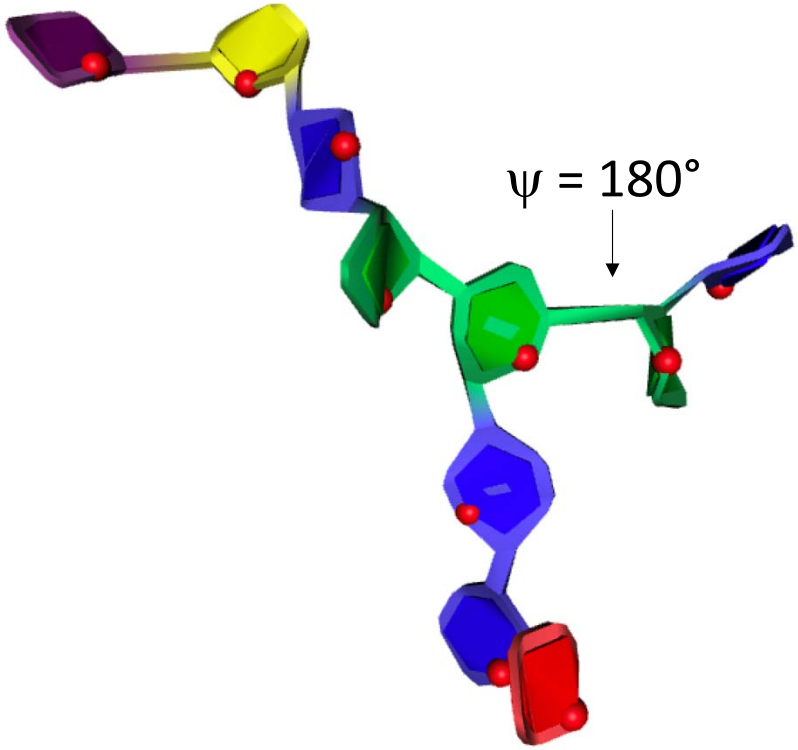
Maynooth University
National University of Ireland Maynooth

Conformational analysis of MD data

Non Gal (1-6) arm is prevalently found in a "**extended**" conformation, with the (1-3) arm also extended (6 μs)

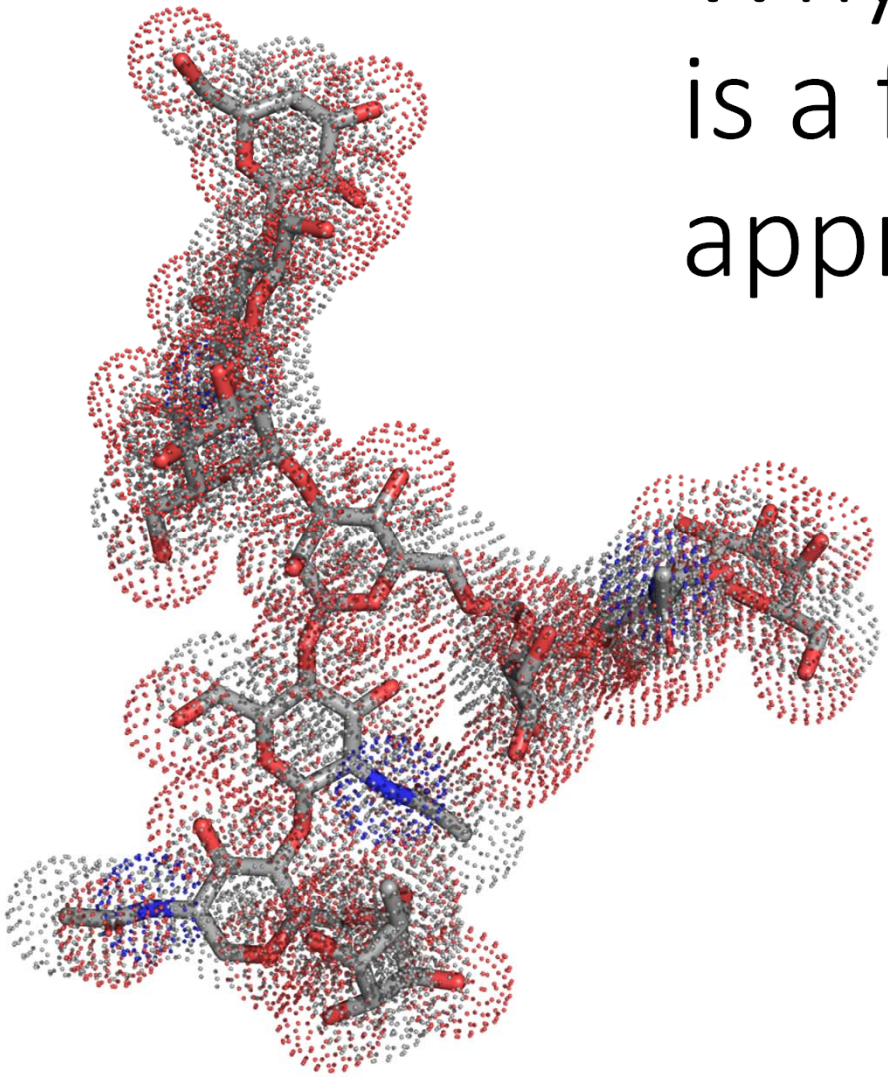


6 μs simulation time



Maynooth University
National University of Ireland Maynooth

Why Classical Mechanics is a fair enough approximation?



According to the Born Oppenheimer (BO) or adiabatic approximation we can separate the motion of the nuclei from the motion of electrons and safely assume that **the electron cloud adjusts instantly** to changes in the nuclear configuration

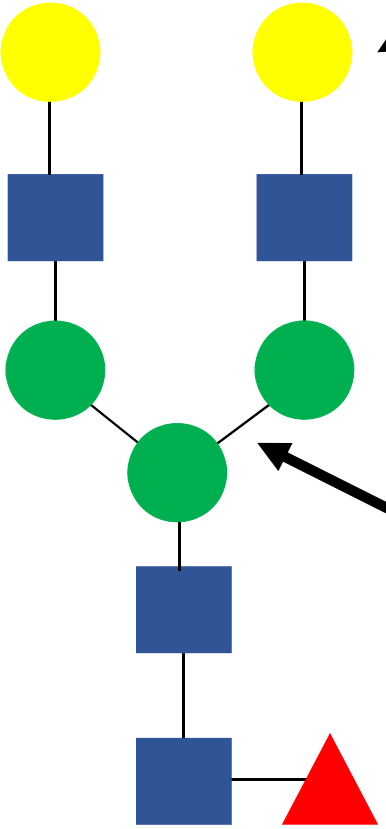
Maynooth University
National University of Ireland Maynooth

Sugars in a Classical Mechanics Representation

Atoms are represented by hard, impenetrable spheres

Electrostatics are reproduced by a Coulomb-type potential

Dispersion interactions (hydrophobic) are reproduced by a van der Waals-type potential



