

Glycans in Rosetta



VANDERBILT
UNIVERSITY

Cristina Elisa Martina
Rosetta Workshop
Meiler Lab



In this introduction:

- Handling Glycans in Rosetta**
- Scoring Glycans in Rosetta**
- Sampling Glycans in Rosetta**
- Tutorial Overview**



Handling Glycans in Rosetta:



Handling Glycans in Rosetta:

Whenever working with glycans:

`-include_sugars`

To read glycans from pdb use these options:

`-alternate_3_letter_codes pdb_sugar`
`-auto_detect_glycan_connections`
`-maintain_links`

To load GLYCAM structure use this option:

`-glycam_pdb_format`



Adding Glycans to Rosetta:

Must include:

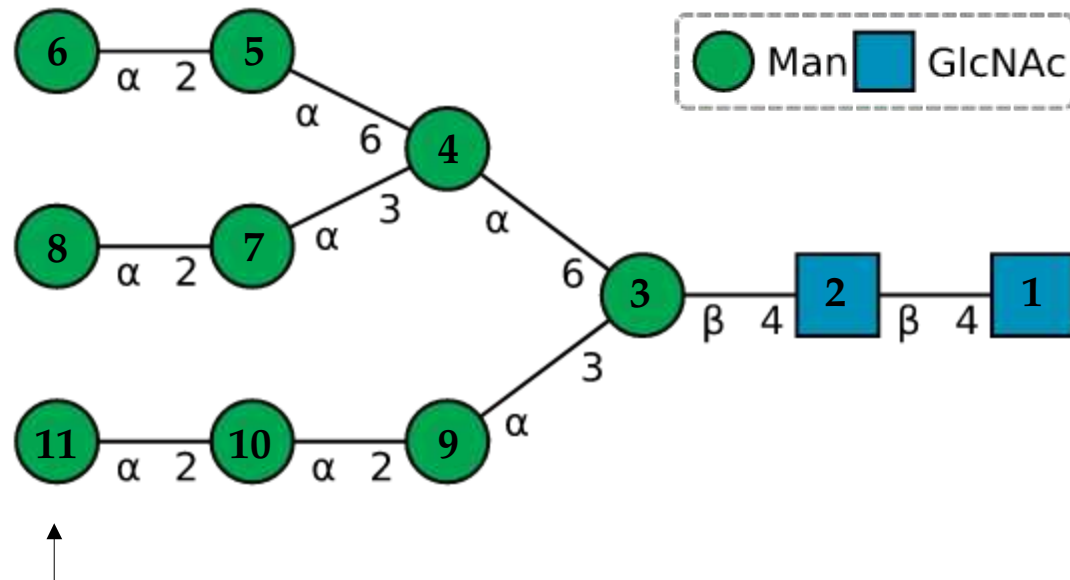
- monosaccharide residue
- stereochemistry
- ring size
- enantiomers
- linkage
- sugar modifications

**IUPAC format (.iupac) or GlycoWorkbench
format (.gws) are accepted**



IUPAC to Rosetta:

1- Sugar residues are read N -> 1, where N is at the extremities and 1 is the root:



In this case N = 11 and we will read from 11 to 1

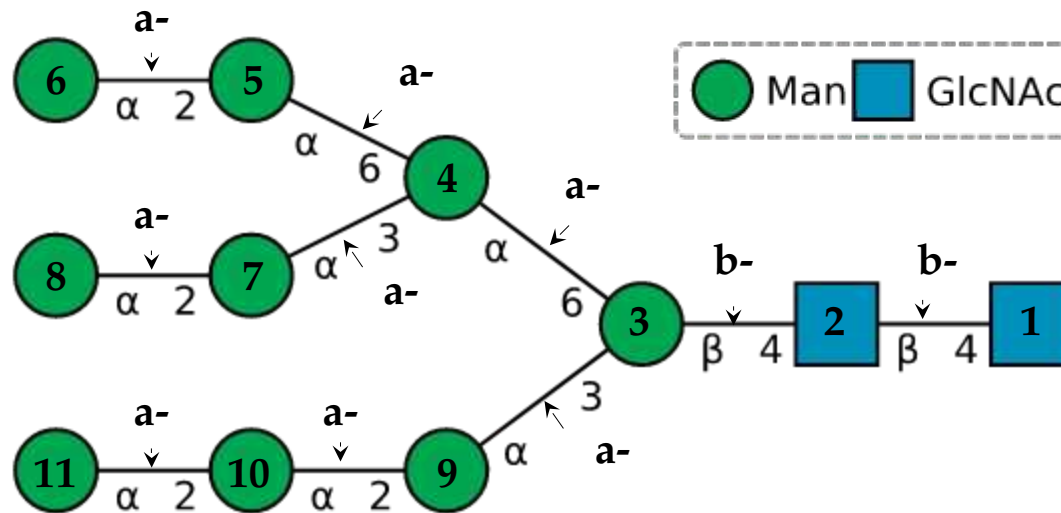


IUPAC to Rosetta:

2- Anomer indication:

α = "a-" or "alpha-"

β = "b-" or "beta-"



IUPAC to Rosetta:

3- Stereochemical indication “L-” or “D-”

4- Sugar name in a 3-letter code format:

Glc = glucose

Man = mannose

Gal = galactose

Fuc = fucose

For a complete list see:

`/rosetta/main/database/chemical/carbohydrates/codes_to_roots.map`



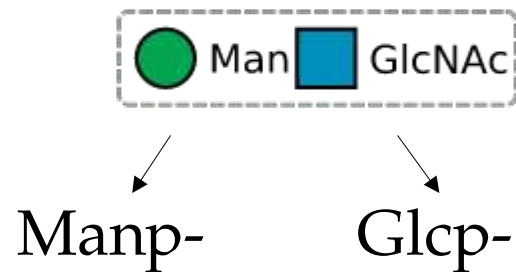
IUPAC to Rosetta:

5- One letter suffix for the ring size:

Furanose = "f-"

Puranose = "p-"

Septanose = "s-"

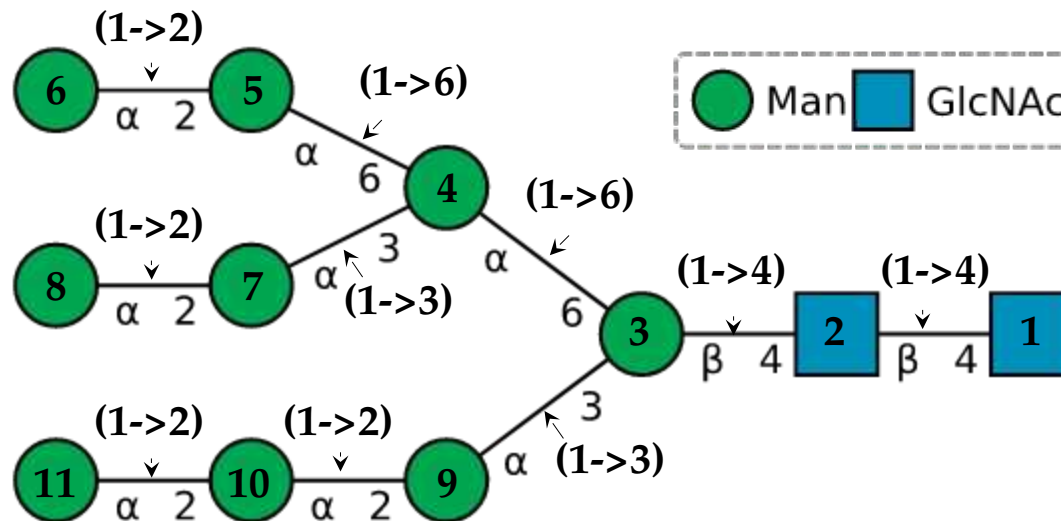


IUPAC to Rosetta:

6- Linkage is as in IUPAC: “(a -> x)”, where:

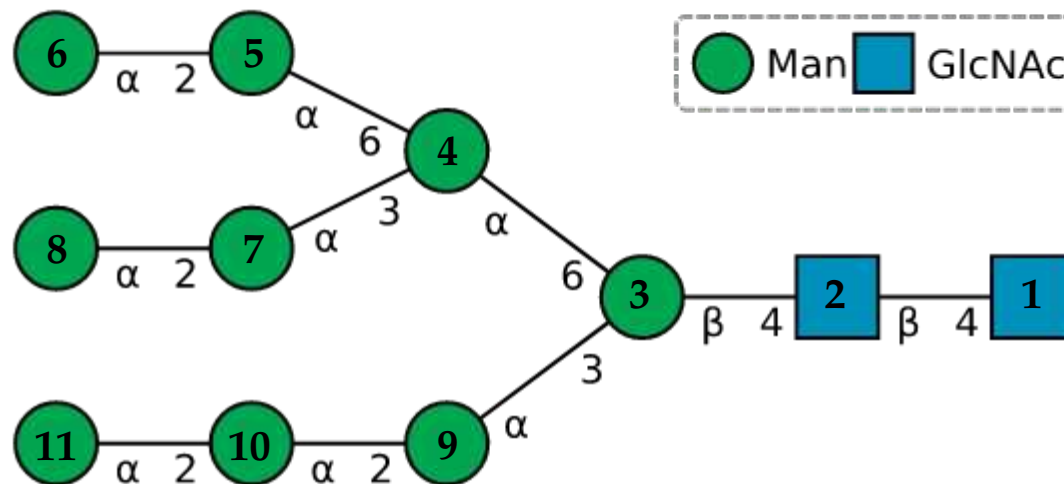
a = anomeric carbon in residue n

x = linked carbon in residue n+1



IUPAC to Rosetta:

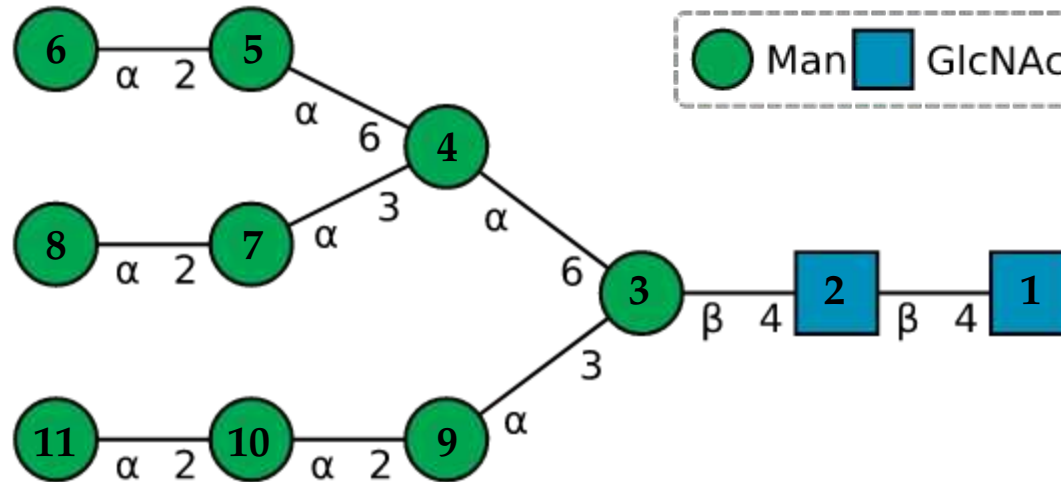
7- Branching is indicated with nested brackets:



11-10-9- [8-7- [6-5] -4] -3-2-1




12



a-D-Manp-(1->2)-a-D-Manp-(1->2)-a-D-Manp-(1->3)-[a-D-Manp-(1->2)-a-D-Manp-(1->3)-[a-D-Manp-(1->2)-a-D-Manp-(1->6)]-a-D-Manp-(1->6)]-b-D-Manp-(1->4)-b-D-GlcpNAc-(1->4)-b-D-GlcpNAc-

