

Characterization of carbohydrates and related products

26 August 2024

All materials deriving from new and/or commercial polysaccharides must be characterized with a sufficient level of detail to allow the work to be reproduced. Please read this document carefully.

1. Essential information for commercially available polymers

- Supplier and catalogue code.
- Purity, e.g., the total carbohydrate content in the sample and the percentage of each monosaccharide residue.
- Molecular weight (please see below the requirements on molecular weight determination).
- For alginates: as a minimum, the fraction of G residues (FG) or equivalently the M/G ratio must be provided. If the more detailed structure is expected to influence the study, the full composition based on NMR must be reported, for example, as a table.
- For pectins: degree of methyl esterification and content of galacturonic acid residues.
- For chitosan: degree of deacetylation.
- For xanthan gum: degree of pyruvylation and acetylation.
- For red seaweed galactans: ratio galactose:3,6-anhydrogalactose:sulfate.
- For glycosaminoglycans or other sulfated glycans: degree of sulfation.
- For heteropolysaccharides: monosaccharide composition.
- For nanocellulose (cellulose nanocrystals, cellulose nano- and microfibrers): how prepared; both lateral and longitudinal dimensions (from AFM or TEM, hydrodynamic radius from DLS data is insufficient); surface charge type and density (or number of charged groups, zeta-potential data is insufficient) and counter ions; composition for nano- and microfibrillated cellulose - cellulose, hemicelluloses and lignin (if any).
- For nanochitin: how prepared; both lateral and longitudinal dimensions (from AFM or TEM, hydrodynamic radius from DLS data is insufficient), surface charge density (or degree of surface deacetylation)
- For cellulose derivatives (e.g. esters, ethers) nature and DS (and MS where pertinent, such as for hydroxyalkyl ethers) of substituents; position of substitution is often pertinent and useful information as well
- For native starch: include source (botanical source, supplier catalogue number), amylose content, purity (non-starch components) and extraction method (if extracted in lab). If the molecular weight/size of amylose and/or amylopectin affects the results, these parameters, including the method of determination, are required.
- For modified starches (chemical, enzymatic and physical): modification process details, degree of modification (or substitution) and molecular/supramolecular structural details, including the methods used, are required.

2. Molecular weight

When reporting molecular weight, the method used must be sufficiently specified, and the following issues should be considered.

- Report molecular weights with units (g/mol or Da).

- The type of molecular weight average must be correctly reported unless essentially uniform (monodisperse) samples are studied. Different methods give different average types. Reporting molecular weight intervals (e.g. 100–300 kDa) is not acceptable.
- For light scattering or combination of light scattering with GPC/SEC the dn/dc must be indicated.
- When using SEC/GPC with standards different from the polymer studied, large systematic errors can be introduced. The values are thus apparent molecular weights (M_{app}). This can be accepted but needs to be clearly demonstrated (in figures and tables). Example: Molecular weights are apparent values obtained from SEC/GPC calibrated with xxx standards (specifying the type of columns used for separation, flow rate and distribution angle used for the quantification as well as the name of the commercial standard used).
- Solution viscosities can be reported, providing concentration, solvent, temperature, and shear rates.
- If providing molecular weight calculated from intrinsic viscosity, all experimental conditions (temperature, solvent) and Mark-Houwink-Sakurada (MHS) constants should be provided with the proper citation. Please note that molecular weights obtained from intrinsic viscosity measurements via the MHS equation are usually not viscosity averages. See e.g. <https://doi.org/10.1002/app.1989.070370822>
- If only a single (average) molecular weight is used, the choice must be justified. In many systems the molecular weight may play a significant role. If molecular weights are not considered or believed not to influence the results/conclusions this must be justified and properly discussed. The same goes for the degree of substitution.
- The term ‘polydispersity index’ (and its abbreviation PDI) was repealed by the IUPAC and should not be used any longer, especially not in scientific communication. The correct term is ‘Dispersity’, abbreviated \mathcal{D} . (Pure Appl. Chem., Vol. 81, No. 2, pp. 351–353, 2009. doi:10.1351/PAC-REC-08-05-02).

3. For polysaccharides, either natural or produced by synthetic procedures, whose structure has been previously reported, or for new polymers derived from polysaccharides with known structure, information comprises:

- Purity.
- Molecular weight.
- Sugars and linkage composition.
- Substitution degree and, where possible, location of the new groups.

4. For polysaccharides whose structure was previously unknown, information comprises:

- Purity.
- Molecular weight.
- Monosaccharide composition inclusive of absolute configuration in case the glycan is from a bacterial or seaweed source. The experimental section should specify whether the ratio between the peaks made use of a calibration or if it was area-based.
- Linkage analysis, preferably with deuterium labeling at the anomeric position and/or C6 in the case of uronic-acid containing polysaccharides. EI-MS spectra of the derivatives must be provided for reviewing purposes as Supplementary Material.
- NMR characterization:
 - The minimal information required is the attribution of the cross-peaks/densities of the ^1H - ^1H COSY, ^1H - ^1H TOCSY, and ^1H - ^{13}C HSQC (or HMQC) NMR spectra to individuate the anomeric configuration (α or β) and ring size (furanose or pyranose) of the monosaccharide residues. In case the polysaccharide is heterogeneous due to the presence of non-stoichiometric substituents, the

assignment can be limited to the most intense signals. The nature and the location of the non-stoichiometric substituents should be defined.

