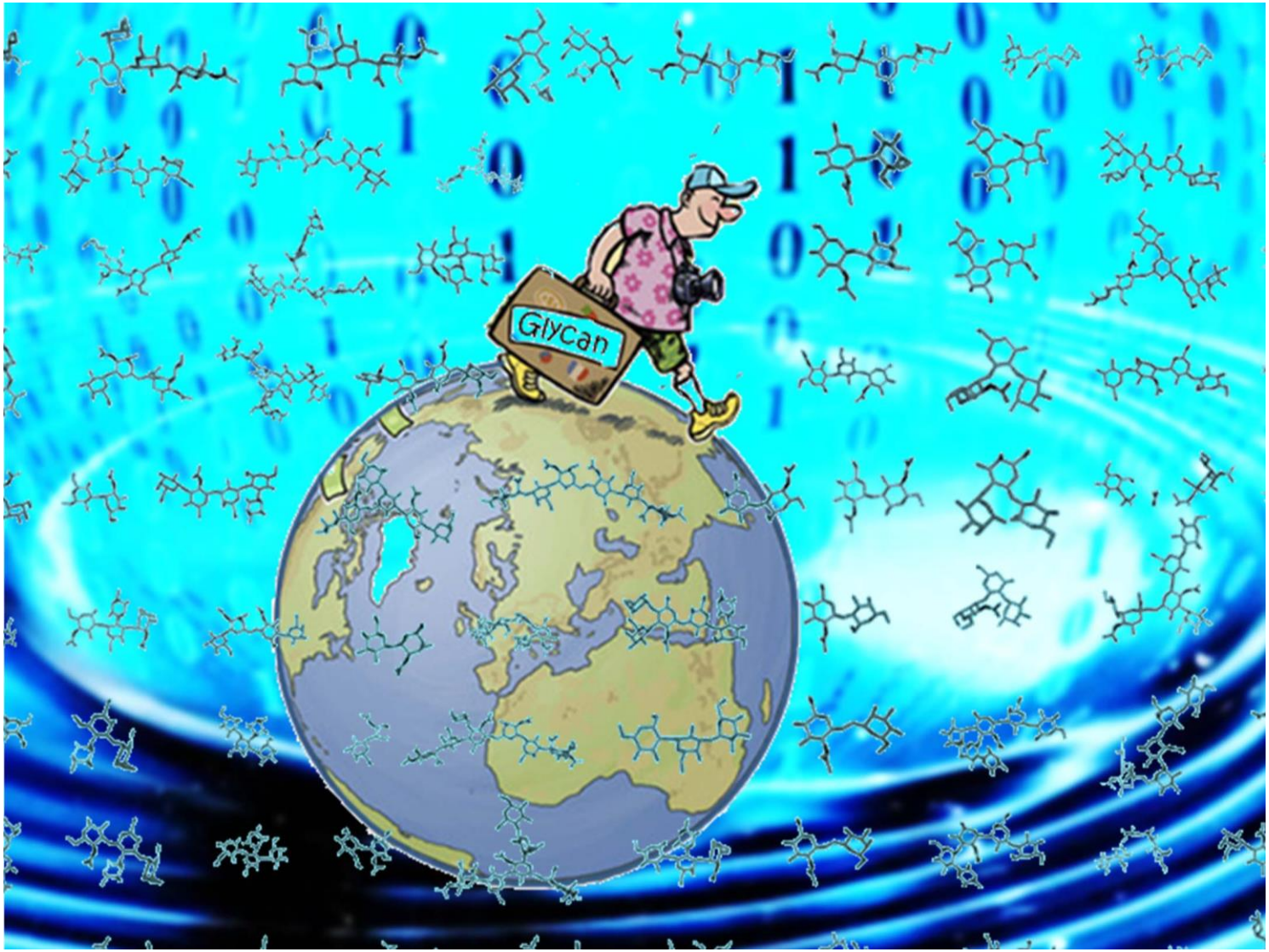
A Traveler-Guide to Carbohydrates in the Cyber-Space



