*Glycobiology*, 2015, vol. 25, no. 12, 1323–1324
doi: 10.1093/glycob/cwv091
Glyco-Forum

Glyco-Forum

# Symbol Nomenclature for Graphical Representations of Glycans

**Ajit Varki[1,2,3], Richard D Cummings[1,2,4,5,6], Markus Aebi[1,2,7], Nicole H Packer[1,2,8], Peter H Seeberger[1,2,9], Jeffrey D Esko[1,3], Pamela Stanley[1,10], Gerald Hart[1,11], Alan Darvill[1,12], Taroh Kinoshita[1,13], James J Prestegard[1,12], Ronald L Schnaar[1,11], Hudson H Freeze[14,15], Jamey D Marth[14,16], Carolyn R Bertozzi[14,17], Marilynn E Etzler[14,18], Martin Frank[2,19,20], Johannes FG Vliegenthart[2,19,21], Thomas Lütteke[2,19,22,23], Serge Perez[2,24,25], Evan Bolton[2,26,27], Pauline Rudd[28,29], James Paulson[30,31], Minoru Kanehisa[32,33], Philip Toukach[2,34,35], Kiyoko F Aoki-Kinoshita[36,37], Anne Dell[38], Hisashi Narimatsu[39], William York[12], Naoyuki Taniguchi[40], and Stuart Kornfeld[41]**

[1]Editor, Third Edition of *Essentials of Glycobiology*, [2]Glycan Symbol Nomenclature Discussion Group, 2014–2015, [3]UC San Diego, La Jolla, CA, USA, [4]Illustrations Editor, *Essentials of Glycobiology*, [5]Chair, Consortium for Functional Glycomics, [6]Harvard Medical School, Boston, MA, USA, [7]ETH Zürich, Zürich, Switzerland, [8]Macquarie University, Sydney, NSW, Australia, [9]Max-Planck-Institute, Potsdam-Golm, Germany, [10]Albert Einstein College of Medicine, New York, NY, USA, [11]Johns Hopkins University, Baltimore, MD, USA, [12]University of Georgia, Athens, GA, USA, [13]Osaka University, Osaka, Japan, [14]Editor, earlier Edition of *Essentials of Glycobiology*, [15]Sanford-Burnham-Prebys Medical Discovery Institute, La Jolla, CA, USA, [16]UC Santa Barbara, Santa Barbara, CA, USA, [17]Stanford University, Stanford, CA, USA, [18]UC Davis, Davis, CA, USA, [19]IUPAC Carbohydrate Nomenclature Committee, [20]Biognos AB, Gothenburg, Sweden, [21]Utrecht University, Utrecht, Netherlands, [22]Curator, MonosaccharideDB, [23]Justus-Liebig-University Giessen, Giessen, Germany, [24]Curator, Glycopedia, [25]CNRS—University Grenoble Alpes, Saint-Martin-d'Hères, France, [26]Lead Scientist, PubChem, National Center for Biotechnology Information, [27]National Library of Medicine, Bethesda, MD, USA, [28]Oxford Glycobiology Institute, [29]National Institute for Bioprocessing, Research and Training, Dublin, Ireland, [30]Past Chair, Consortium for Functional Glycomics, [31]The Scripps Research Institute La Jolla, La Jolla, CA, USA, [32]Principal Investigator, KEGG Database, [33]Kyoto University, Kyoto, Japan, [34]Curator, Carbohydrate Structure Database (CSDB), [35]Zelinsky Institute of Organic Chemistry, Moscow, Russia, [36]Lead Scientist, GlyTouCan Database, [37]Soka University, Tokyo, Japan, [38]Imperial College, London, UK, [39]National Institute of Advanced Industrial Science and Technology, Tsukuba, Japan, [40]RIKEN, Saitama, Japan, and [41]Washington University School of Medicine, St. Louis, MO, USA

Glycans are among the most complex biological molecules found in nature, being composed of monosaccharide units in various combinations and linkages, and featuring diverse and asymmetric types of branching. Until the 1970s, glycan structures were generally represented in scientific publications as chemically accurate drawings, or using standard IUPAC-approved abbreviations (Sharon 1986; McNaught 1997; Chester 1998). However, given the complexity of many glycan structures, it became difficult to represent such molecules in publication figures, especially those involving multistep biosynthetic pathways or multiple structures. Thus, various investigators began to use symbols to represent monosaccharides in glycans. In 1978, Kornfeld and colleagues put forward an elegant and simple system for representation of vertebrate glycans (Kornfeld et al. 1978), and it entered into popular use over the next two decades. This system was eventually adopted and standardized by the editors of the first edition of the textbook *Essentials of Glycobiology* (Varki et al. 1999).

While this adoption increased the popularity of the symbols, the system had some limitations and inconsistencies and did not take advantage of color. Anticipating a revision, the editors developed an updated symbol nomenclature, which was made available to the community in 2005 before the publication of the Second Edition (Varki, Cummings, Esko, Freeze, Stanley, Bertozzi, et al. 2009). This nomenclature was adopted and widely disseminated, especially by the NIGMS-funded Consortium for Functional Glycomics (CFG), which also consulted the scientific community on this issue, thus enhancing acceptance. The official version of the Second Edition nomenclature has since been available in the online version of the textbook. A suggested extended version was then presented at the website Glycopedia. Meanwhile, alternate symbol nomenclature systems have been proposed (Harvey et al. 2009). Issues regarding nomenclature tend to be more controversial than the underlying science (Varki, Cummings, Esko, Freeze, Stanley, Marth, et al. 2009), as there is never one correct answer, and some aspects are matters of taste. For example, the two systems mentioned each have some relative advantages and disadvantages.

