Glycans in Pathway Tools

by

Markus Krummenacker

SRI International

Q1 2013

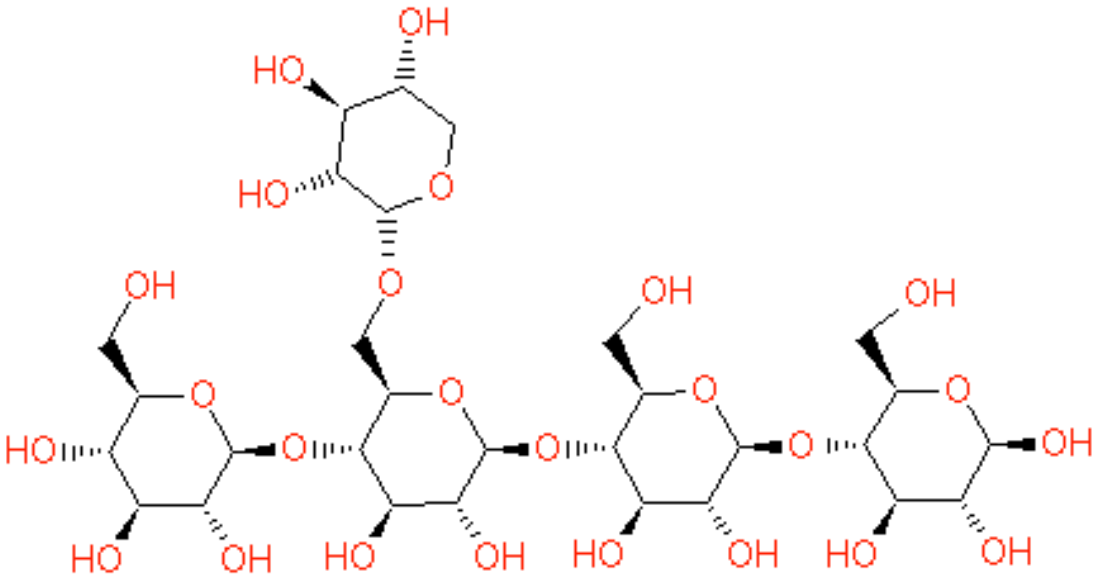
What are Glycans ?

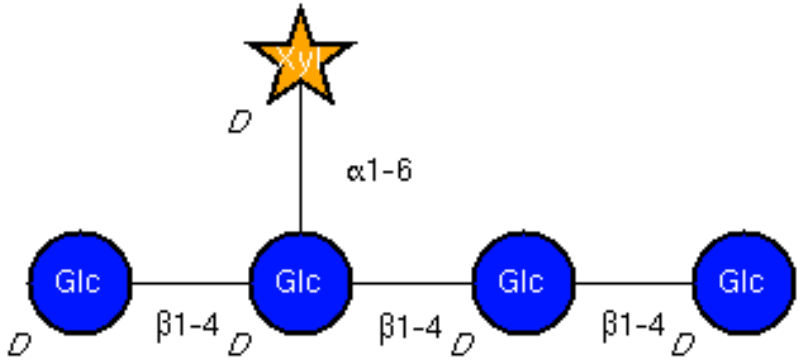
- Macromolecules: Polysaccharides or Oligosaccharides
- Decoration of proteins or lipids
- Structural materials, cell walls: cellulose, chitin
- Storage polymers: glycogen, amylose
- Structures often contain 100s or 1000s of saccharide units
- Important biochemically, and for Biofuel production

Goals for Pathway Tools

- Glycan structures can be very big
- Showing big glycans in atomic detail results in overwhelming clutter
- Instead, icons with colors clarify what the building-blocks are and how they are connected
- Support editing and display of glycans
- For Biofuel production, the goal is to display summarized degradation pathways