Serge Pérez & Frédérique Lisacek

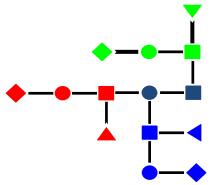
Information Content

Composition

NeuAc)3 (Fuc)3 (GlcNac)3 (Gal)4 (GalNac)1 Ser

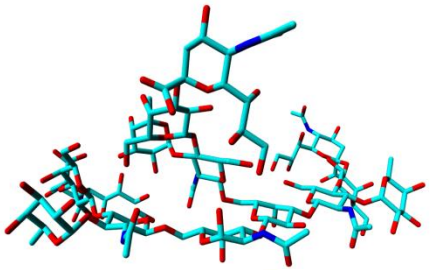
MS, HPLC, CE

Topology



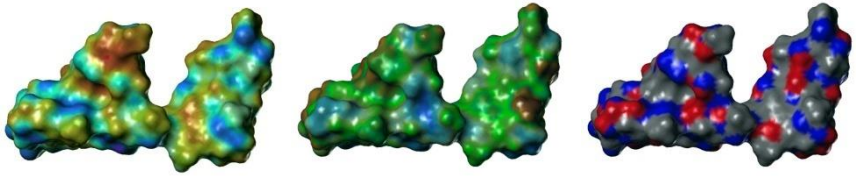
MS, HPLC
CE, NMR

Conformation



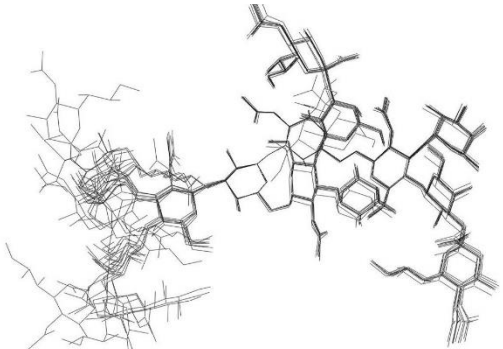
NMR
XRAY
MS

Properties



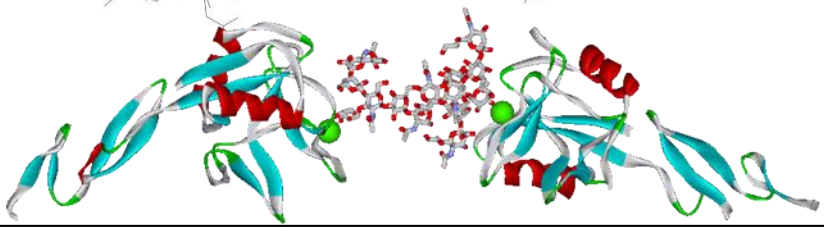
MM
CAMD

Dynamics



NMR
MM
MD

Interactions

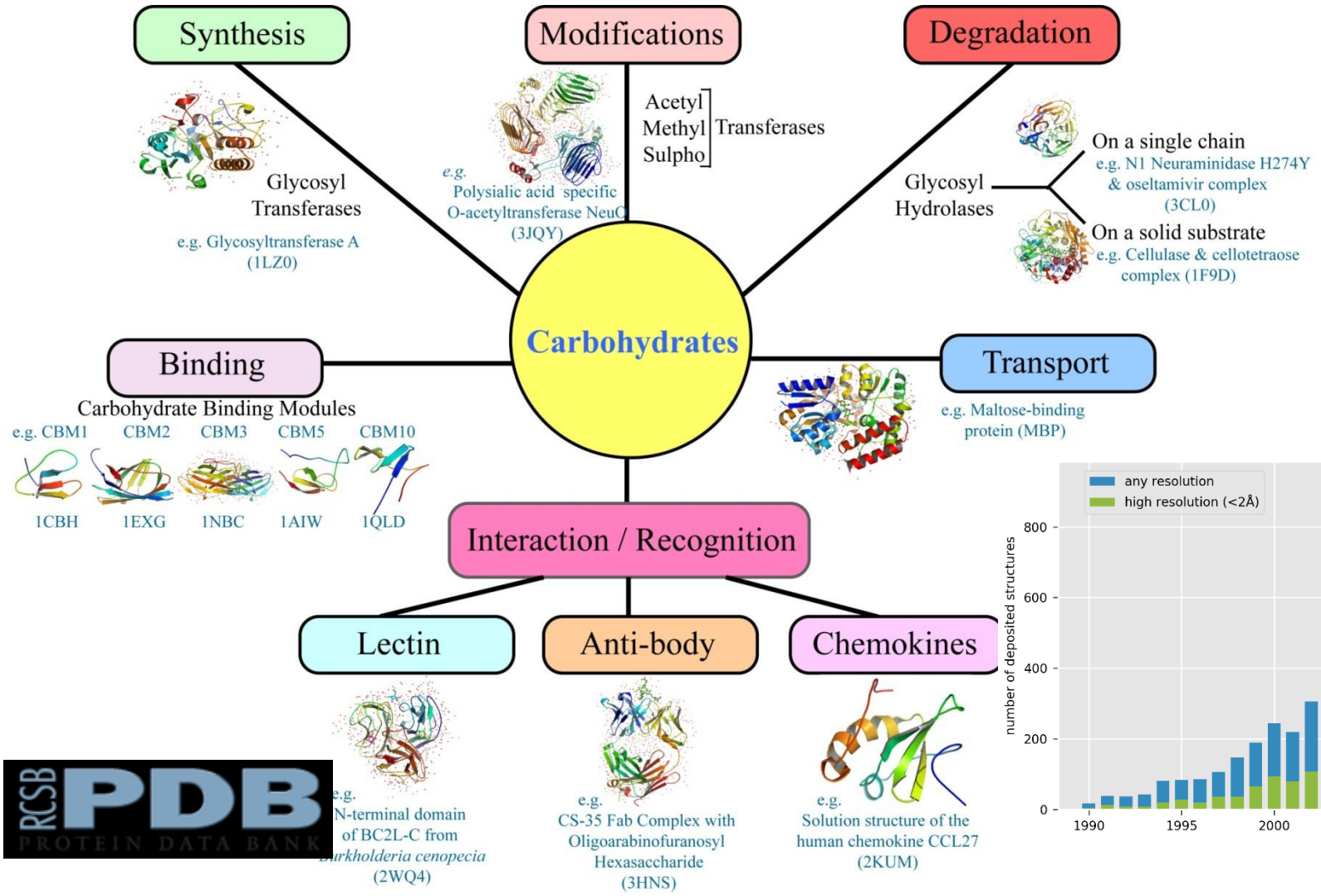


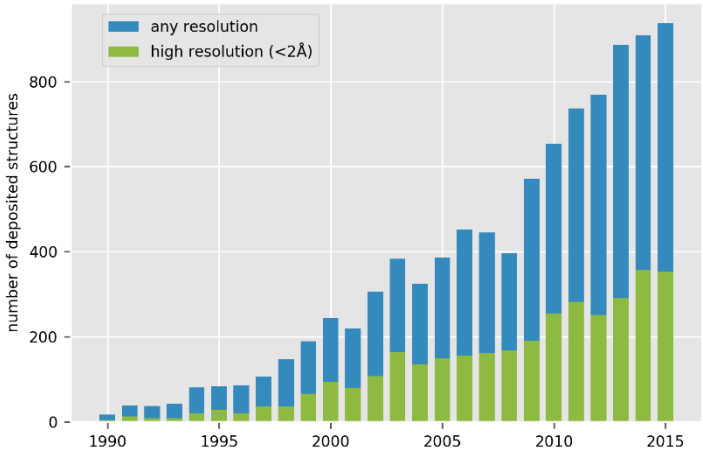
XRAY
MM

Methods

Crystalline Conformations of Oligosaccharides in Proteins

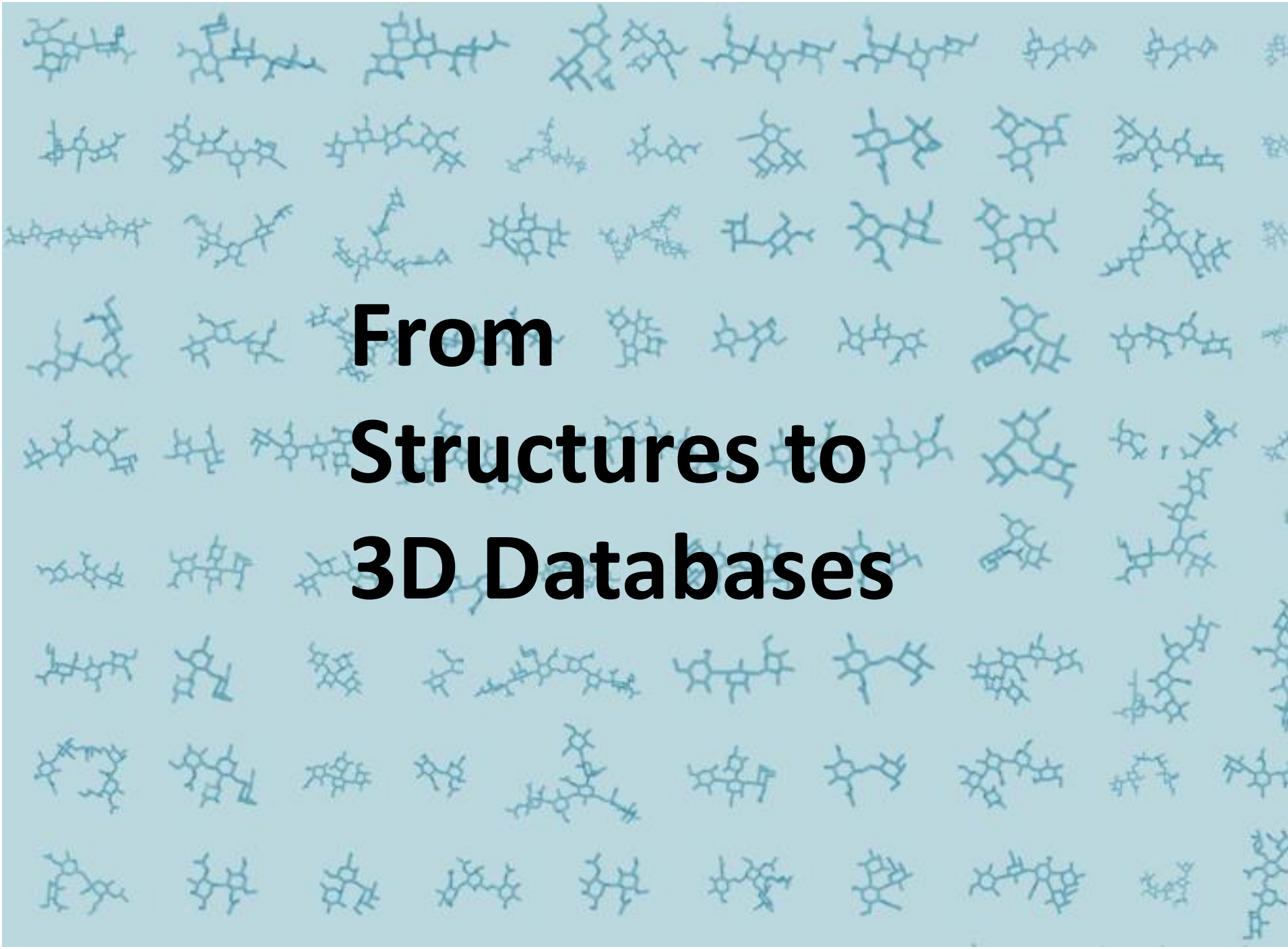
Protein-Carbohydrate Interactions



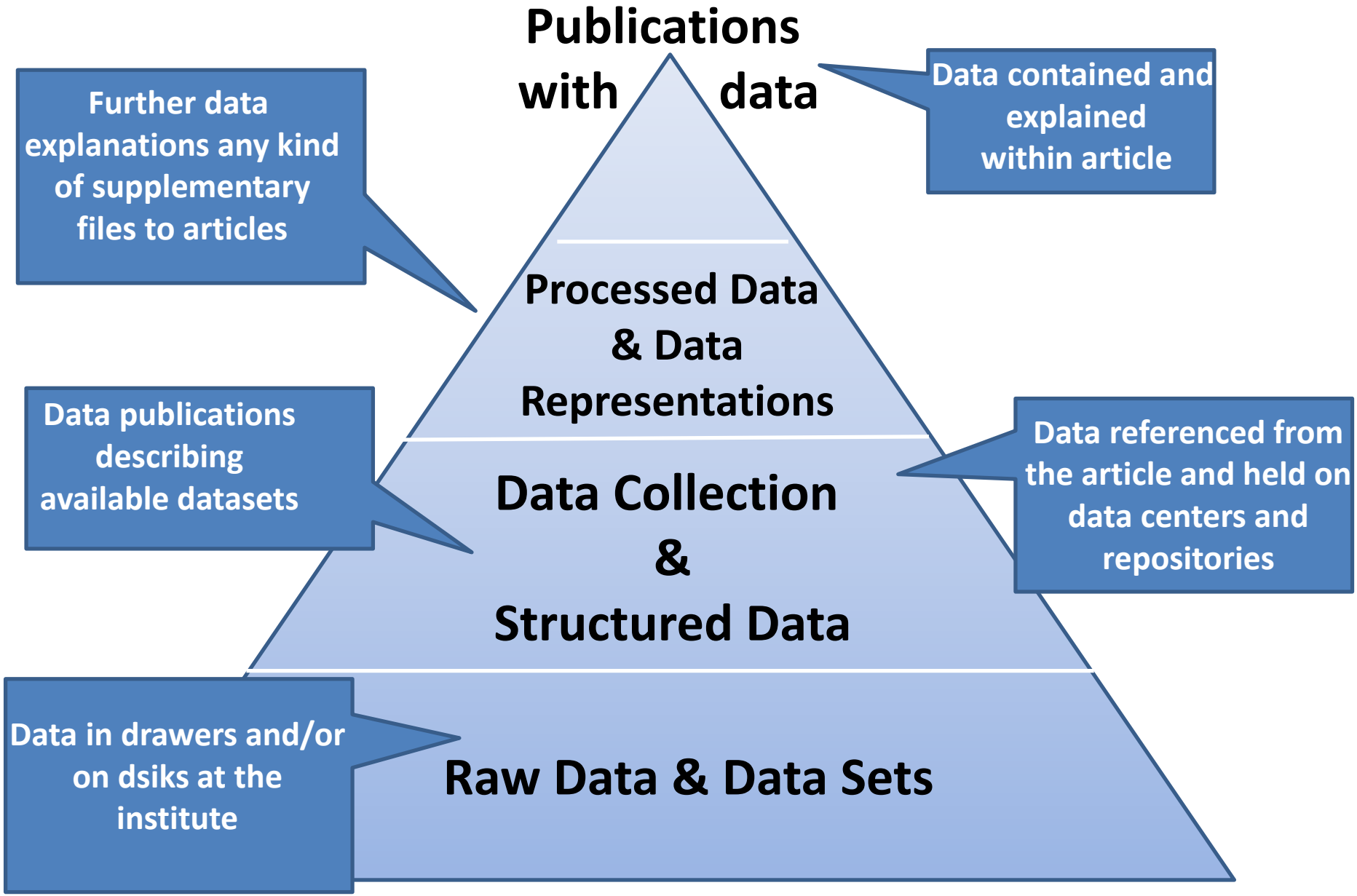


RCSB PDB PROTEIN DATA BANK

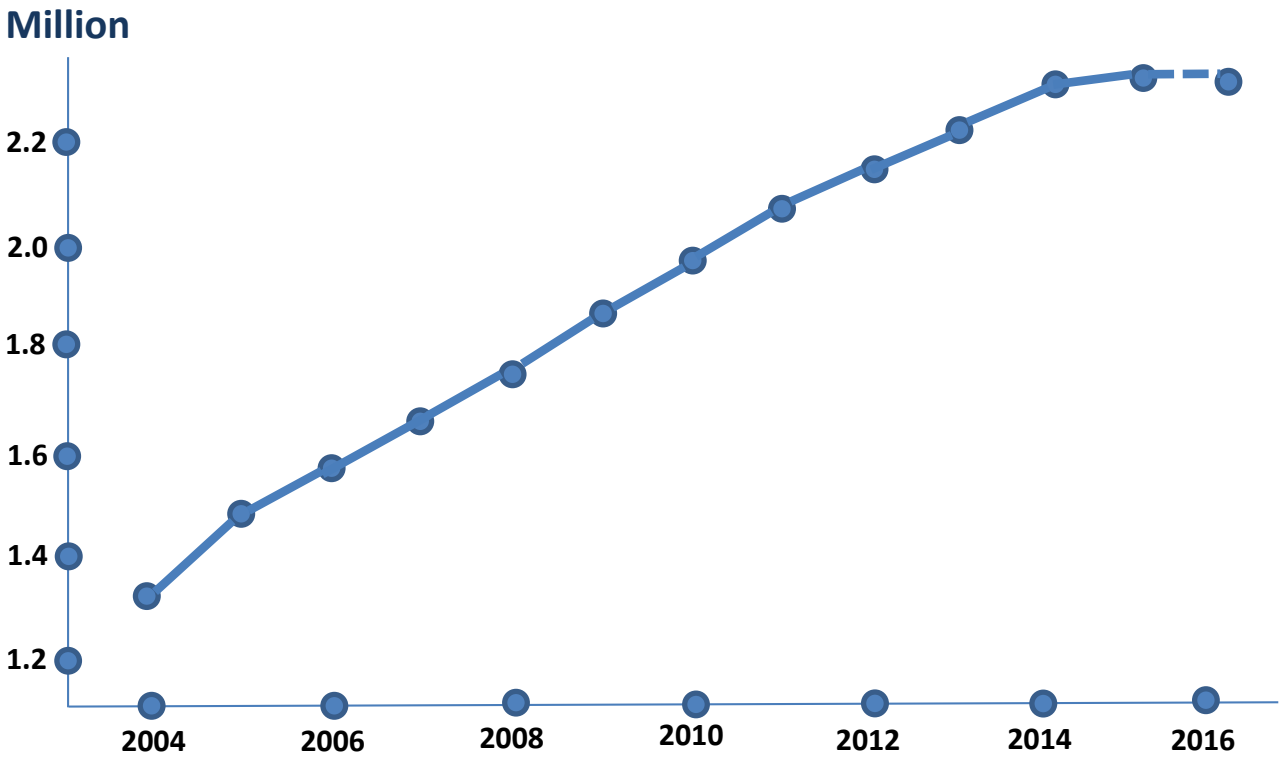
Protein Data Bank : http://www.rcsb.org/pdb/home/home.do