Bonds/Angles/Torsions are reproduced by a Hooke-type potential

Maynooth University
National University of Ireland Maynooth

Molecular Dynamics

Initial positions given by the **PDB**
Initial velocities determined based on a Boltzmann distribution of velocities at the target temperature

$$\vec{F} = m\vec{a} = -\frac{dU}{dr}$$

New positions and velocities through integration

MD run → trajectory

THE FORCE FIELD

$$v^{\text{Coulomb}}(r) = \frac{Q_1 Q_2}{4\pi\epsilon_0 r},$$

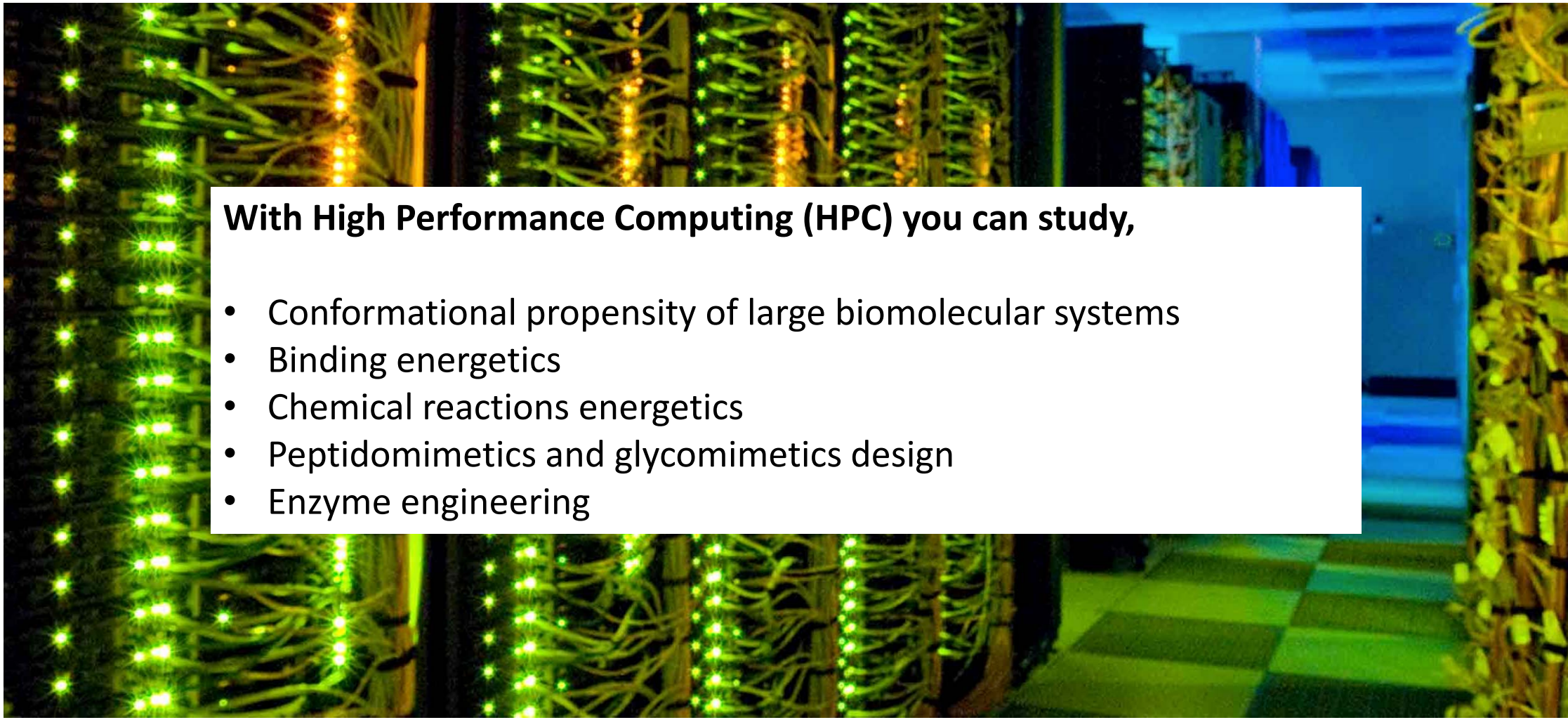
$$v^{\text{LJ}}(r) = 4\varepsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right].$$

$$\begin{aligned} \mathcal{U}_{\text{intramolecular}} = & \frac{1}{2}\sum_{\text{bonds}} k_{ij}^{r}\left(r_{ij} - r_{\text{eq}}\right)^2 \\ & + \frac{1}{2}\sum_{\substack{\text{bend}\\ \text{angles}}} k_{ijk}^{\theta}\left(\theta_{ijk} - \theta_{\text{eq}}\right)^2 \\ & + \frac{1}{2}\sum_{\substack{\text{torsion}\\ \text{angles}}}\sum_{m} k_{ijkl}^{\phi,\,m}\left(1 + \cos(m\phi_{ijkl} - \gamma_m)\right) \end{aligned}$$

Maynooth University
National University of Ireland Maynooth

Computational Biophysics

ICHEC
Irish Centre for High-End Computing



With High Performance Computing (HPC) you can study,

- Conformational propensity of large biomolecular systems
- Binding energetics
- Chemical reactions energetics
- Peptidomimetics and glycomimetics design
- Enzyme engineering

Maynooth University
National University of Ireland Maynooth

GLYCAM06 Carbohydrate Force Field

A complete, self-contained and transferable set of parameters for the simulation of carbohydrates and glycol-conjugates

- Carbohydrates of **all ring sizes and conformations** for both monosaccharides and oligosaccharides (not open ring conformations)
- Parameters for **N-** and **O-glycosidic linkages**, suitable for combination with the *param99sb* version of the AMBER protein force field
- Parameters for sulfated (and not) **GAGs**

- **GLYCAM06 is a 'stand alone' force field** that can in principle be combined to any other protein ff (non-covalent linkages)
- It is the only ff that uses the **same atom type (CG) for α and β anomers** allowing for ring conformational changes (puckering)

Kirschner *et al J Comp Chem* (2008) 29(4): 622–655
Fadda and Woods *Drug Discov Today* (2010) 15(15-16):596-609

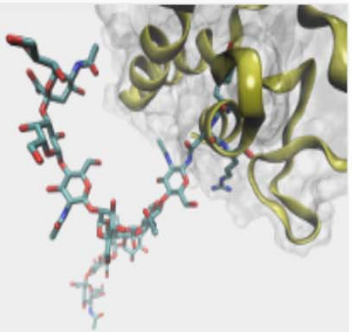
Maynooth University
National University of Ireland Maynooth

GLYCAM-Web (glycam.org)

GLYCAM

HOME ABOUT US NEWS LEGACY TOOLS HELP DOWNLOADS

3D Structure Prediction Tools | 3D Structure Libraries | Other Tools | Force Field | Documentation | Report a Problem



GLYCAM-Web is dedicated to simplifying the prediction of three-dimensional structures of carbohydrates and macromolecular structures involving carbohydrates. Click on the tabs above to learn the current capabilities of the site.