Icons clarify glycan structure

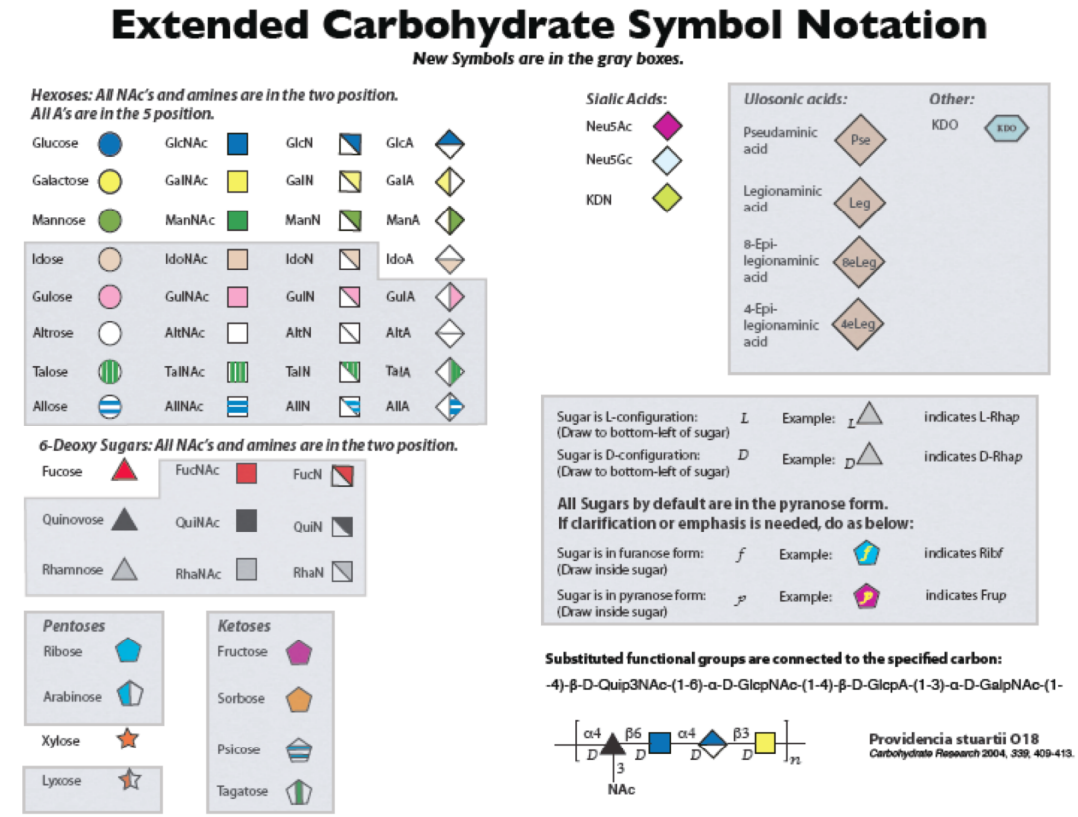




Chosen Representation Standards

- Display:
  - CFG icons (Consortium for Functional Glycomics)
  - Different shapes and colors are combined to represent saccharide units
  - Standard chemical modifications have their own icons