From
Structures to
3D Databases

An avalanche of data…



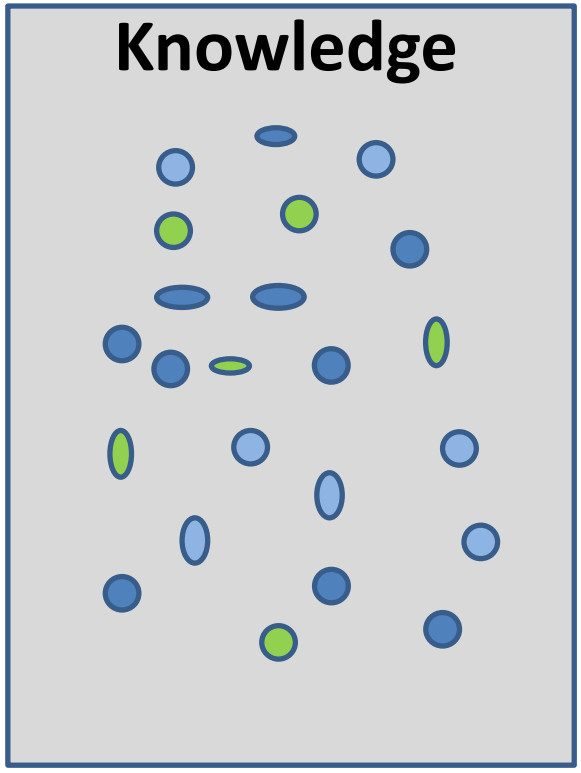
An avalanche of data…

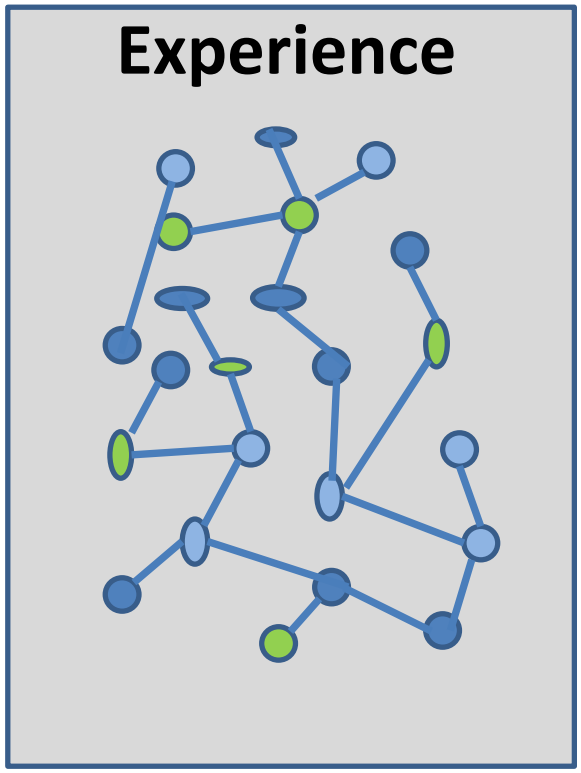


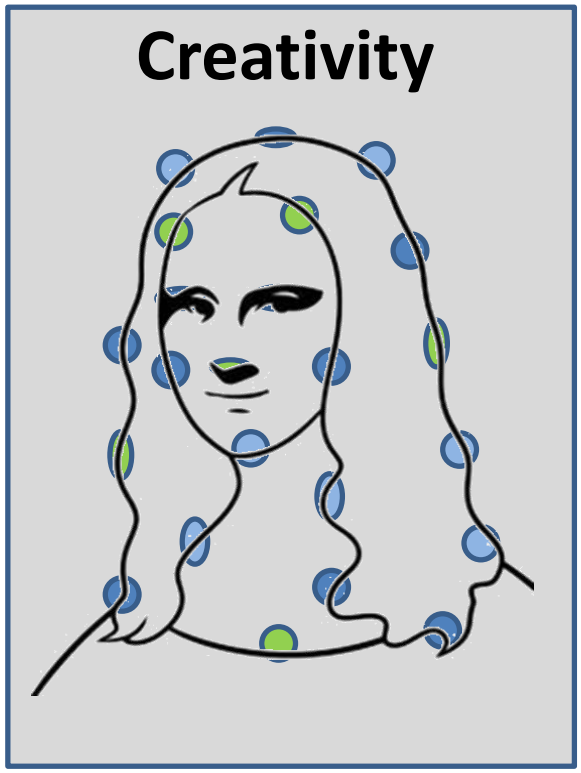
Global scientific output doubles every nine-years