***You will build a simple and a complex carbohydrate this afternoon**

Currently available tools:

Carbohydrate Builder | Glycoprotein Builder | Oligosaccharide Libraries | PDB Preprocessor | Build via Text | Build via URL | GAG Builder

Auto Docking | Gly-Spec (Grafting) | Carbohydrate Visualization

Maynooth University
National University of Ireland Maynooth

CHARMM + CHARMM-GUI

- CHARMM36 Carb parameter set (PAR_ALL36_CARB.PRM)

42. Guvench O, Mallajosyula SS, Raman EP, Hatcher E, Vanommeslaeghe K, Foster TJ, Jamison FW, MacKerell AD., Jr CHARMM Additive All-Atom Force Field for Carbohydrate Derivatives and Its Utility in Polysaccharide and Carbohydrate-Protein Modeling. J Chem Theory Comput. 2011;7 (10):3162–3180. [PMC free article] [PubMed]

43. Mallajosyula SS, MacKerell AD., Jr Influence of Solvent and Intramolecular Hydrogen Bonding on the Conformational Properties of O-Linked Glycopeptides. J Phys Chem B. 2011;115 (38):11215–11229. [PMC free article] [PubMed]

44. Mallajosyula SS, Guvench O, Hatcher E, MacKerell AD., Jr CHARMM Additive All-Atom Force Field for Phosphate and Sulfate Linked to Carbohydrates. J Chem Theory Comput. 2012 doi: 10.1021/ct200792v. [PMC free article] [PubMed] [Cross Ref]

J Chem Theory Comput. 2018 Jun 12;14(6):3132-3143. doi: 10.1021/acs.jctc.8b00175. Epub 2018 May 4.

CHARMM Drude Polarizable Force Field for Glycosidic Linkages Involving Pyranoses and Furanoses.

Aytenfisu AH[1], Yang M[1,2], MacKerell AD Jr[1].

TABLE 2

A summary of the parameterization protocol used for the development of four carbohydrate force fields reviewed

	CHARMM	GLYCAM06	GROMOS-45A4	OPLS-AA-SEI
Valence terms				
Equilibrium bond lengths (r) and angles (θ)	Chosen to reproduce crystal internal and unit-cell geometries	Chosen to reproduce neutron-diffraction geometries	GROMOS-45A3	OPLS-AA
Force constants $kb/k\theta$	Fit to QM data	Fit to QM data	GROMOS-45A3	OPLS-AA
Torsion terms	Fit to QM rotational energy curves	Fit to QM rotational energy curves	Fit to QM rotational energy curves	Fit to QM rotational energy curves
Partial charges	Empirically fit for carbohydrate fragments, and refined to reproduce: QM solute–water E_{int}, and experimental V_m of carbohydrate solutions	QM RESP fit and ensemble averaged over multiple conformations. RESP scaling to reproduce crystal unit-cell geometries	QM RESP fit with averaging over atom types	OPLS-AA (empirically fit to reproduce heat of vaporization and densities of pure liquids)
vdW terms	CHARMM22	AMBER PARM94	GROMOS-45A3	OPLS-AA
1,4 scaling (Elec/vdW)	No/no	No/no	No/no	Yes/yes
Unique charge sets for α- and β-anomers	No	Yes	No	No
Unique charges on each atom	No	Yes	No	No
Unique atom types for α- and β-anomers	Yes	No	Yes	Yes

Fadda and Woods *Drug Discov Today* (2010) 15(15-16):596-609

Maynooth University
National University of Ireland Maynooth

What do I need to run an MD simulation?

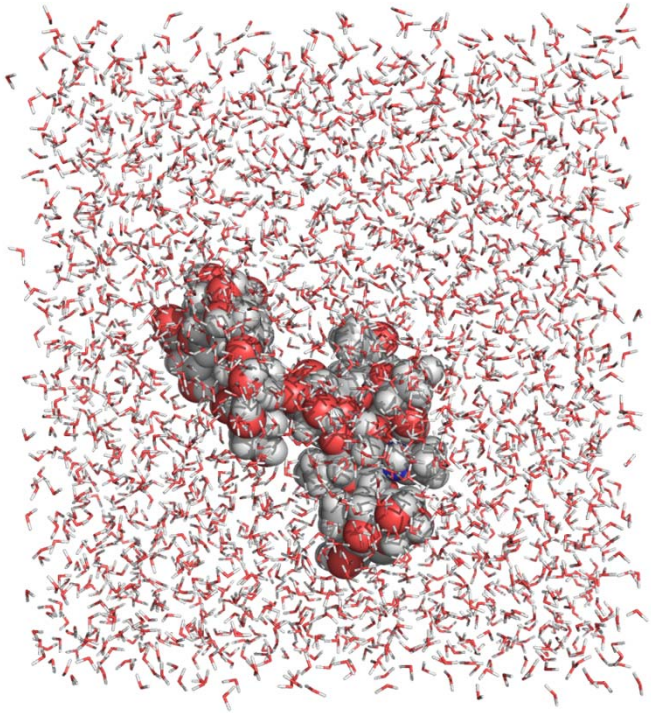
1. Coordinate file (pdb)
2. Decide what force field your want to use

Carbohydrate	Protein/Ions	Water	MD software
GLYCAM06	AMBER99-SB-ILDN	Tip3P	AMBER v.12/16/18
CHARMM36	CHARMM36	Tip4P(Ew)	CHARMM
		Tip5P	GROMACS

3. Convert the coord file to be read by the MD running software
4. In AMBER this is done by a tool called ***tleap*** which produces ***.rst*** and ***.prm7*** files, i.e. the amber coordinate file and parameters (topology) file

Maynooth University
National University of Ireland Maynooth

An MD run protocol



1. Simulation box and adjust ionic concentration
2. Energy minimization

3. Equilibration phase

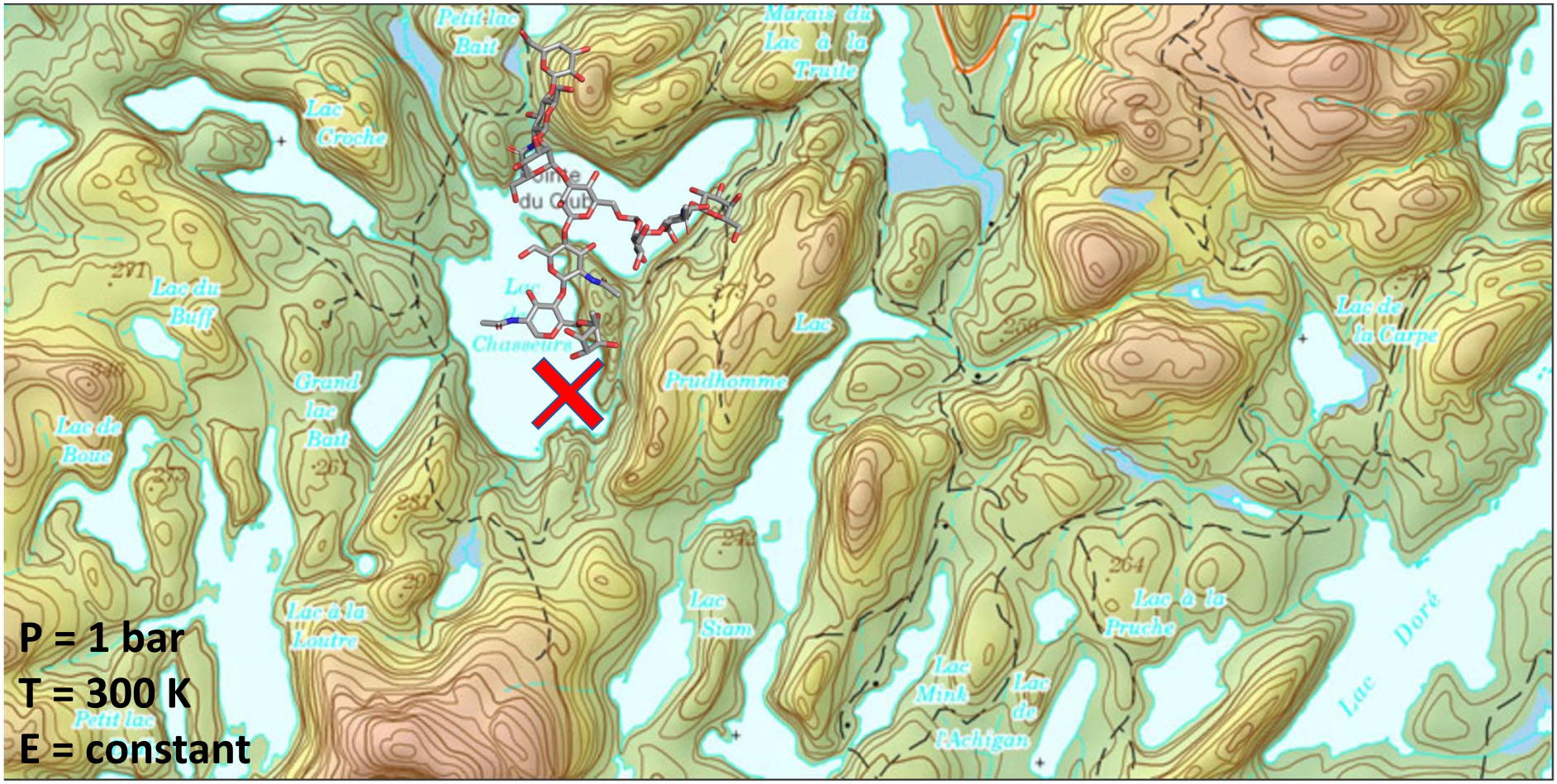
1. Heating 0 -> 300 K (NVT)
2. Equilibration of the pressure (NPT)
3. Equilibration of the conf. degrees of freedom (NPT) *