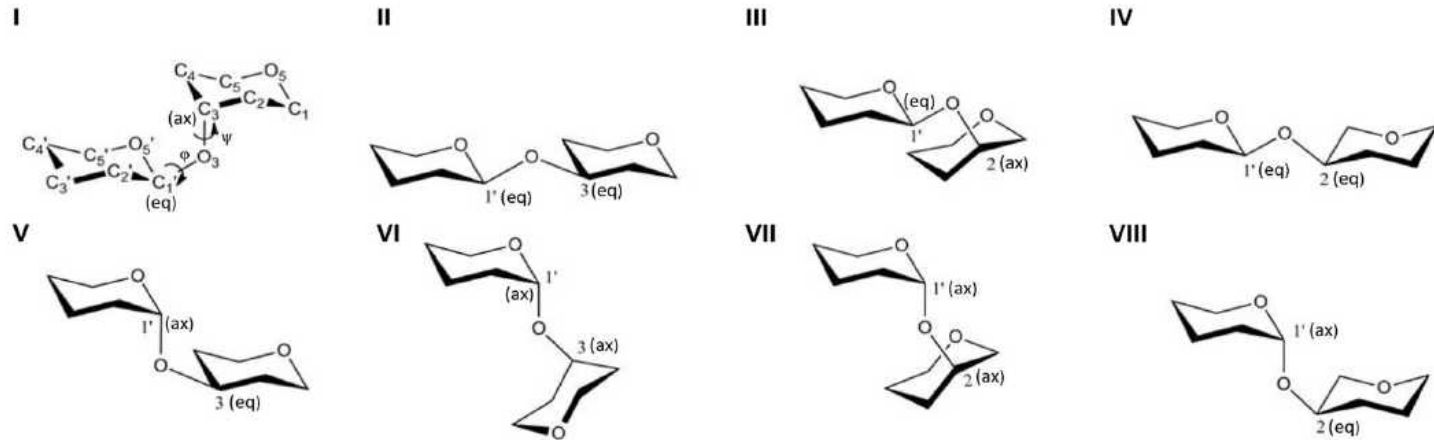
File = /Rosetta/main/database/chemical/carbohydrates/common_glycans/**man9.iupac**

Scoring Glycans in Rosetta:



CarboHydrate Intrinsic energy function:

CHI energy function



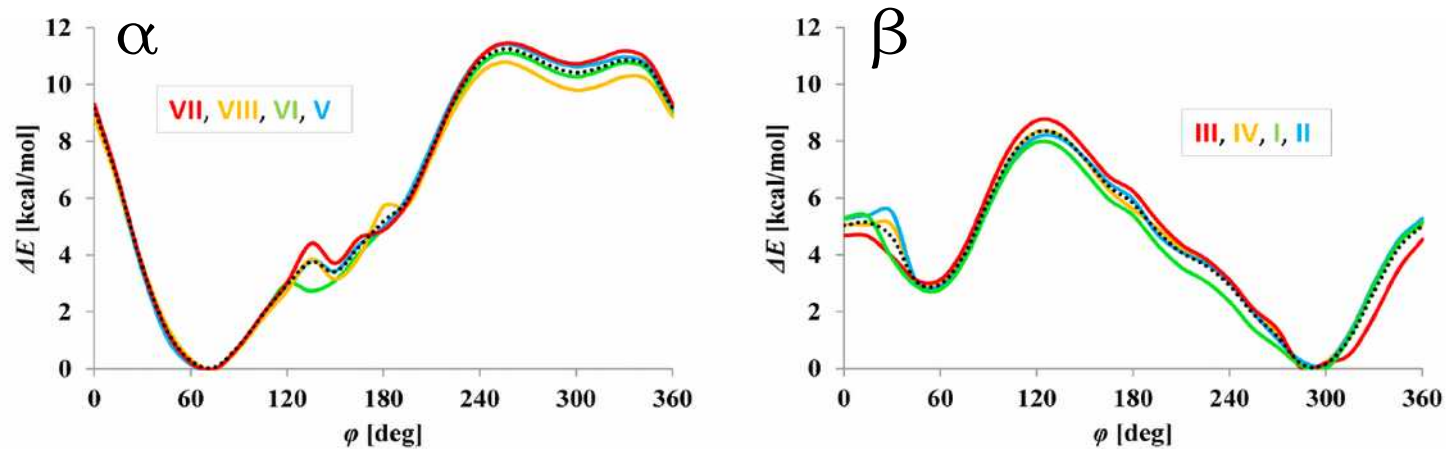
Quantum Mechanical Calculations

- glycosidic linkage: 1-2, 1-3 and 1-4
- anomeric configurations: α and β
- hydroxyl configurations: axial and equatorial