Number of active researchers world-wide 8 Millions

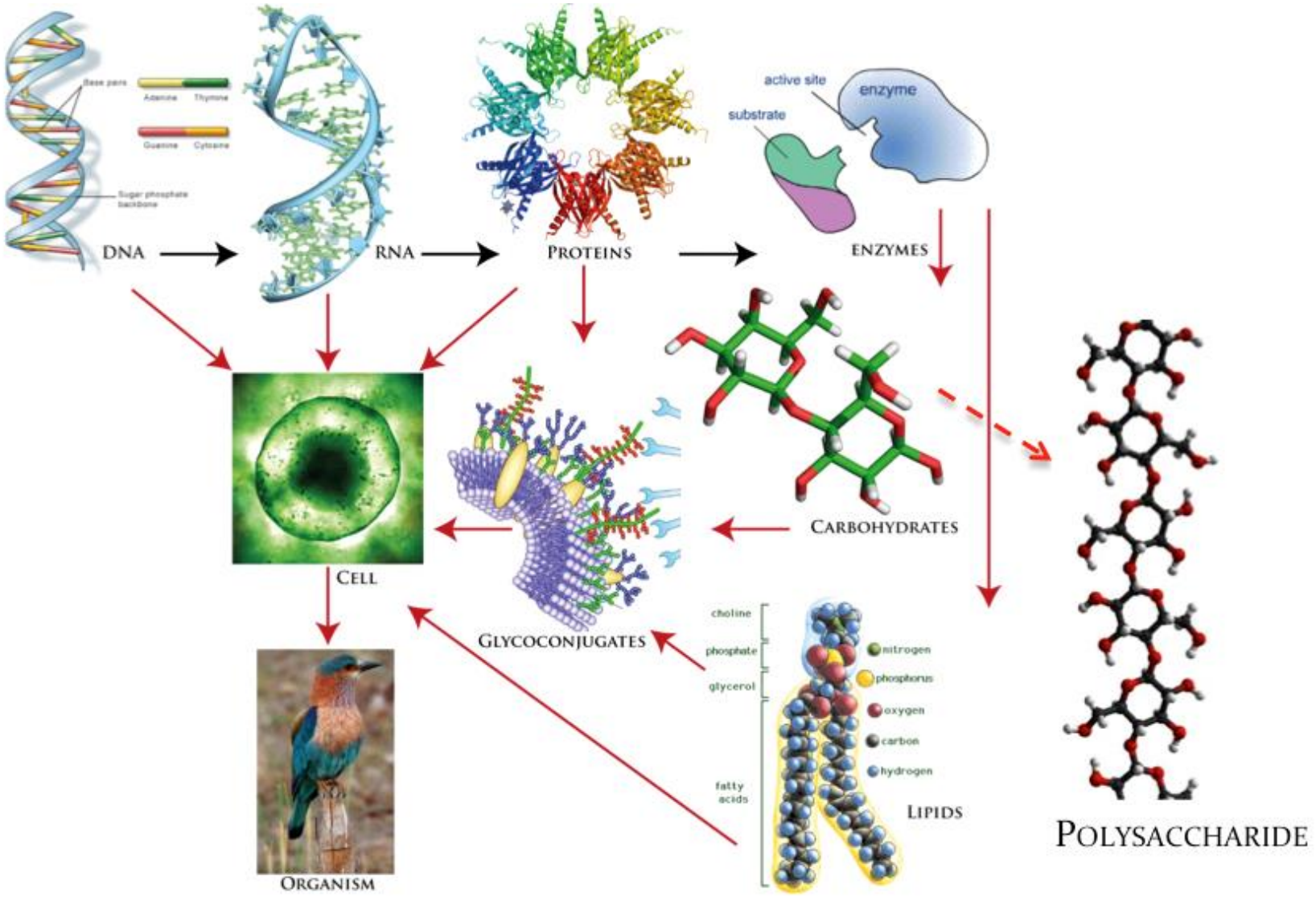
Knowledge, Experience, Creativity





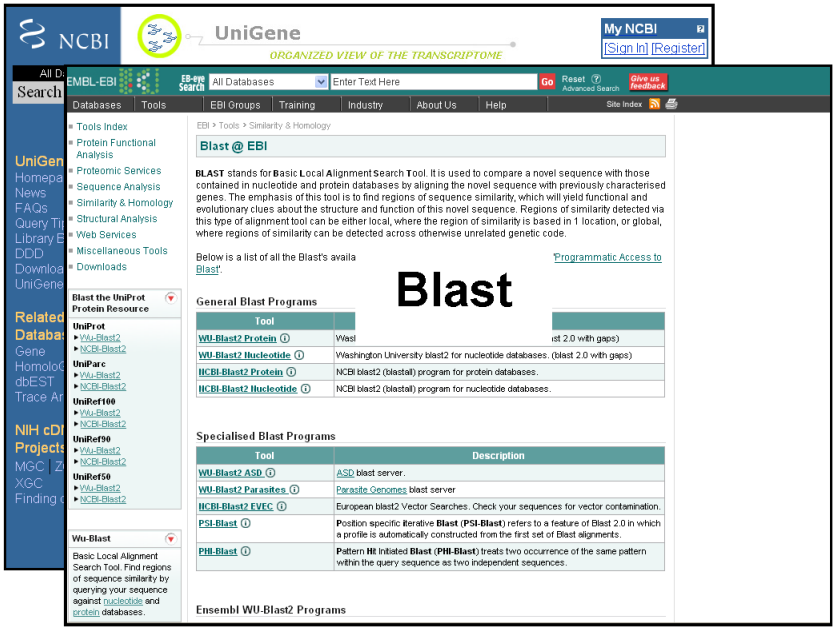


Carbohydrates in the Scheme of the Central Dogma of Life

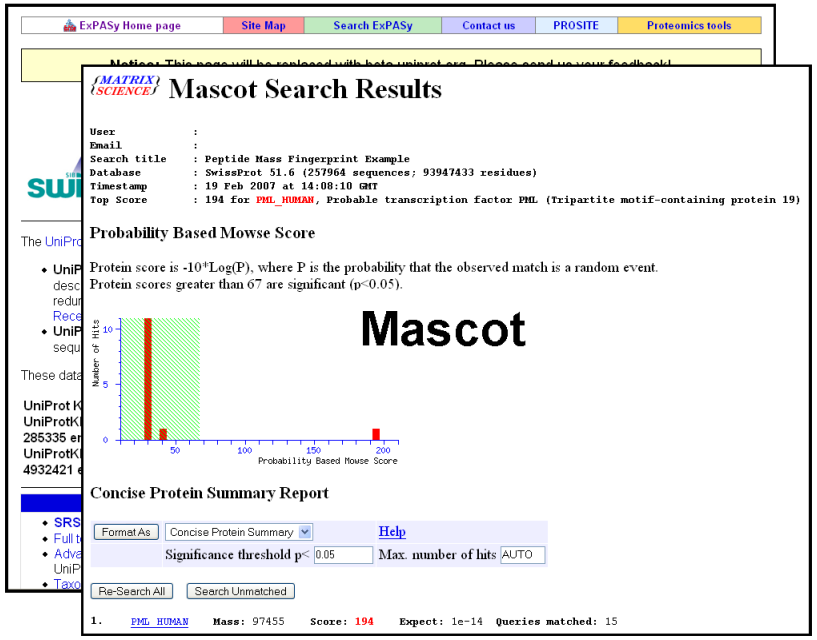


Glycoinformatics

Genomics



Proteomics



Glycomics

Challenges for Glycoinformatics

Structures as Primary Access Key

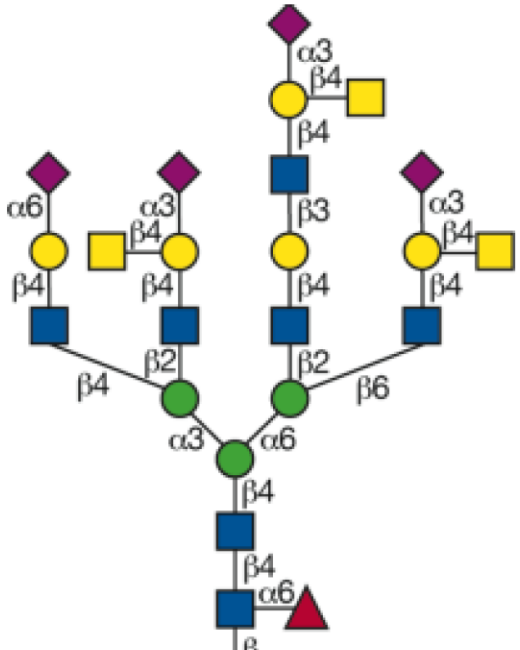
Bioinformatic

Galectin-1
source organism="Homo sapiens"
gene gene="LGALS1"
Site /site_type="binding"
/note="Beta-galactoside (Potential)."

Sequences of residues

1 MACGLVASNL NLKPGECLRV RGEVAPDAKS
31 FVLNLGKDSN NLCLHFNPRF NAHGDANTIV
61 CNSKDGGAWG TEQREAVFPF QPGSVAEVCI
91 TFDQANLTVK LPDGYEFKFP NRLNLEAINY
121 MAADGDFKIK CVAFD

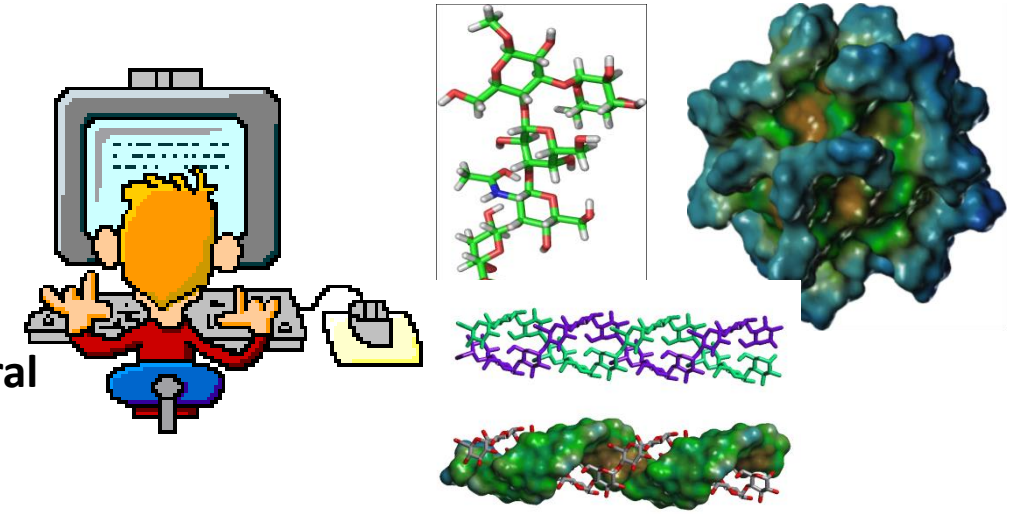
Glycoinformatic



Topology of Residues

e-Glycoscience

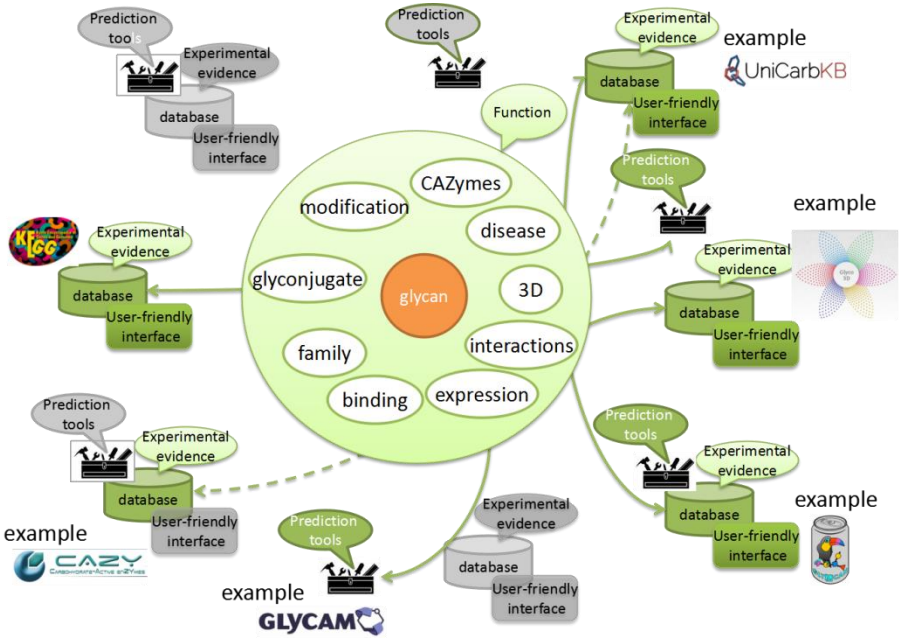
Continued advances in molecular modeling has generated insights for understanding glycan structures and properties. Robust, validated informatics tools are developed in to enable accurate and fast determination of complex **carbohydrate and glycoconjugate structural prediction, computational modeling, and data mining.**



Database have been developedand cover including mammalian, plant and microbial carbohydrates and glycoconjugates.

The carbohydrate **structural database needs to be fully cross-referenced with databases that provide complementary biological information**.

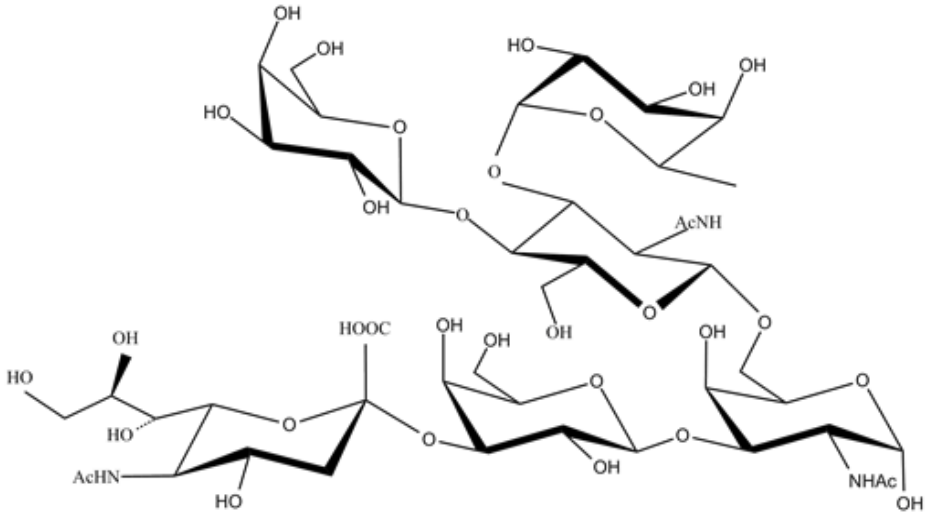
There should be a requirement for deposition of new structures into the database using a reporting standard for minimal information.