4. Production (data collection) *

How long should I run steps 3.3 and 4?

Maynooth University
National University of Ireland Maynooth

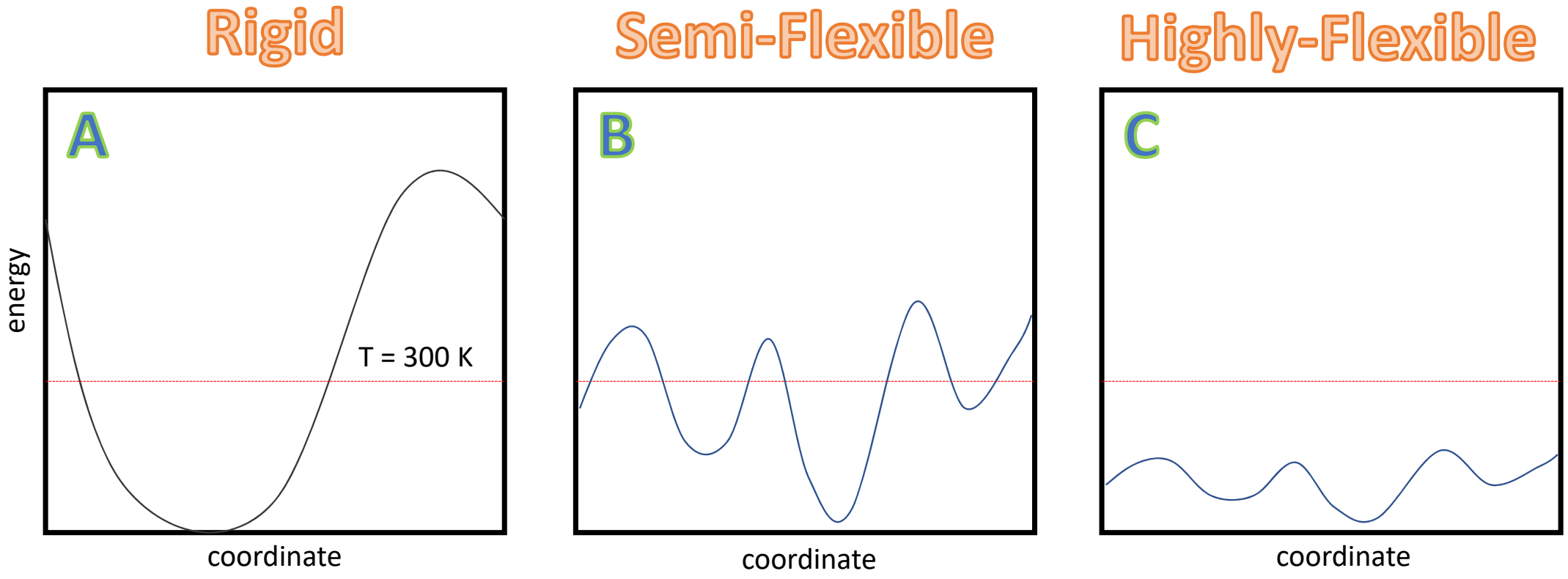
Reaching thermodynamic equilibrium



Maynooth University
National University of Ireland Maynooth

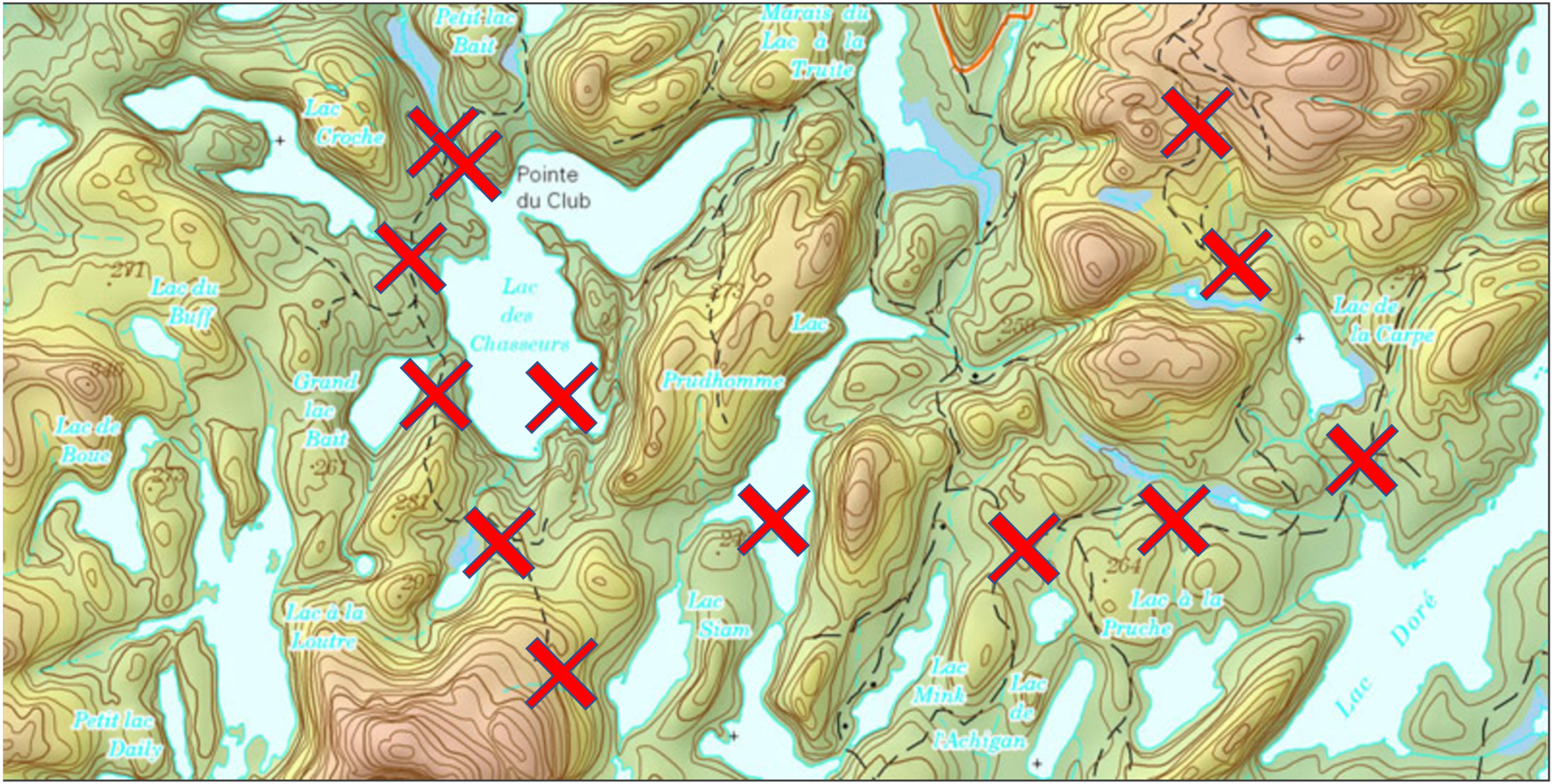
Sampling: How long is long enough?