(Nivedha AK, 2014)



CarboHydrate Intrinsic energy function:

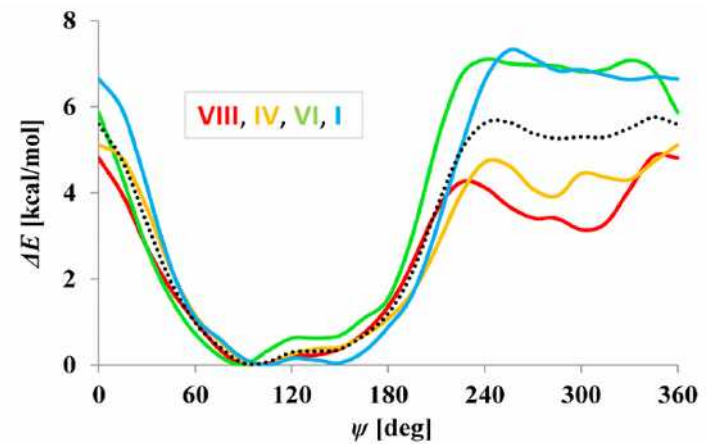
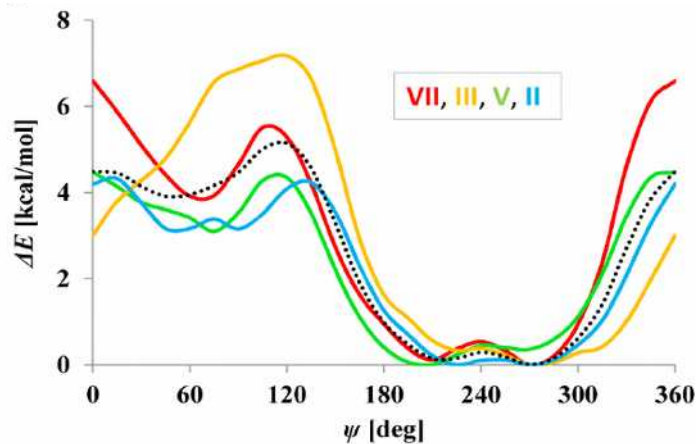


Φ (phi) depends mainly on the anomeric conformation (α or β)



(Nivedha AK, 2014)

CarboHydrate Intrinsic energy function:



ψ (psi) and ω (omega) depend mainly on the hydroxyl configurations (axial or equatorial)

(Nivedha AK, 2014)



Sugar score term in Rosetta:

CHI energy function was implemented in
Rosetta in the `sugar_bb` score term

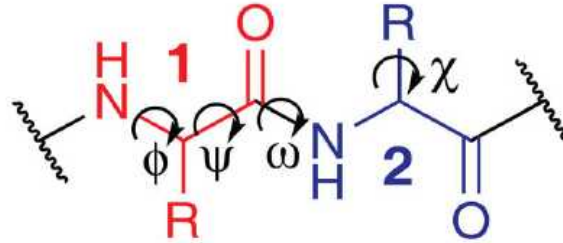
Automatically called with the option:
`-include_sugars`



Sampling Glycans in Rosetta:



Rosetta Sampling for proteins:



Backbone torsion angles:

- Φ (phi)
- ψ (psi)
- ω (omega)

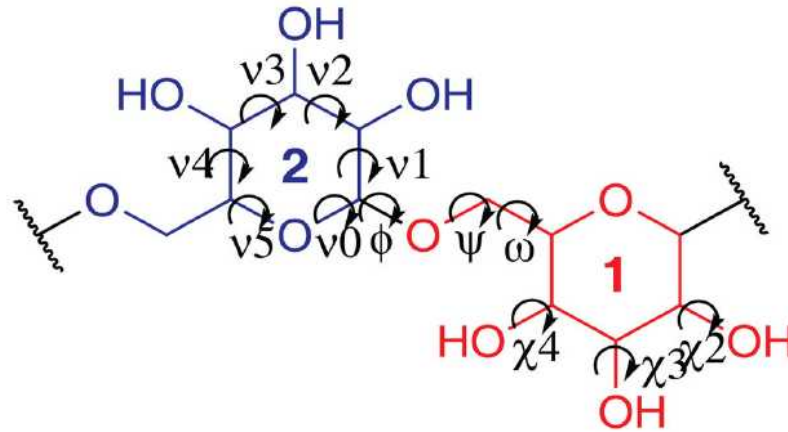
Side-chain torsion angles:

- χ (chi)
- extra torsion angles for 5 residues

Proteins Backbone is linear (no branching) in most cases!