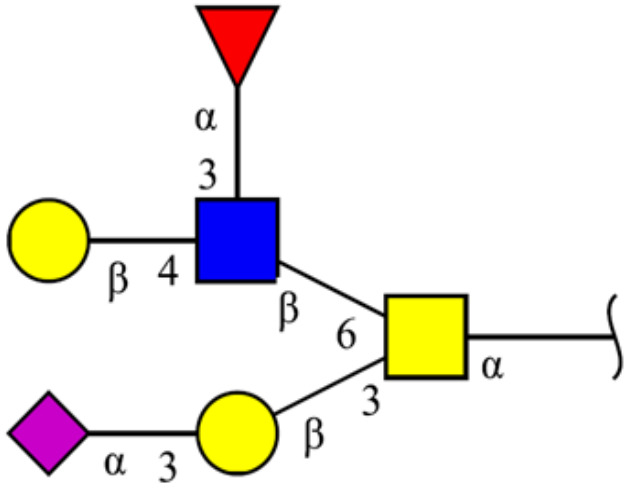


Encoding of Glycan Structures

Lewis X and Sialyl Acid on Core 2

Neu5Ac a2-3 Gal b1-3 (Gal b1-4 (Fuc a1-3) GlcNAc b1-6) GalNAc





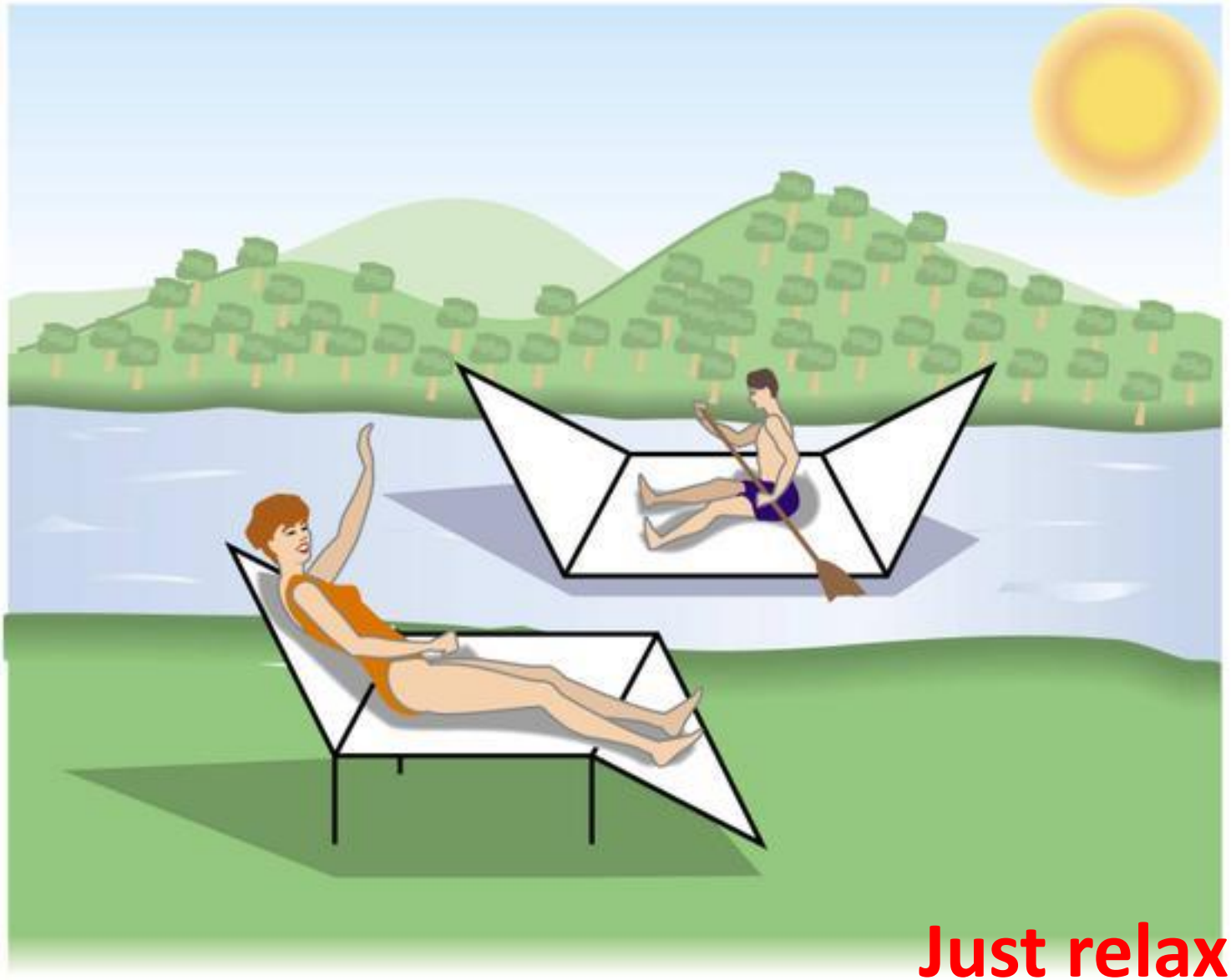
RES
1b:a-dgal-HEX-1:5
2s:n-acetyl
3b:b-dgal-HEX-1:5
4b:a-dgro-dgal-NON-2:6|1:a|2:keto|3:d
5s:n-acetyl
6b:b-dglc-HEX-1:5
7s:n-acetyl
8b:a-lgal-HEX-1:5|6:d
9b:b-dgal-HEX-1:5
LIN
1:1d(2+1)2n
2:1o(3+3)3d
3:3o(3+2)4d
4:4d(5+1)5n
5:1o(6+1)6d
6:6d(2+1)7n
7:6o(3+1)8d
8:6o(4+1)9d

GlycoCT

Symbol Nomenclature for Graphical Representation of Glycans (2015), *Glycobiology*, **25,** 1323-1324

A. VARKI, R.D. CUMMINGS, M. AEBI, N.H. PARKER, P.H. SEEBERGER, J.D. ESKO, P. STANLEY, G. HART, A. DARVILL, T. KINOSHITA, J.J. PRESTEGARD, R.L. SCHNAAR, H.H. FREEZE, J.D. MARTH, C.R. BERTOZZI, M.E. ETZLER, M. FRANK, J.F.G. VLIEGENTHART, T. LUTTEKE, S. PEREZ, E. BOLTON, P. RUDD, J. PAULSON, M. KANEHISA, P. TOUKACH, K.F. AOKI-KINOSHITA, A. DELL, H. NARIMATSU, W. YORK, N. TANIGUCHI & S. KORNFELD,

No Comments …….



Just relax ……..

Extending the Symbolic Representation of Monosaccharides

Residue Letter Name: Rib, Ara, Xyl, Lyx, All, Alt, Glc, Man, Gul, Ido, Gal, Tal,….

[O-ester and ethers]: (when present) are shown attached to the symbol with a number, e.g.
6Ac for 6-*O*-acetyl group, 3S for 3-*O*-sulfate group
6P for 6-*O*-phosphate group, 6Me for 6-*O*-Methyl group
36Anh for 3,6-anhydro, Pyr for pyruvate group

Absolute Configuration: D or L

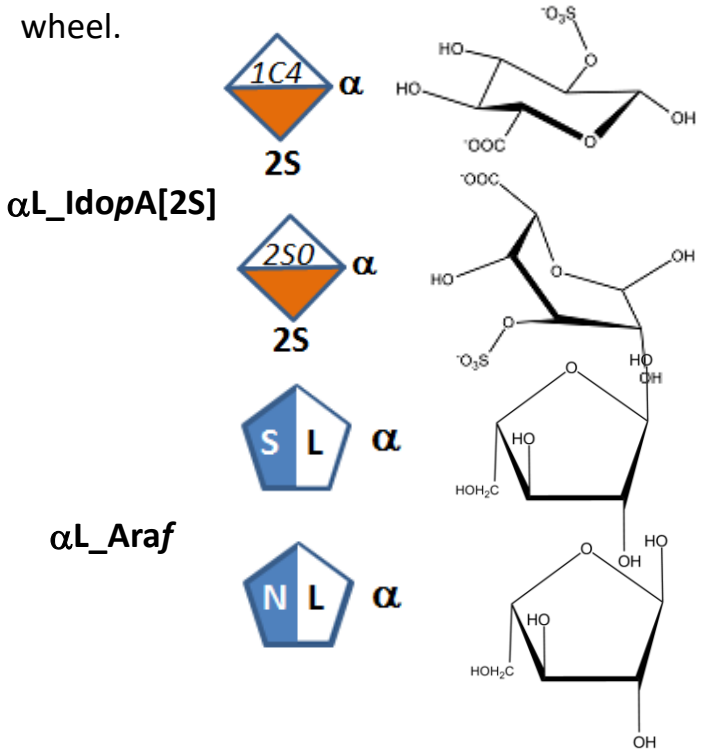
The D-configuration for monosaccharide and the L configuration for Fucose and Idose are implicit and does not appear in the symbol. Otherwise the L configuration, is indicated in the symbol, as in the case of Arabinose or L-Galactose.
For those occurring in the furanose form, a letter *N* or *S* is inserted in the symbol, indicating the northern *(N)* or Southern (*S*) conformation of the five membered ring.

Anomeric Configuration.

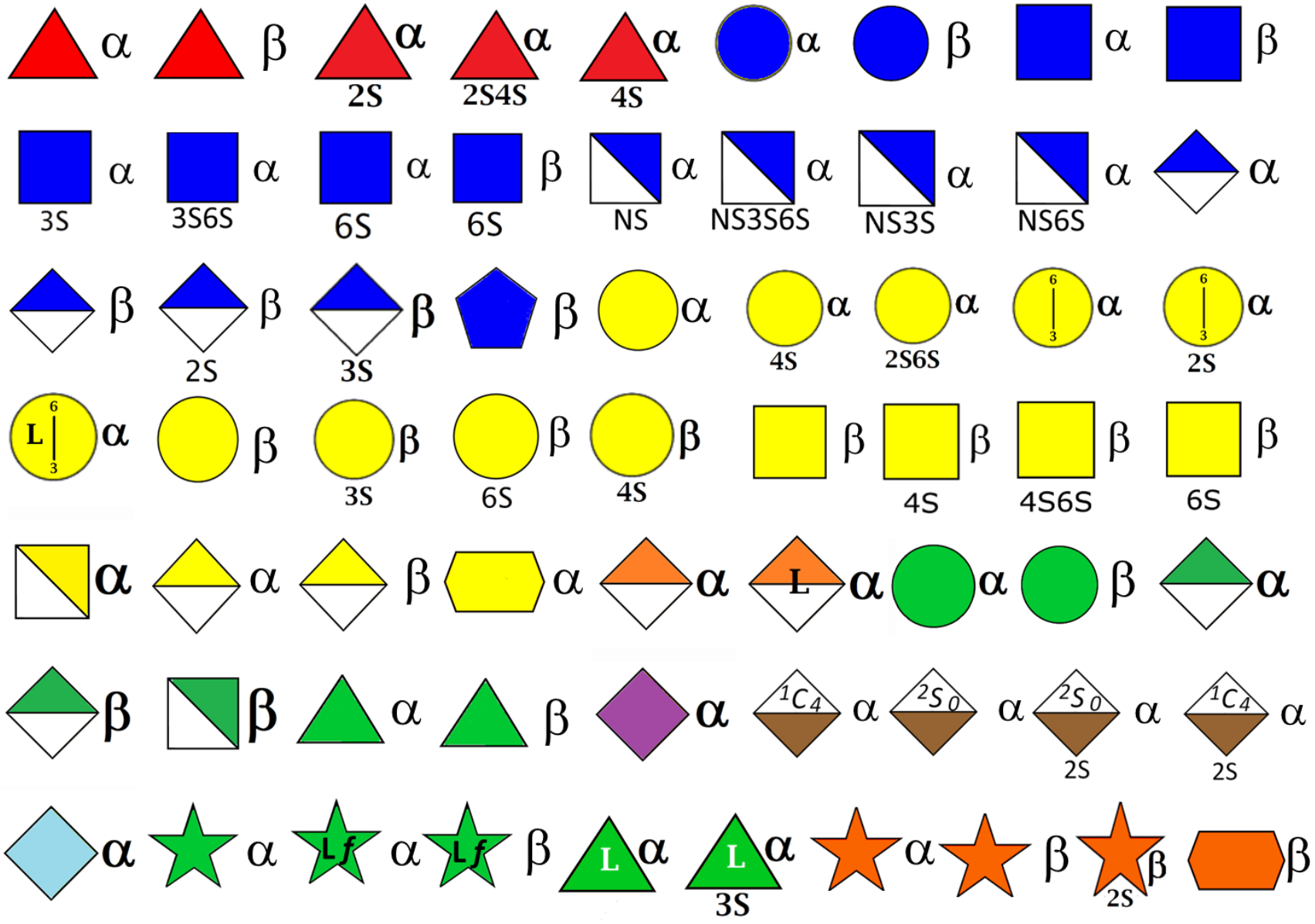
The nature of the glycosidic configuration (α or β) is explicitly set within the symbol.