It depends on the system, on its intrinsic flexibility, and on the chosen ff



Maynooth University
National University of Ireland Maynooth

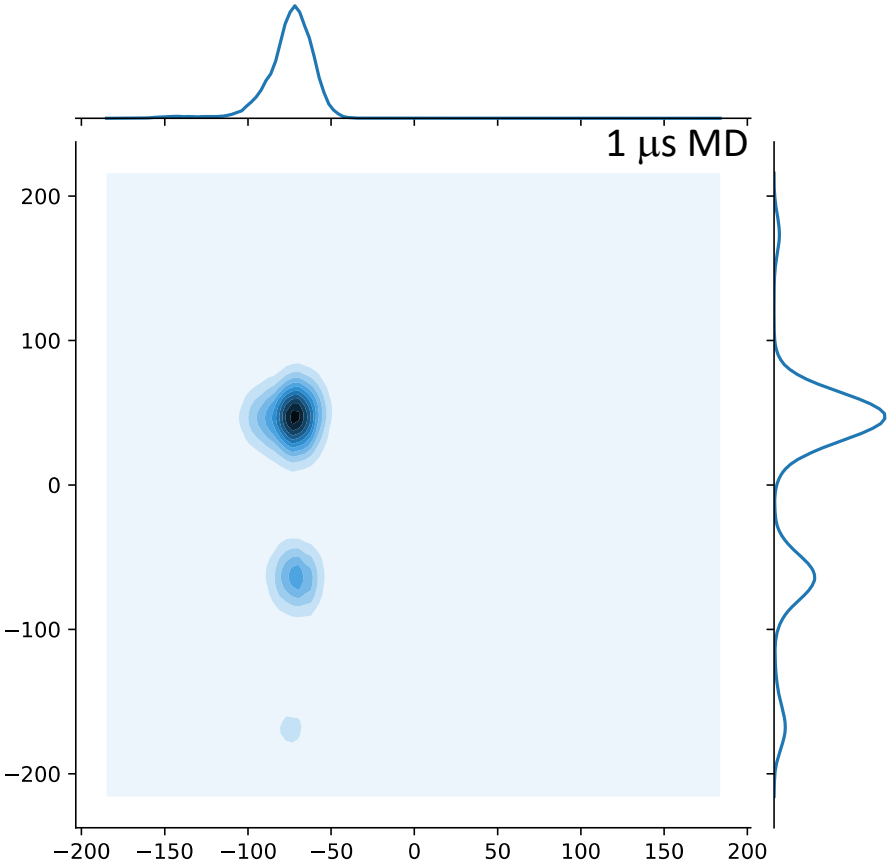
Enough sampling: Exploring the PES exhaustively



Maynooth University
National University of Ireland Maynooth

Tough cases: Sampling rare events

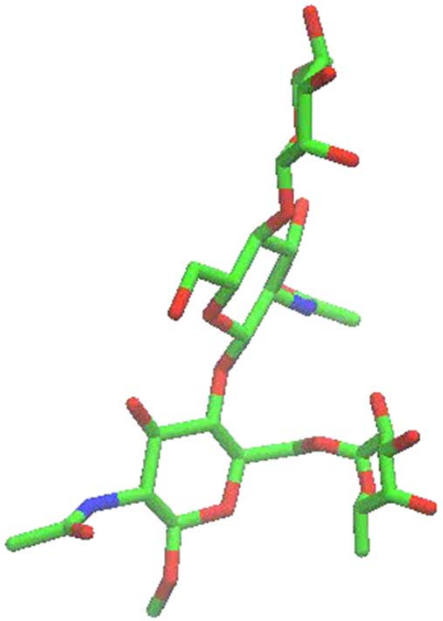
Man-β(1-4)-GlcNAc-β(1-4)-[α(1-6)-Fuc]-GlcNAc-β-OH



gg (67%)

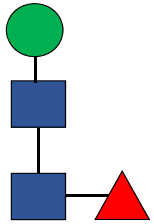
tg (24%)

gt (9%)