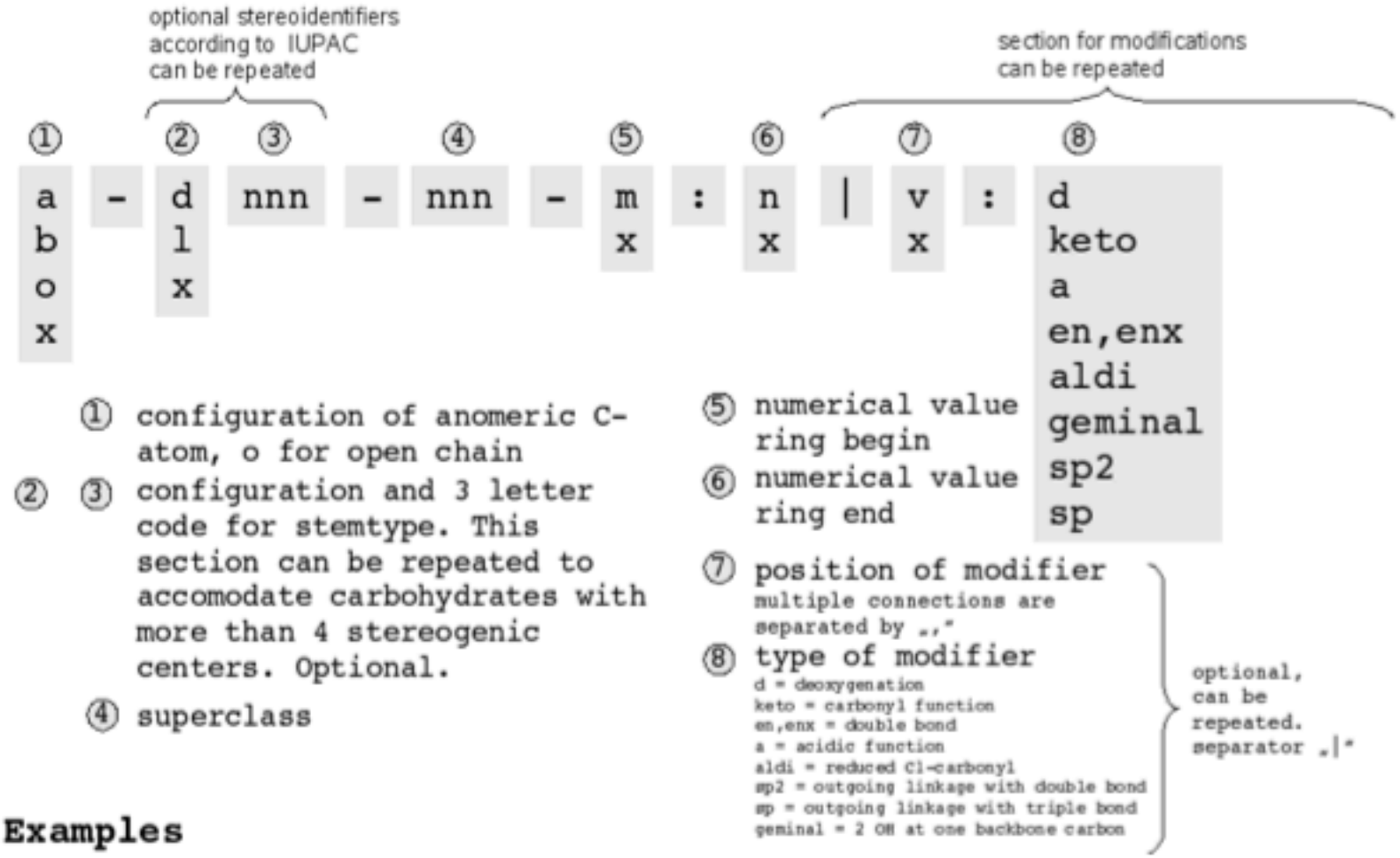
CFG Icons


Extended Carbohydrate Symbol Notation
New Symbols are in the gray boxes.
Hexoses: All NAc's and amines are in the two position.
All A's are in the 5 position.
Glucose GlcNAc GlcN GlcA
Galactose GalNAc GalN GalA
Mannose ManNAc ManN ManA
Idose IdoNAc IdoN IdoA
Gulose GulNAc GulN GulA
Altrose AltNAc AltN AltA
Talose TalNAc TalN TalA
Allose AllNAc AllN AllA
6-Deoxy Sugars: All NAc's and amines are in the two position.
Fucose FucNAc FucN
Quinovose QuiNAc QuiN
Rhamnose RhaNAc RhaN
Pentoses
Ribose
Arabinose
Xylose
Lyxose
Ketoses
Fructose
Sorbose
Psicose
Tagatose
Sialic Acids:
Neu5Ac
Neu5Gc
KDN
Ulosonic acids:
Pseudaminic acid Pse
Legionaminic acid Leg
8-Epi-legionaminic acid 8eLeg
4-Epi-legionaminic acid 4eLeg
Other:
KDO KDO
Sugar is L-configuration: (Draw to bottom-left of sugar) L Example: L indicates L-Rhap
Sugar is D-configuration: (Draw to bottom-left of sugar) D Example: D indicates D-Rhap
All Sugars by default are in the pyranose form.
If clarification or emphasis is needed, do as below:
Sugar is in furanose form: (Draw inside sugar) f Example: indicates Ribf
Sugar is in pyranose form: (Draw inside sugar) p Example: indicates Frup
Substituted functional groups are connected to the specified carbon:
-4)-β-D-Quip3NAc-(1-6)-α-D-GlcpNAc-(1-4)-β-D-GlcpA-(1-3)-α-D-GalpNAc-(1-
α4 β6 α4 β3
D D D D
3
NAc
n
Providencia stuartii O18
Carbohydrate Research 2004, 339, 409-413.
O. Berger et al. Symbol Notation Extension for Pathogen Polysaccharides. July 3, 2008. Consortium For Functional Glycomics, Core D. The Scripps Research Institute.

Chosen Representation Standards

- File exchange:
  - GlycoCT XML
  - GlycoCT has 3 flavors: XML, condensed, compressed
  - Namespaces of all entities are controlled
  - Canonical numbering of residues and linkages
  - Allows a variety of chemical modifications of residues
  - Allows repetitive segments, structural ambiguities