Ring Conformation.

All pyranoses in the D-configuration are assumed to have $^4C_1$ chair conformation; those in the L configuration are assumed to have $^1C_4$ chair conformation. Otherwise, the ring conformation is indicated in the symbol, as $^2S_0$ in the case of α-L-Idopyranose.
N or *S* indicates the conformation of the five membered rings on the conformational wheel.



More than 150 Monosaccharides



Tools and DataBases

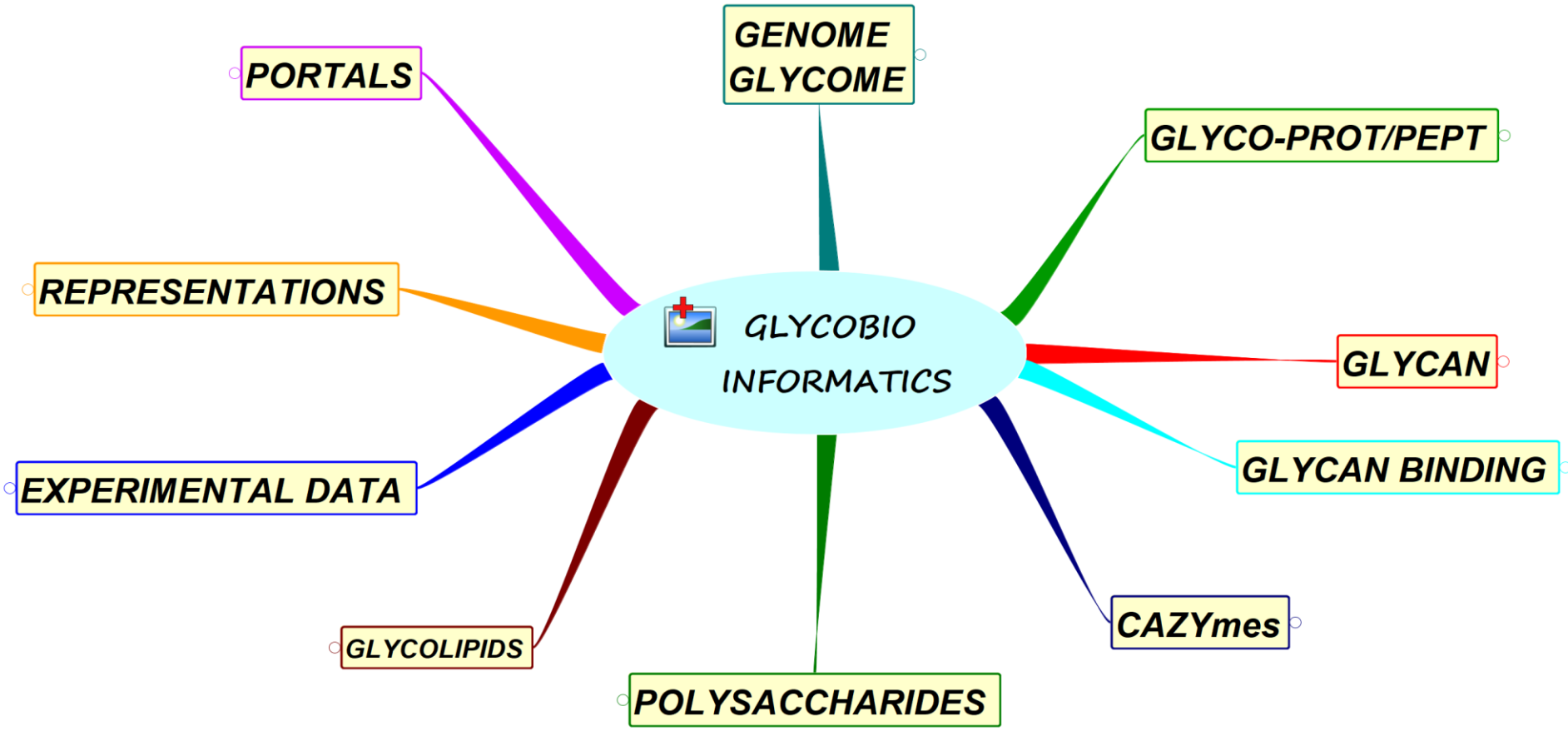
Protocole: Search on Internet Tools and DataBases in Glycosciences

Sites should be up to date (checked) and freely available.