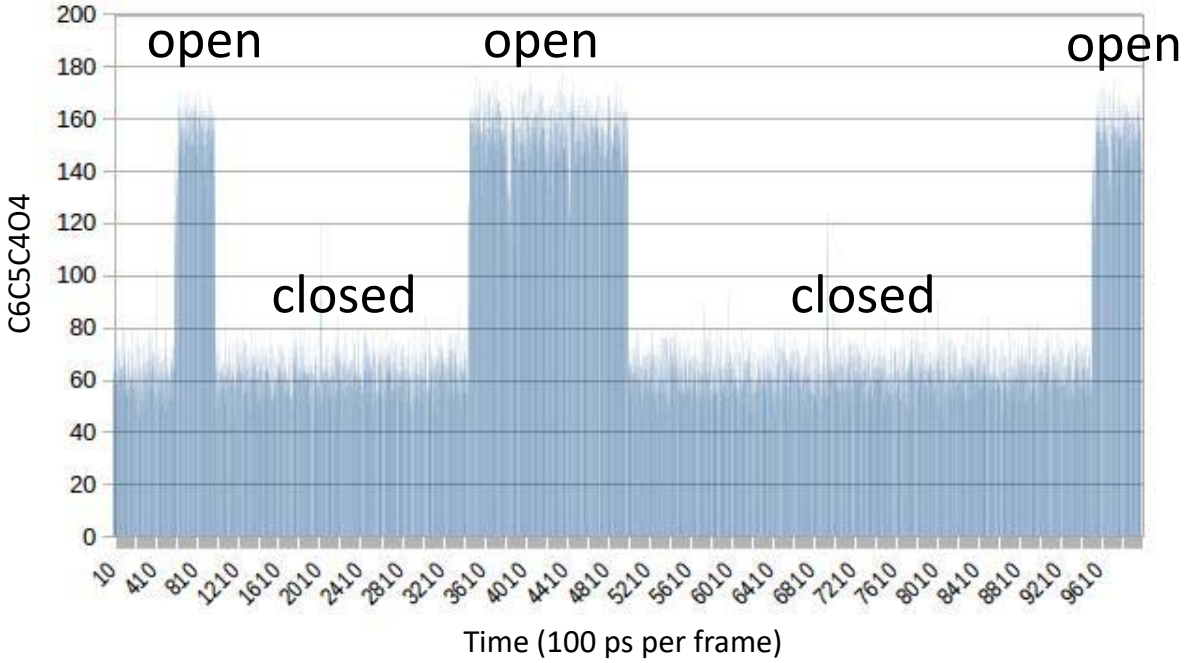
Confomational sampling from MD

Man-β(1-4)-GlcNAc-β(1-4)-[α(1-6)-Fuc]-GlcNAc-β-OH



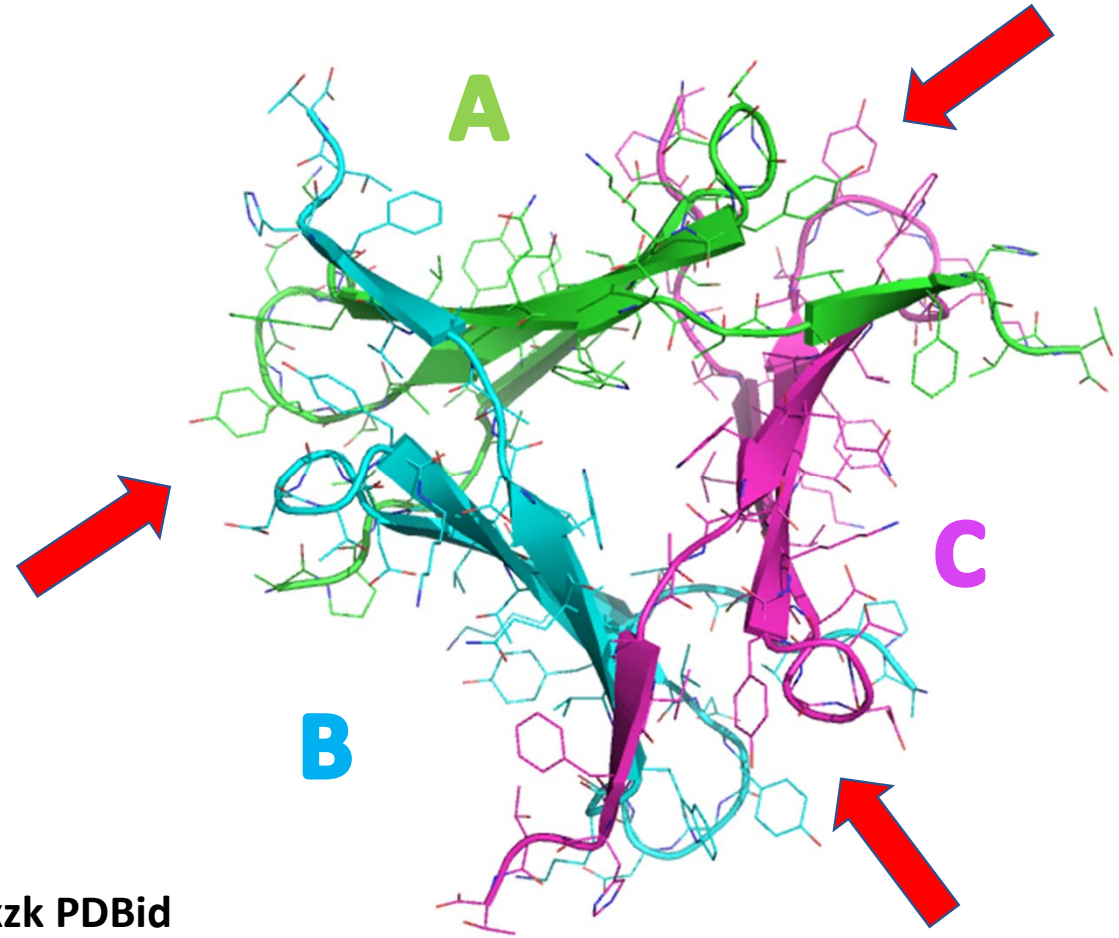
The MD trajectory was extended to 1 μs to analyze the conformational behavior of the sugar. A visual analysis has shown that there is **a significant conformational change occurring** at first around 60 ns.

The conformational change lasts for about 40 ns. If the MD trajectory was of 100 ns, **the relative stability of the 2 conformations would have been 60:40**, largely overestimated.