GlycoCT technical details



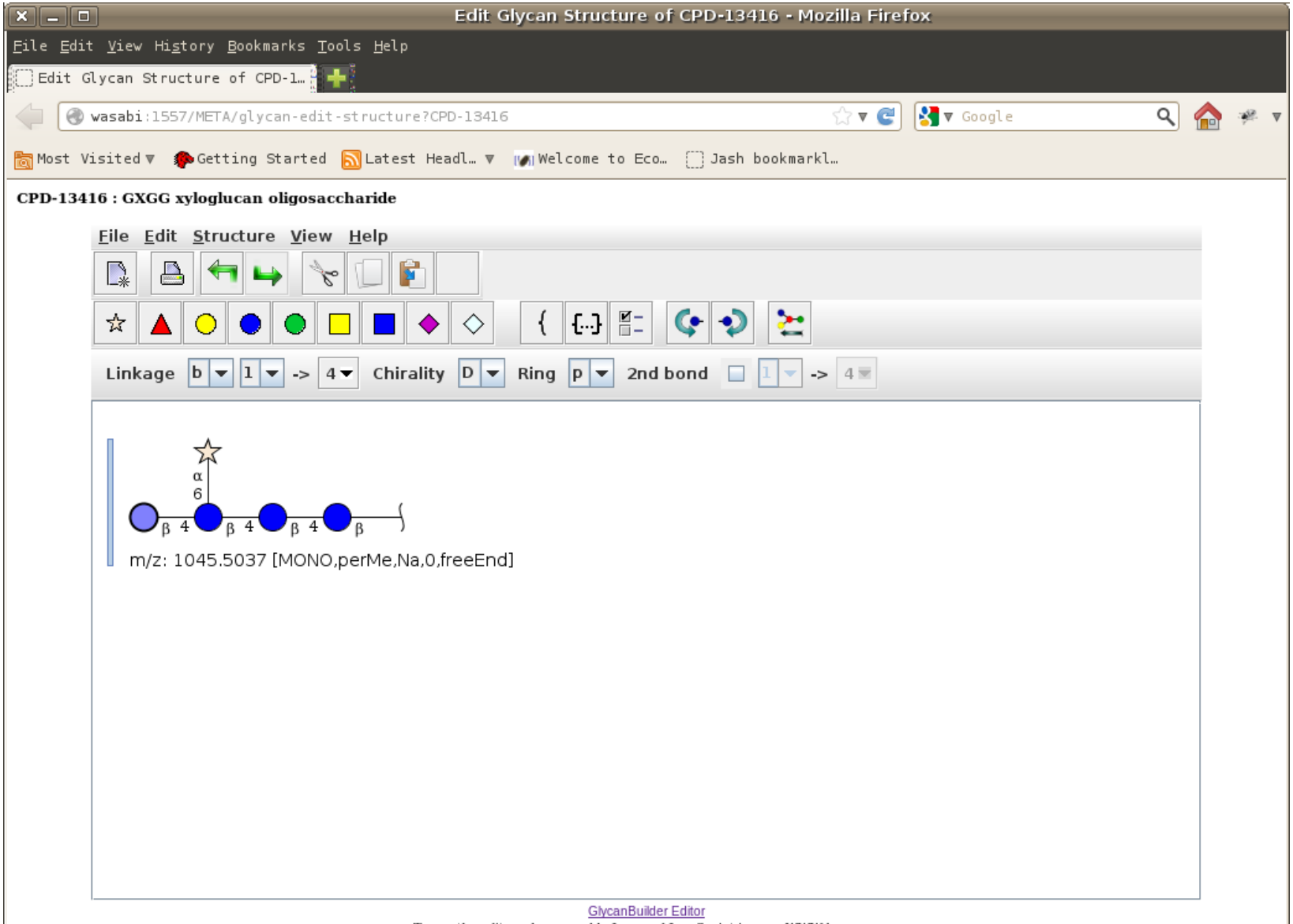
Examples

b-dglc-hex-1:5	β-D-Glucose, pyranose form

GlycanBuilder

- Open source JAVA applet, developed by the EUROCarbDB project
- Uses layout algorithms to draw glycans in standardized ways
- Pathway Tools communicates with GB by starting a local HTTP server, launching a Web browser, and displaying GB in the browser, while exchanging the glycan structure as GlycoCT XML (similar to the Marvin compound editor applet)
- Glycan icon structures are available for compounds under the Glycans class

Editing with GlycanBuilder



Pathway Tools technical details

- Compounds under the class Glycans can have both a traditional atomic structure and a icon structure
- Layout coordinates are sent by GlycanBuilder (special extension)
- Additional slots record icon structure:
  - GROUP-COORDS-2D: (186 82), (134 82), (82 82), (30 82), (82 30)
  - STRUCTURE-GROUPS: |B-DGLC-HEX-1:5|, |B-DGLC-HEX-1:5|, |B-DGLC-HEX-1:5|,|B-DGLC-HEX-1:5|, |A-DXYL-PEN-1:5|
  - STRUCTURE-LINKS: (1 2 O 4 D 1), (2 3 O 4 D 1), (3 4 O 4 D 1), (3 5 O 6 D 1)

Future Work

- Support for non-canonical NON residues. Extend GlycanBuilder to enter free-form text for them.
- Support polymer repeat units
- Implement Glycan (degradation) pathway displays that show at which linkages enzymes act
- Consistency checks between traditional compound structure versus icon structure, if both are present