Results: 94 Tools 32 DataBases

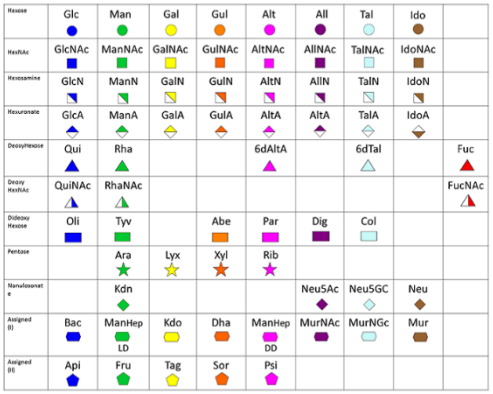
Analysis and Clustering

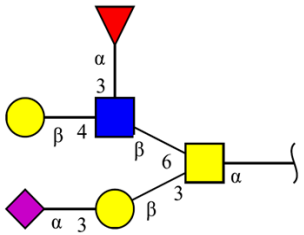
Tools and DataBases



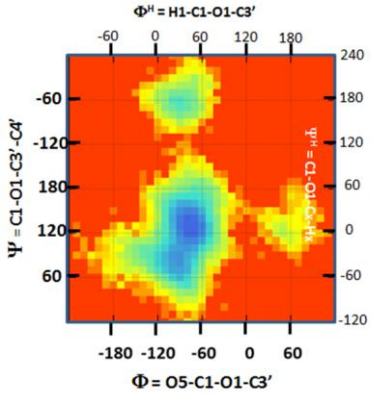
Representations

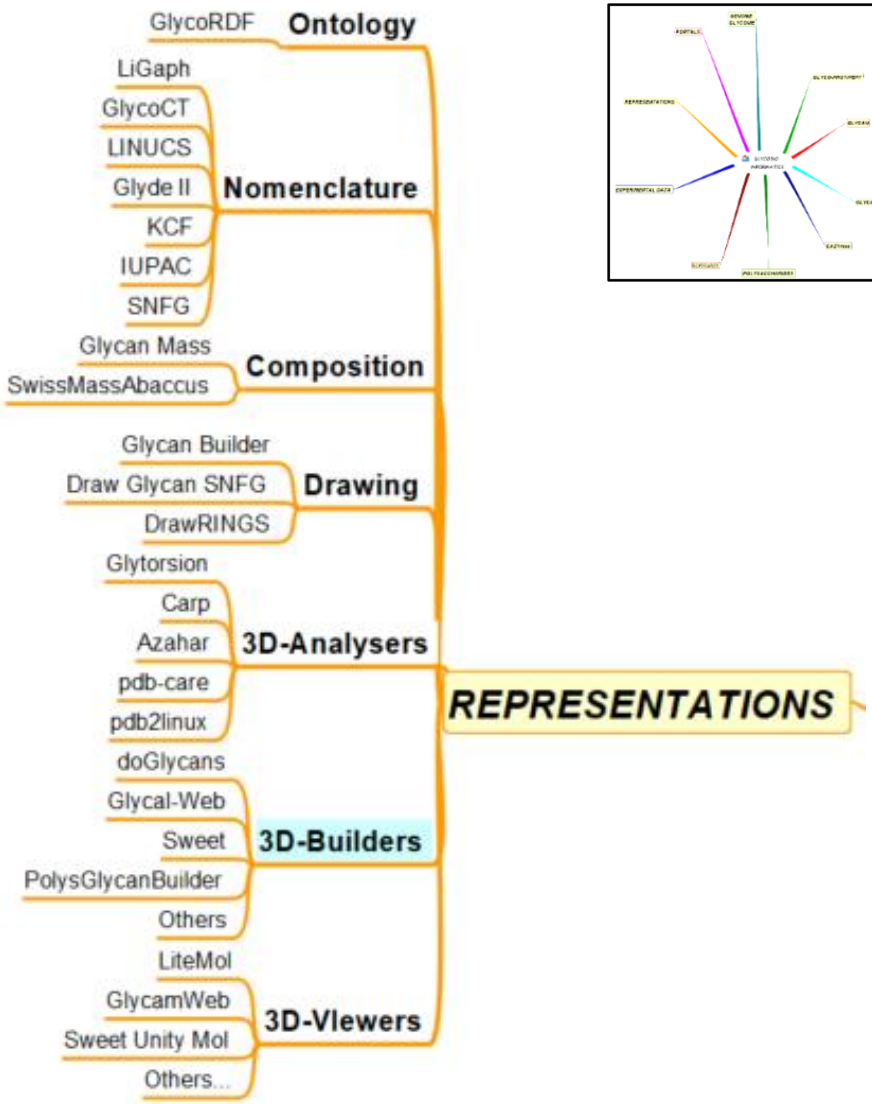
Symbol Nomenclature for Glycans: SNFG



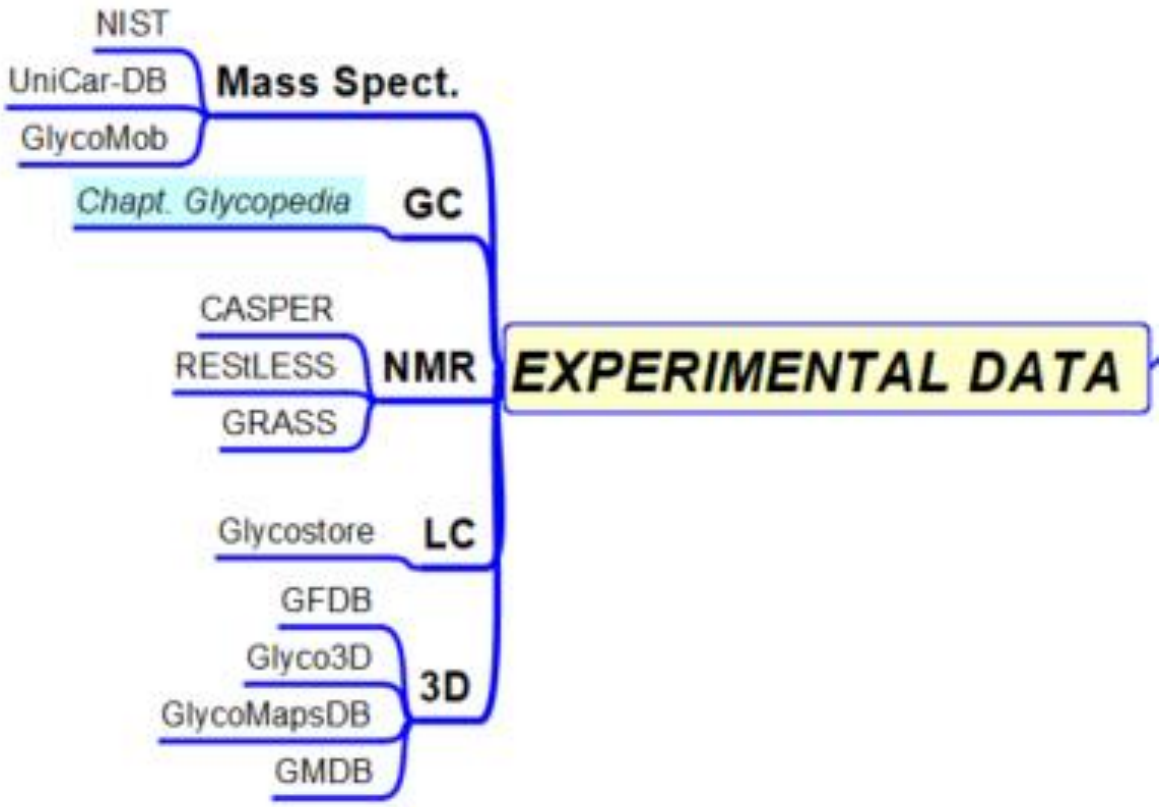


RES
1b:a-dgal-HEX-1:5
2s:n-acetyl
3b:b-dgal-HEX-1:5
4b:a-dgro-dgal-NON-2:6|1:a|2:keto|3:d
5s:n-acetyl
6b:b-dglc-HEX-1:5
7s:n-acetyl
8b:a-lgal-HEX-1:5|6:d
9b:b-dgal-HEX-1:5
LIN
1:1d(2+1)2n
2:1o(3+3)3d
3:3o(3+2)4d
4:4d(5+1)5n
5:1o(6+1)6d
6:6d(2+1)7n
7:6o(3+1)8d
8:6o(4+1)9d