Over 1 μs the ratio is 76:24 => $\Delta G° \approx$ - 3 kJ/mol

Conformational sampling and molecular recognition



5xzk PDBid

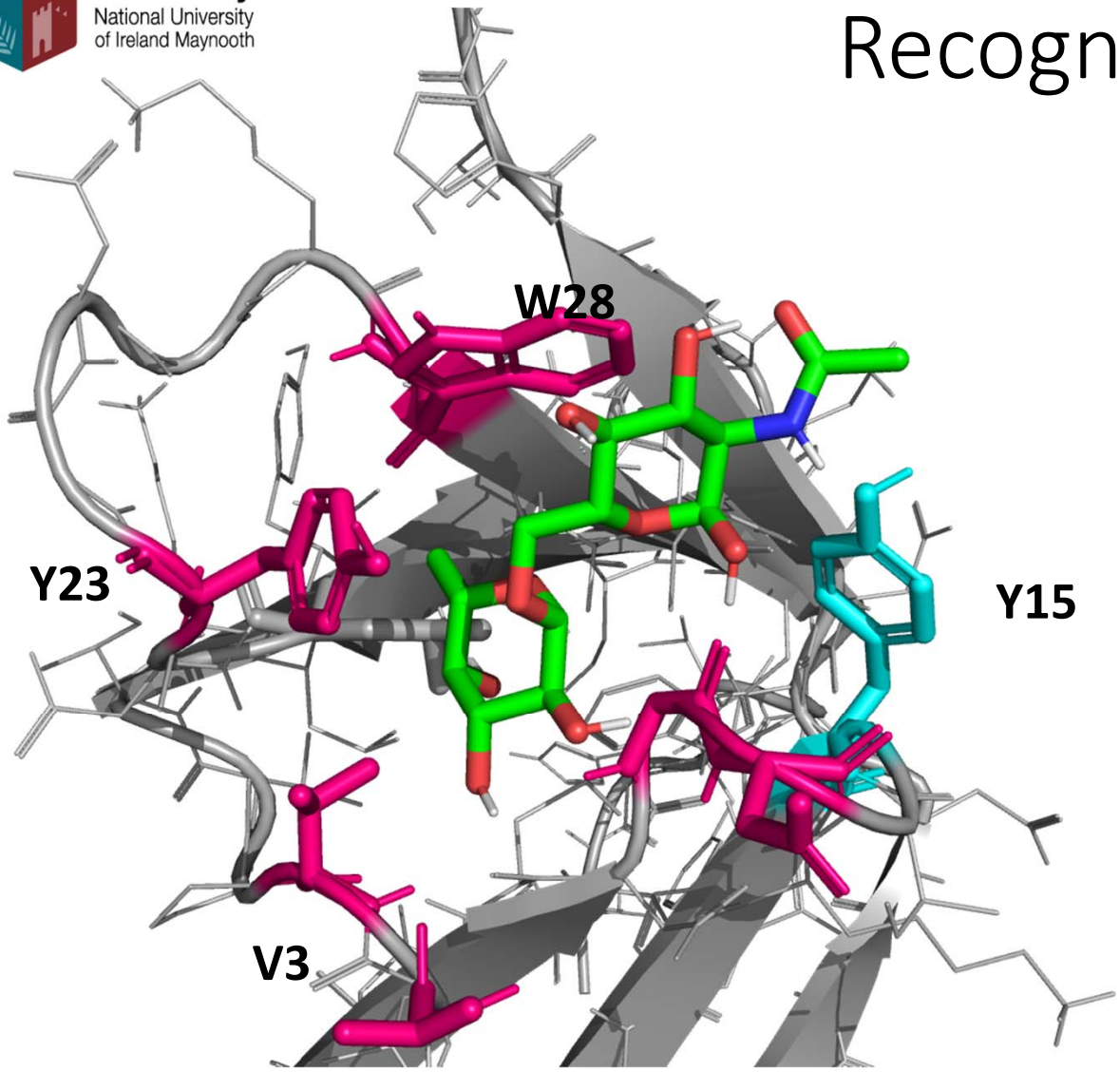
Yamasaki *et al Sci Rep* (2018) 8:7740

Pholiota squarrosa lectin (PhoSL)

- Highly selective for α(1-6)-fucose
- Very low binding affinity for L-fucose (**5.8 – 6.2 mM**)
- Higher affinity for N-glycans (**3 μM**), prob interactions with GlcNAc moieties

Maynooth University
National University of Ireland Maynooth

Recognition of *gt* conformer



Pholiota squarrosa lectin (PhoSL) HADDOCK docking to *gg, tg and gt* conformers

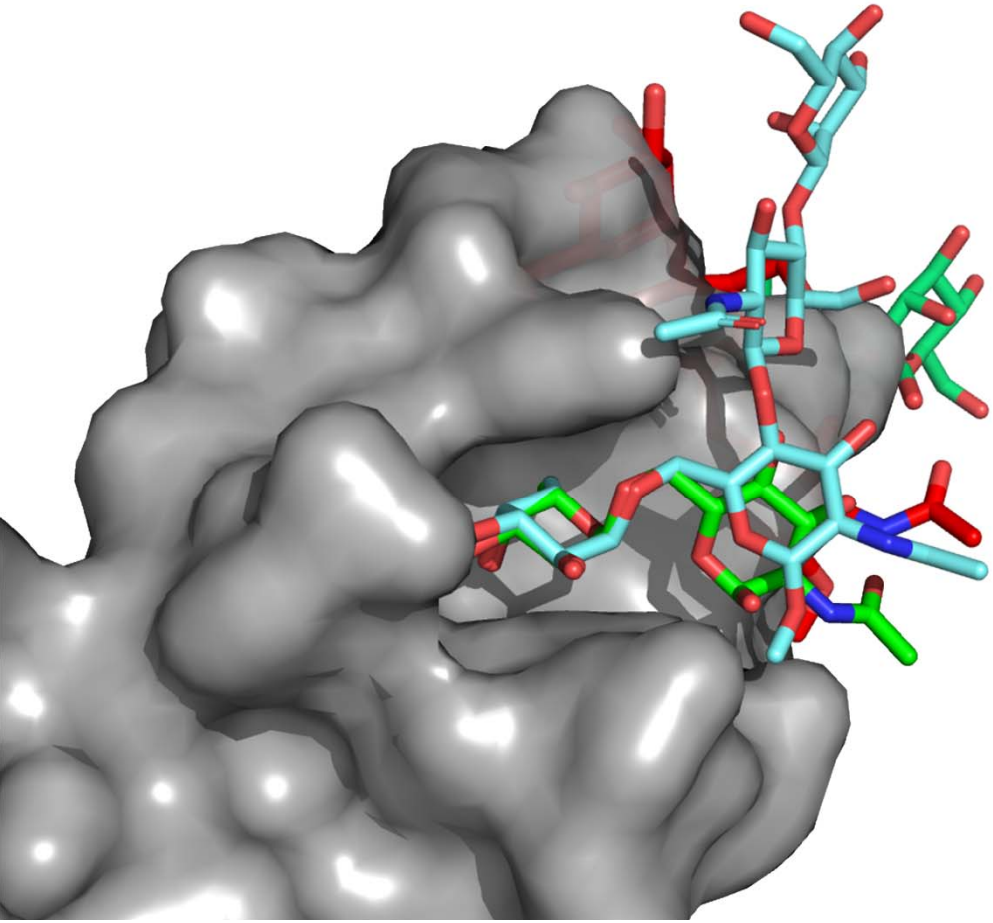
gt disaccharide (4AGT*) binds in the conformation suggested based on NOE (shown)

The bound conformation is consistent with additional contacts to the GlcNAc

gg tetrasaccharide does not seem to have the correct conformation to bind

*Houser *et al* (2015) *Acta Crystallogr.,Sect.D* 71: 442

Recognition of *gt* conformer in the *sugar d* tetrasaccharide



Pholiota squarrosa lectin (PhoSL) structural alignment of *gt conformers*

gt tetrasaccharide presents minor steric clashes.

The ”recognition complexes” will need to be studied via MD simulation to release the clashes

Docking alone is not always the answer!

Maynooth University
National University of Ireland Maynooth

A few take home messages

- MD is a powerful structural biology tool that allows us to understand molecular behavior at the atomistic level of detail
- The above is true provided that the system is properly equilibrated and that the simulation is converged
- Through identification of the glycan conformational propensity we can understand molecular recognition
- Docking is a very helpful technique, but is often not plug-and-play

Acknowledgements

Anne Imberty, Serge Pérez and the Organizers

Students
Aoife Harbison (PhD)
Matthew Nixon (PhD)
Philip Hanley (4th year BSc)
Keith Fenlon (4th year BSc)
Lorna Brosnan (4th year BSc)

N-Glycans/Carb Collaborators
Roisin O'Flaherty (NIBRT, UCD)
Pauline Rudd (NIBRT, UCD)
Chris Taron (NEB)
Sandrine Py (UGA)
Paul V. Murphy (NUIG)
Chris Rowley (Memorial Univ, NL)

$$$
John and Pat Hume Scholarships, MU

Molecule:

LFucpa

Total S

Solvation Options

Solvate Structures:

(•) Yes ( ) No

Choose the shape of the solvent box:

( ) Rectangular (•) Cubic

Enter the size of the solvent buffer in Angstroms:

8.0

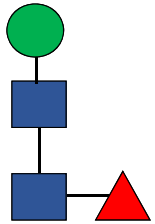
Enter the minimum distance between the center of a solute atom and a solvent atom:

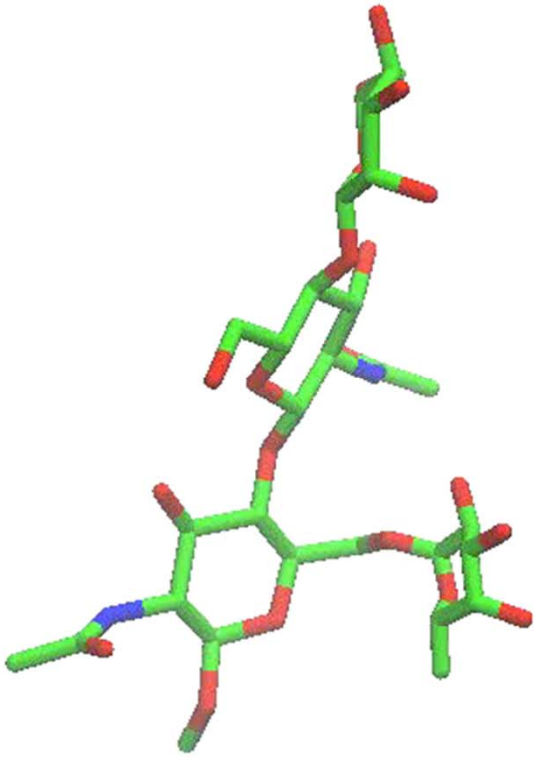
1.5

Save and continue

Ring conf and (1-6) linkage torsion angles

Man-β(1-4)-GlcNAc-β(1-4)-[α(1-6)-Fuc]-GlcNAc-β-OH (**sugar d**)

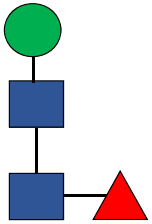




Sampling through high and low populated conformational states during a 1 μs MD trajectory

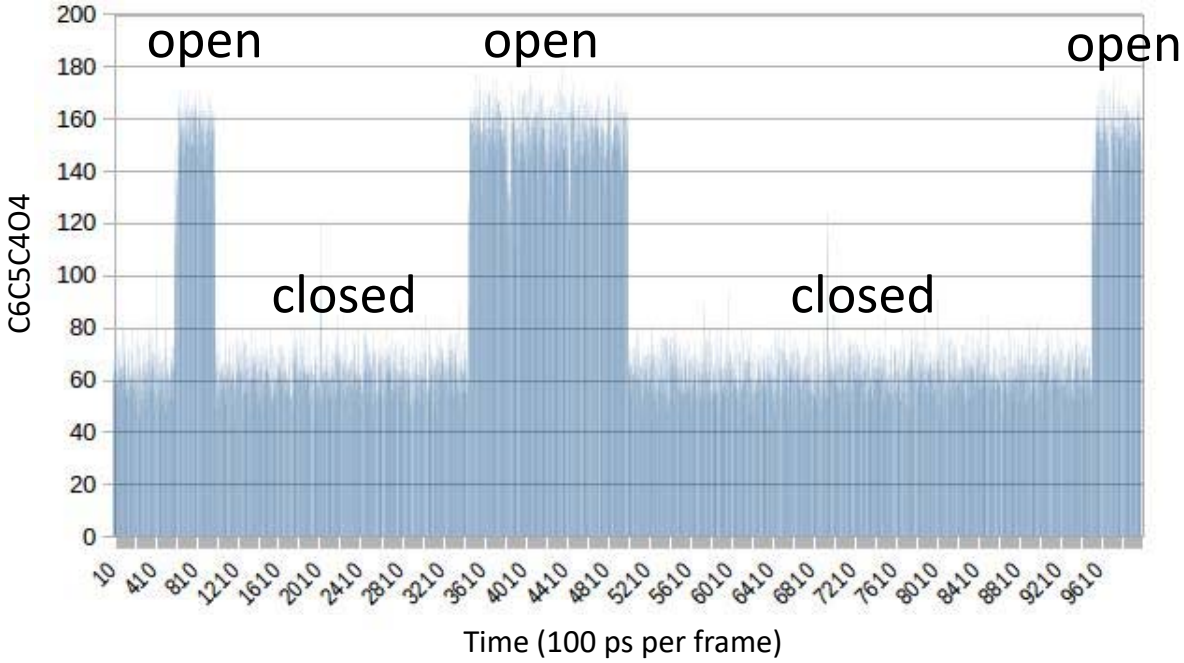
Confomations from MD

Man-β(1-4)-GlcNAc-β(1-4)-[α(1-6)-Fuc]-GlcNAc-β-OH (**sugar d**)



The MD trajectory was extended to 1 μs to analyze the conformational behavior of the sugar. A visual analysis has shown that there is **a significant conformational change occurring** at first around 60 ns.

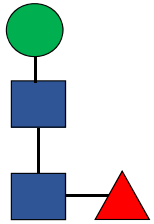
The conformational change lasts for about 40 ns. If the MD trajectory was of 100 ns, **the relative stability of the 2 conformations would have been 60:40**, largely overestimated.



Over 1 μs the ratio is 76:24 => $\Delta G° \approx -3$ kJ/mol

Confomations from MD

Man-β(1-4)-GlcNAc-β(1-4)-[α(1-6)-Fuc]-GlcNAc-β-OH (**sugar d**)



VMD

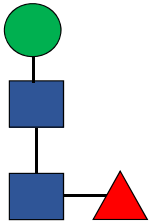
File>New Molecule> Browse: *closed_conf_f2785.pdb*

Mouse > Label > Dihedrals 4
Select the following atoms represented in the VMD OpenGL Display and write down the corresponding value for the dihedral angle (torsions):
on GlcNAc(1) **O4 – C4 – C5 – C6**
on GlcNAc(1)-Fuc(2) **O5 – C1 - O6 – C6 (ϕ)**, **C1 – O6 – C6 – C5 (ψ)**, **O6 – C6 – C5 – C4 (ω)**

> VMD Main Window: deselect “D” which makes the indicated molecule not visible, then load *open_conf_f4810.pdb* and repeat the steps above

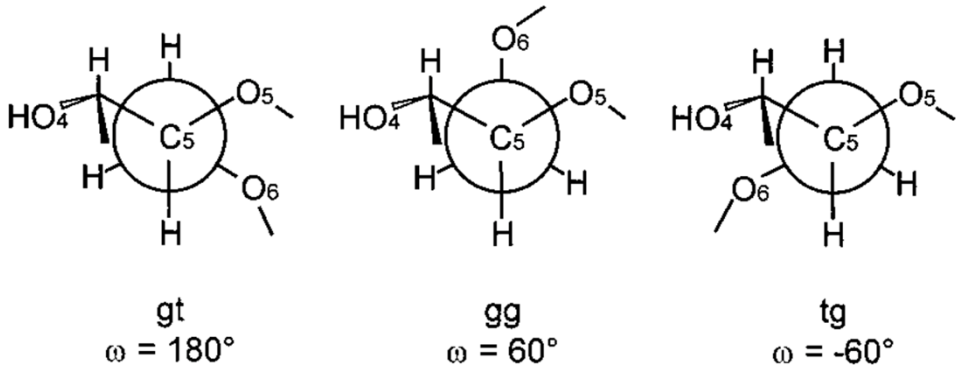
Confomations from MD

Man-β(1-4)-GlcNAc-β(1-4)-[α(1-6)-Fuc]-GlcNAc-β-OH (**sugar d**)



Write your results in a table like the one shown below,

Trajectory frame	**O4C4C5C6**	**O5C1O6C6 (ϕ)**	**C1O6C6C5 (ψ)**	**O6C6C5C4 (ω)**
2785	**59.95**	-87.86	-178.54	53.15 (*gg*)
4810	**158.11**	-80.11	122.24	-69.97 (*tg*)



For nomenclature and other info see also Wormald *et al Chem Rev* (2002) 102:371-386