Anticipating preparation of another revision of the *Essentials* textbook, a newly reorganized group of editors has agreed on an updated monosaccharide symbol nomenclature for depicting glycans in the Third Edition. The goal was to produce a system that goes beyond vertebrate glycans, while also considering input from others using related systems, coordinating with the IUPAC Carbohydrate Nomenclature committee, linking each monosaccharide symbol to the corresponding entry in PubChem at NCBI/NLM and initiating coordination with other long-term online resources such as UniCarbKB, MonosaccharideDB, KEGG Pathways, CSDB, JCGGDB, ACGG-DB and GlyTouCan. The system is now being presented as an online advance Appendix to the Third Edition of *Essentials* to be published at NCBI, and by Cold Spring Harbor Laboratory Press. As with the previous version, there is no attempt to require researchers to use the proposed system, unless they voluntarily wish to do adopt it, ahead of the formal publication of the Third Edition.

To ensure compatibility with many prior publications that have already used the Second Edition symbol set, no changes were made to that system, but new symbols have been added. For this historical reason, shapes and colors are internally consistent only for Hexoses, Hexosamines, HexNAcs, Hexuronates and Pentoses, and only shapes are consistent for Nonulosonates, Deoxyhexoses, DeoxyhexNAcs and Dideoxyhexoses (e.g. L-Fucose is not assigned the most logical color). It was felt useful to designate a generic monosaccharide class using an all-white symbol, to aid mass spectrometric depictions wherein only the mass is known (e.g. a hexose of unknown type would be a white circle). As mentioned previously, D and L configurations, pyranose versus furanose and ring conformations are not represented, but are assumed to be either implicit in their trivial name, or the most common ones found in nature (less common configurations need to be indicated in the figure or stated in a figure legend). Abbreviations for modifications (e.g. *O*-acetyl and sulfate esters) will also follow the style of the Second Edition, with additions as needed. Monosaccharides not represented in the current set can be shown with a pentagon symbol and a letter inside, and defined in the figure legend. Symbol colors are specified in CMYK settings currently used by printers, and assignable in standard drawing programs. Linkages can be shown as in the Second Edition system using IUPAC style, with the originating carbon assumed, and hyphens (not commas) used as an option. Examples of depictions of common glycans using the new system are shown. There is now also the option of using the Oxford method, which depicts the monosaccharide linkages with embedded specificity and anomericity (Harvey et al. 2009). This Third Edition system will be considered as a foundation for a symbolic nomenclature to be included in a future update of the IUPAC Recommendations for Carbohydrate Nomenclature. Further details on all of the above aspects can be found in the book Appendix at NCBI. The online and pdf versions of this article also contains many other useful links.

In addition to a substantially expanded number of assignments of colored symbols to specific monosaccharides, a special feature of the new system is that the symbols at the website will eventually have embedded in them the following information: hovering a tooltip/pointer over a symbol will reveal the full name of the monosaccharide, and a click on a symbol will link the reader to the corresponding PubChem page, with full details about the monosaccharide. For user convenience, symbol sets are also available in downloadable files as high quality SVG objects, and a table of text word objects, the latter file including symbol-embedded links to PubChem. The symbols may also be eventually featured at the UniCarbKB database, MonosaccharideDB (a site that lists 700+ natural and synthetic monosaccharides found in current databases), in the CSDB, within an updated version at Glycopedia, and possibly at other sites such as the KEGG pathways database, GlyTouCan, and ACGG-DB.

We hope that this extended graphical representation of glycans will cope better with the rapidly growing information on the structure and functions of glycans in chemical and biological systems. In particular, the symbol nomenclature might help students and researchers to more easily discover and appreciate the relevance and beauty of glycan diversity in living systems, and to communicate this exciting information to others.

## References


Chester MA. 1998. IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN). Nomenclature of glycolipids—Recommendations 1997. *Eur J Biochem*. 257:293–298.

Harvey DJ, Merry AH, Royle L, Campbell MP, Dwek RA, Rudd PM. 2009. Proposal for a standard system for drawing structural diagrams of N- and O-linked carbohydrates and related compounds. *Proteomics*. 9:3796–3801.

Kornfeld S, Li E, Tabas I. 1978. The synthesis of complex-type oligosaccharides. II. Characterization of the processing intermediates in the synthesis of the complex oligosaccharide units of the vesicular stomatitis virus G protein. *J Biol Chem*. 253:7771–7778.

McNaught AD. 1997. International Union of Pure and Applied Chemistry and International Union of Biochemistry and Molecular Biology. Joint Commission on Biochemical Nomenclature. Nomenclature of carbohydrates. *Carbohydr Res*. 297:1–92.

Sharon N. 1986. IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN). Nomenclature of glycoproteins, glycopeptides and peptidoglycans. Recommendations 1985. *Eur J Biochem*. 159:1–6.

Varki A, Cummings RD, Esko JD, Freeze H, Hart GW, Marth JD. 1999. *Essentials of Glycobiology*. First Edition. Plainview (NY): Cold Spring Harbor Laboratory Press.

Varki A, Cummings RD, Esko JD, Freeze HH, Stanley P, Bertozzi CR, Hart GW, Etzler ME. 2009. *Essentials of Glycobiology*. Second Edition. Plainview (NY): Cold Spring Harbor Laboratory Press.

Varki A, Cummings RD, Esko JD, Freeze HH, Stanley P, Marth JD, Bertozzi CR, Hart GW, Etzler ME. 2009. Symbol nomenclature for glycan representation. *Proteomics*. 9:5398–5399.