Rosetta Sampling for glycans:



Backbone torsion angles:

- Φ (phi)
- ψ (psi)
- ω (omega)
- ν (nu) multiple

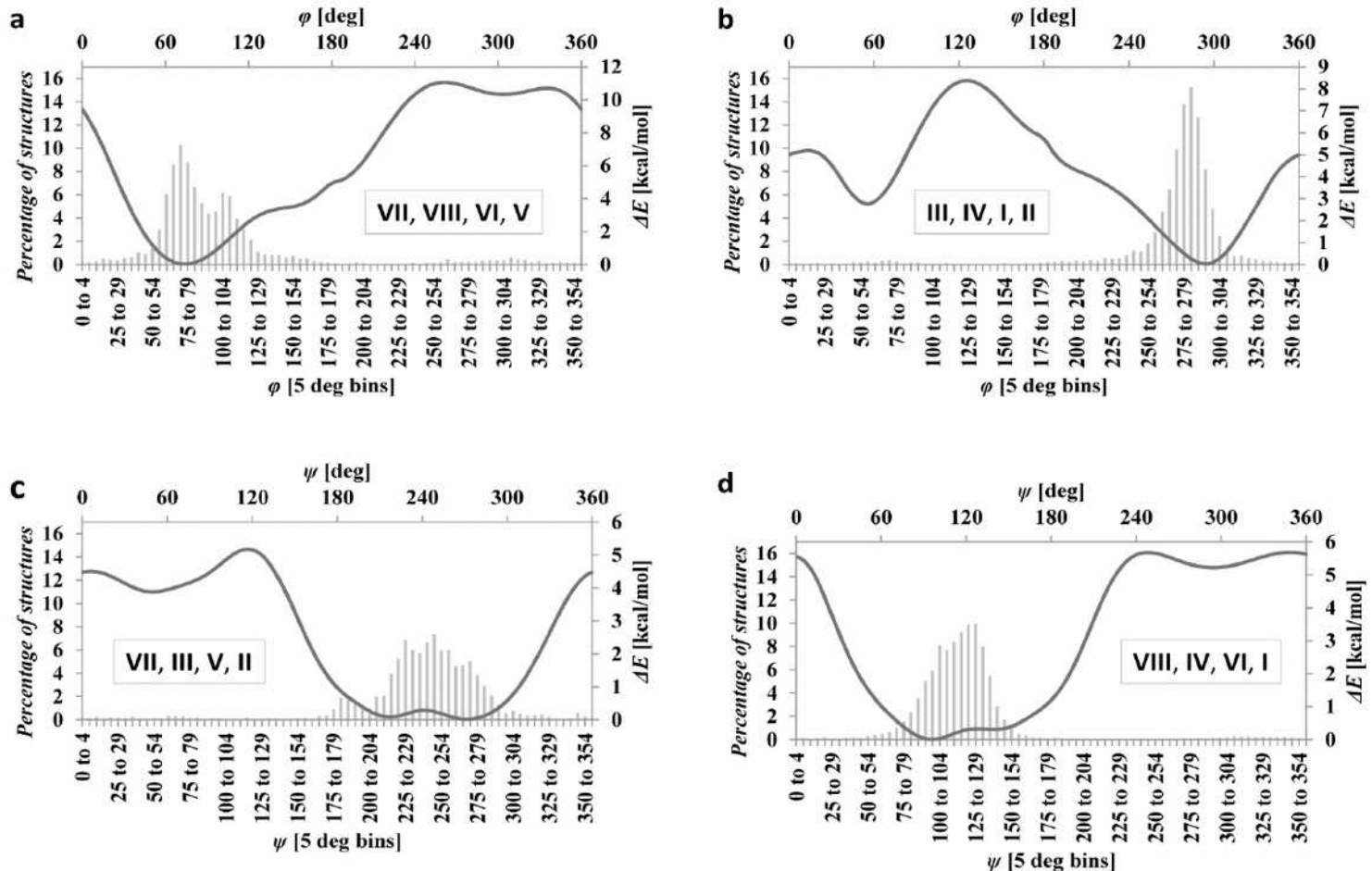
Side-chain torsion angles:

- χ (chi) multiple
- extra torsion angles for modified sugars

Glycan Backbone can be branched!