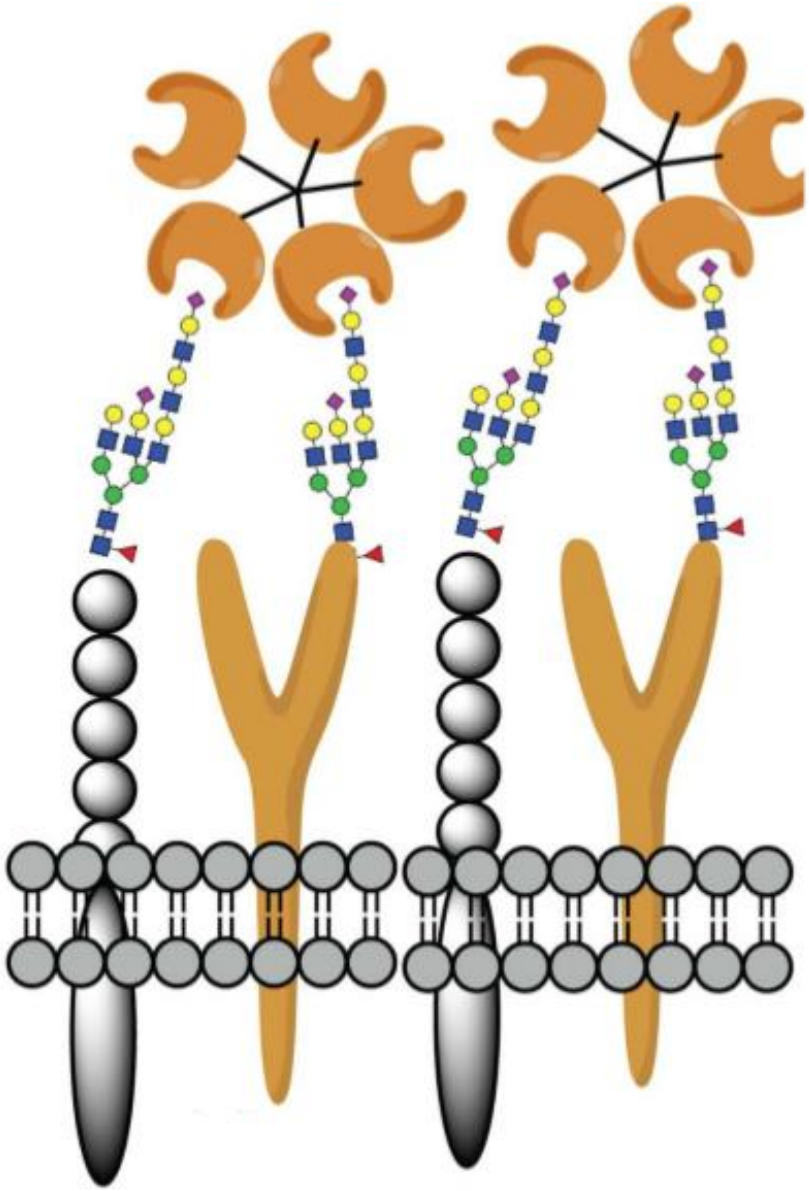




A Traveller-Guide in Glycoscience Cyber-Space

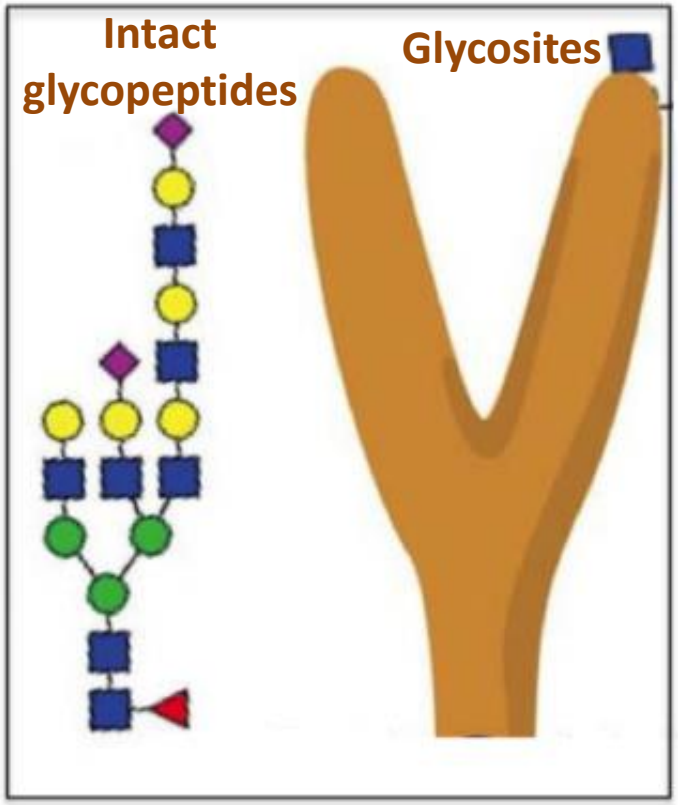


Glycomics

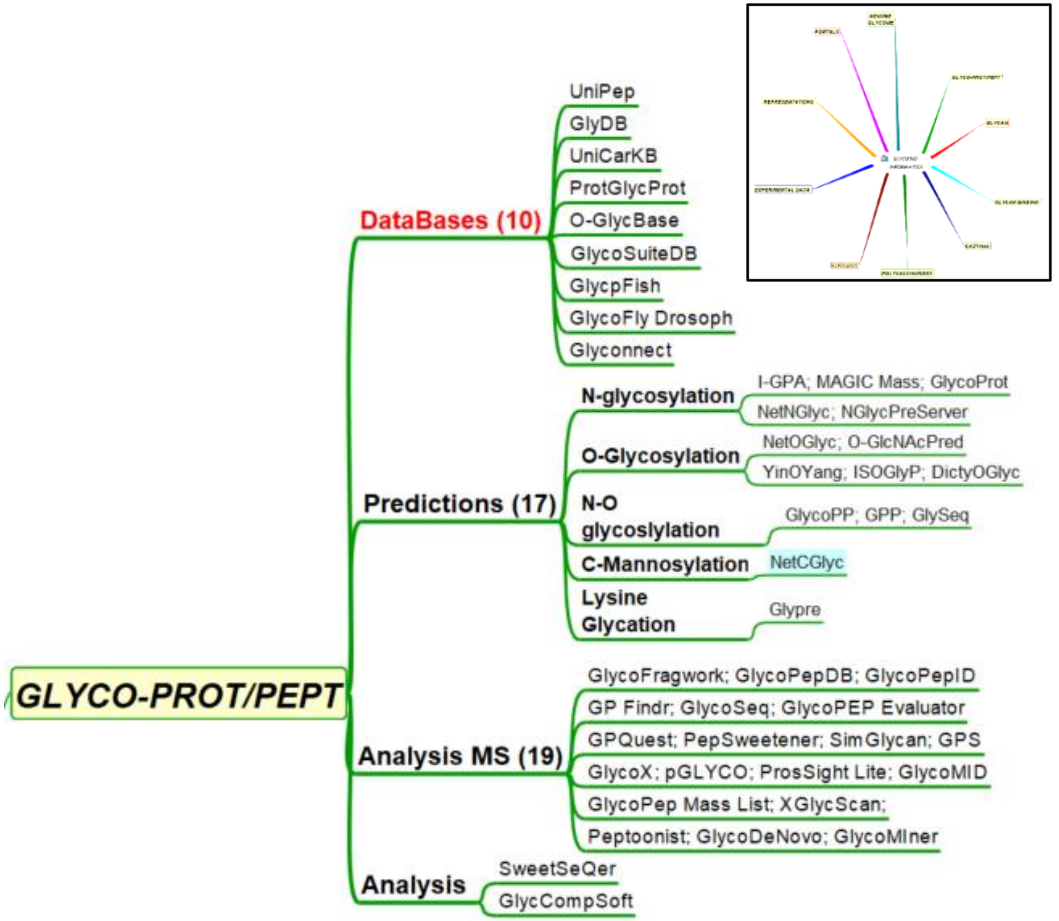


Glycoproteomics

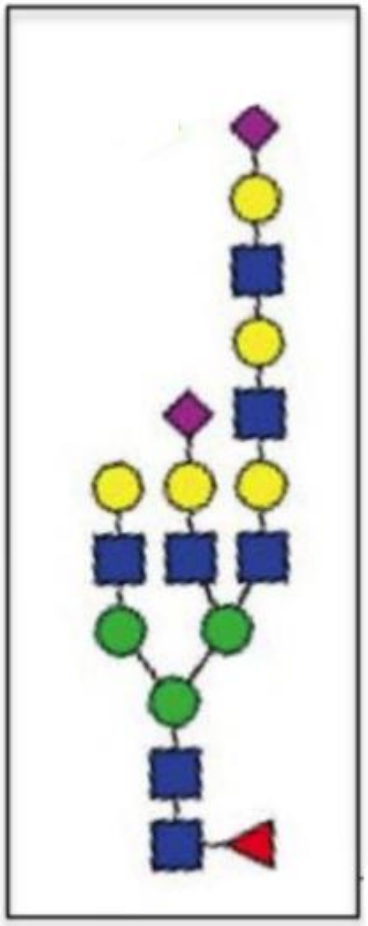
Glycoproteomics

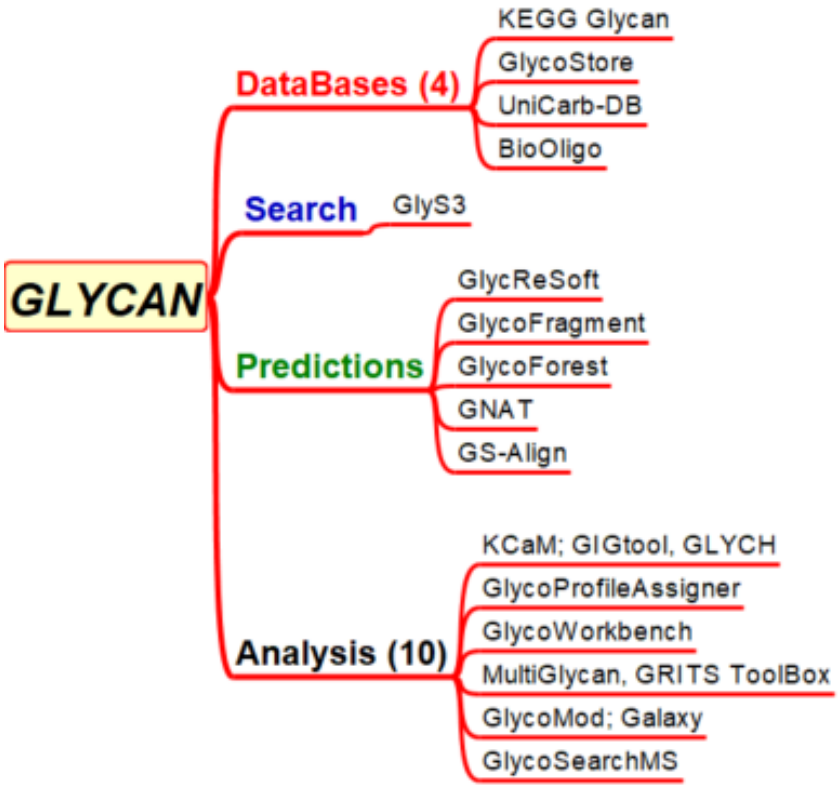


Glycan profiles

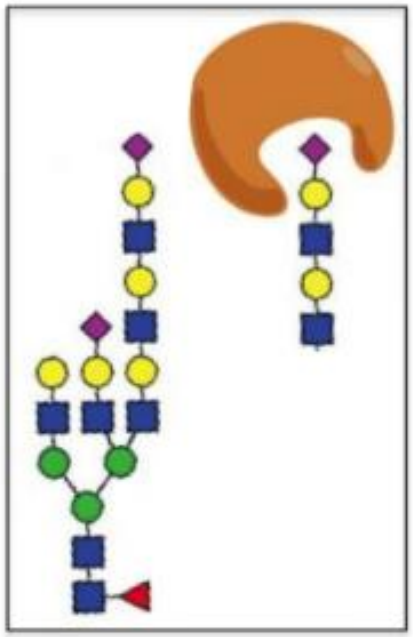


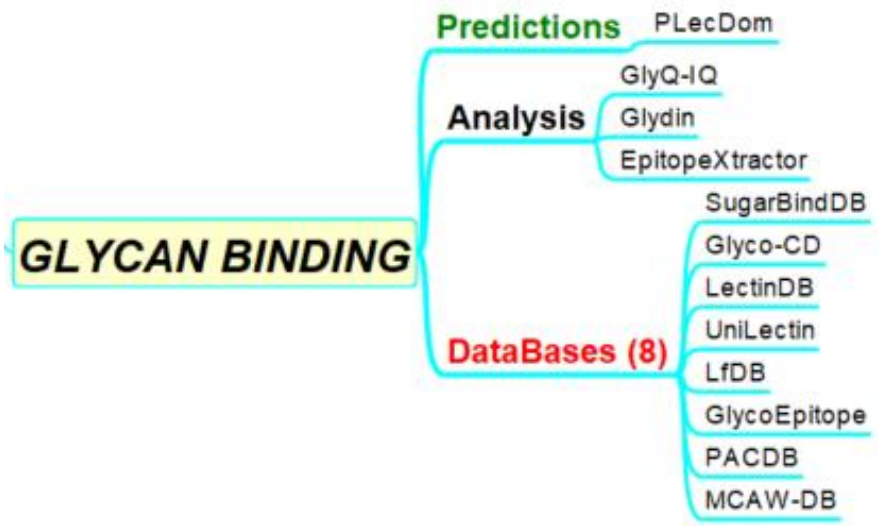
Glycans




GLYCAN
DataBases (4)
KEGG Glycan
GlycoStore
UniCarb-DB
BioOligo
Search
GlyS3
Predictions
GlycReSoft
GlycoFragment
GlycoForest
GNAT
GS-Align
Analysis (10)
KCaM; GIGtool, GLYCH
GlycoProfileAssigner
GlycoWorkbench
MultiGlycan, GRITS ToolBox
GlycoMod; Galaxy
GlycoSearchMS

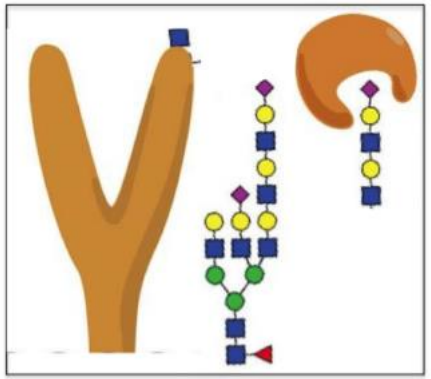
Functional Glycomics





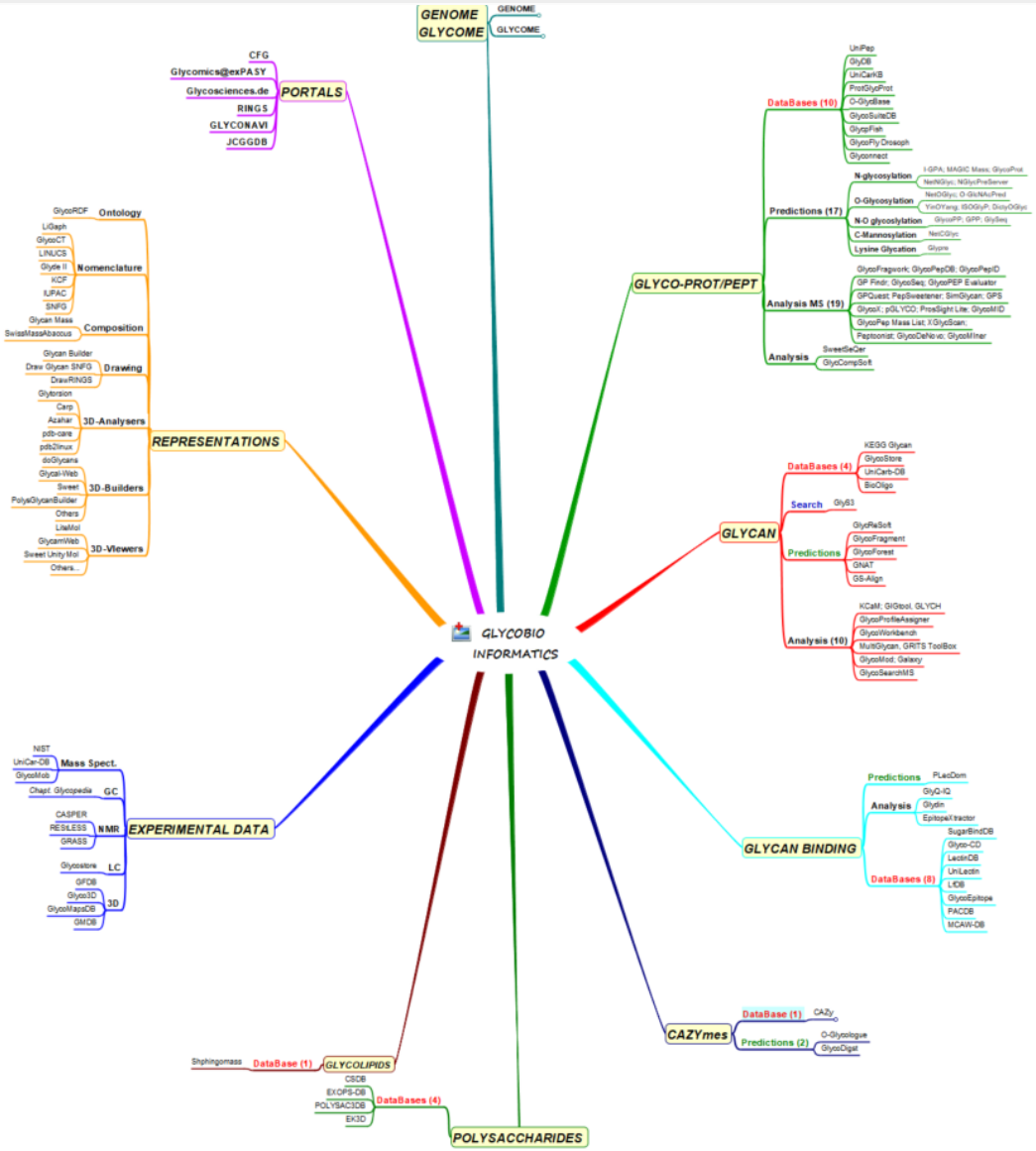
Glycoproteomics

Functional Glycomics



Glyconnect

A Traveller-Guide in Glycoscience Cyber-Space



available soon at : https://glycopedia.eu


Yes
we
Glycan !