Rosetta Sampling for glycans:



(Nivedha AK, 2014)

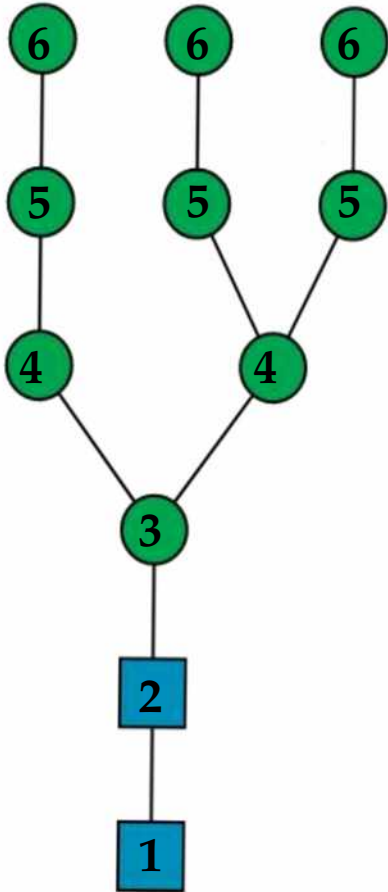


Rosetta Sampling for glycans:

**CHI energy function
distributions are converted
into phi/psi/omega
probabilities**



GlycanTree in Rosetta:



For modeling, glycans are divided in layers and sampled individually:

1- all sugars are set as virtual (no scoring)

2- First layer is sampled and scored

3- Second layer is sampled and scored

• • •

n- all tree is sampled and scored

To select the layer:

- GlycanLayerSelector within Residue Selector

Tutorial Overview:



Tutorial Overview:

Bee Hyaluronidase is an interesting enzyme for therapy.

It can break down sugars in the extracellular matrix and allow the delivery of larger drugs to region of interest.

However it present an highly immunogenic site, that is targeted by the immune system.

We will use glycans to shield the immunogenic epitope in order to hide it from the immune system.



Tutorial Overview:

1. Download, trim, clean and score the complex in pdb 2J88
2. Relax and score again the complex
3. Insert the sequon to add the glycosylation site (in two ways)
4. Add the glycan to the structure and model its conformation



Tutorial Overview:

1. Download, trim, clean and score the complex in pdb 2J88
2. Relax and score again the complex

You should already be familiar with these steps to set up the protein for Rosetta



Tutorial Overview:

3. Insert the sequon to add the glycosylation site (in two ways)

The sequon is the amino acid triplet recognized by the enzyme Glycosyl Transferase. This enzyme add a glycan chain to an Asparagine through a N-glycosidic linkage.

The sequon for N-glycosylation is:

NxT/S

(x = any aa except Proline)



Tutorial Overview:

4. Add the glycan to the structure and model its conformation

After mutating the aa sequence with the sequon triplet, we will add a glycan chain (man5) and then model it.



Bibliography - Glycans in Rosetta:

J. W. Labonte et al., **“Residue-Centric Modeling and Design of Saccharide and Glycoconjugate Structures”**, J Comput Chem, vol. 38, no. 5, pp. 276–287, Feb. 2017, doi: 10.1002/jcc.24679.

B. Frenz et al., **“Automatically Fixing Errors in Glycoprotein Structures with Rosetta”**, Structure, vol. 27, no. 1, pp. 134-139.e3, Jan. 2019, doi: 10.1016/j.str.2018.09.006.

M. L. Nance, et al. **“Development and Evaluation of GlycanDock: A Protein-Glycoligand Docking Refinement Algorithm in Rosetta”** J Phys Chem B, Jun. 2021, doi: 10.1021/acs.jpcb.1c00910.