- Reporting the sequence between the monosaccharide residues is not mandatory but considered beneficial for the quality of the manuscript, and it should be established by ^1H - ^{13}C HMBC and/or ^1H - ^1H NOESY (or ROESY) spectra.
- The inclusion of other spectra (as ^1H - ^{13}C HSQCTOCSY, H2BC, ^{31}P -based NMR spectra for phosphorous containing samples) that can countercheck the deductions from the other spectra, is desirable.
- Rules for NMR data presentation are reported in the “Tables and Figures” section.
- If present, nature of non-carbohydrate substituents like (but not limited to) acyl, alkyl, sulfate and phosphate groups.
- To assess the novelty of a newly isolated polysaccharide from bacterial or plant source, the authors should consult the public repository available at <http://csdb.glycoscience.ru/database>.

5. Order-disorder transitions and sample history

Polysaccharides like xanthan and scleroglucan/schizophyllan can exist in different solvent- and temperature-dependent conformations. They often show hysteresis in heating-cooling cycles and may be mixtures of multistrands, single strands and complex intermediate forms. Therefore, the sample history and treatments need to be specified.

6. Use of certain techniques:

The following techniques/approaches may not be accepted to define the fine structural features (monosaccharide residues, ring size, anomeric configuration) of new and/or commercial polysaccharides:

- Microscopy (SEM, TEM or AFM) images. Of note, these measurements are suitable to define the features of materials derived from the glycans and/or the interaction with suitable molecules used as probes, as well as dimensions of nanocellulose.
- Standard FT-IR measurements, since spectra of different oligo/polysaccharides usually do not have sufficient relevant differences. Of note, characterization by this technique is accepted for chemically modified polysaccharides to qualitatively demonstrate the derivatization of the sample. For FT-IR to be used quantitatively, an appropriate calibration should be performed and reported in the manuscript.
- Standard UV measurements of isolated oligo/polysaccharides, since these molecules do not have sufficient diagnostic UV absorbing moieties.
- Hydrodynamic radius from dynamic light scattering (DLS) is insufficient in defining the dimensions of nanocellulose. Zeta-potential cannot be solely relied to determine the charge amount of the nanocellulose.

7. Methods and units

Methods already published should be indicated by a reference, only relevant modifications should be described. Yet, the text must report a minimum set of information to enable the reader to understand in what the method consisted of.

Example of an acceptable reference to a previously published method: “Monosaccharide composition was performed by acid hydrolysis with TFA 2 M for 2 h at 120 °C and derivatization to the acetylated alditols”.

Example of an unacceptable reference to a previously published method: “Monosaccharide composition was performed as reported by (citation)”. This way to refer to a previous method is not appropriate because it forces the readership to back to the reference to understand the approach used.

The routine chromatographic conditions (equipment, column, temperature ramps, carrier gas or eluant, detector, etc.) should be provided by adding a reference, and the chromatographic profiles shown in Supplementary Material.

For studies reporting the molecular modeling of glycans and/or their docking with enzymes/receptors, the coordinates (see protein-data-bank, PDB, or other accessible formats) of the molecules discussed in the manuscript, and proposed as figures, should be provided as Supplementary Data.

Mechanical properties of materials should contain the information on sample dimensions and temperature and humidity conditions.

Please follow internationally accepted rules and conventions: use the international system of units (SI). If other units are mentioned, please give their equivalent in SI.

Absolute configuration convention:

- descriptors D or L used to define the absolute configuration of monosaccharides should be two points smaller in size than the rest of the text or typed in SMALL CAPS.

Anomeric configuration convention:

- the stereochemistry of the anomeric carbon should be indicated as α or β . These letters are not written in italic and must be done with the “symbol” font of the MS Word.

NMR data convention:

- number of digits in NMR spectra: proton chemical shifts must be reported with two decimals while carbon chemical shifts with only one.
- Chemical shift can be expressed in δ or ppm, never both, and the same convention should be used throughout the manuscript
- The first column of the chemical shift Table must indicate the letter associated with a certain residue together with its anomeric configuration, its mode of substitution and the location of non-carbohydrate substituents (if any). The chemical shift of the non-carbohydrate substituents can be indicated in separate lines of the table, in the heading or as a note at the end of the table.

Linkage analysis convention:

- data should be reported preferably in a Table stating the methylated and acetylated positions of the alditol (or PMAA) and/or the deduced linkage, their proportions with integer numbers if expressed as %, or up to 2 decimal figures if expressed as ratios, specifying in the heading of the column if the numbers are area- or mole-based ratios, and the fragments diagnostic of the derivative (note that the EI-MS spectra must be reported as Supplementary Material).
- The abbreviation to be used for the deduced linkage of the PMAA derivative should omit the anomeric position: 2-linked glucopyranose should be written as 2-Glcp (or 2-Glc but not 1,2-Glcp), 3-linked glucosamine as 3-GlcN (or 3-GlcpN). The “p” of pyranose can be omitted since most of the residues are in this form, while the furanose form, if deduced from the linkage analysis or by other approaches, must be indicated. In cases where the linkage analysis does not define a unique ring size, like 1,3,4,5-tetra-O-acetyl-2-O-methylarabinitol whose deduced linkage is compatible with both 3,5-Araf or 3,4-Arap, and if the correct ring size is not deduced by other methods, the table should list both possibilities.

8. Depicting polysaccharides

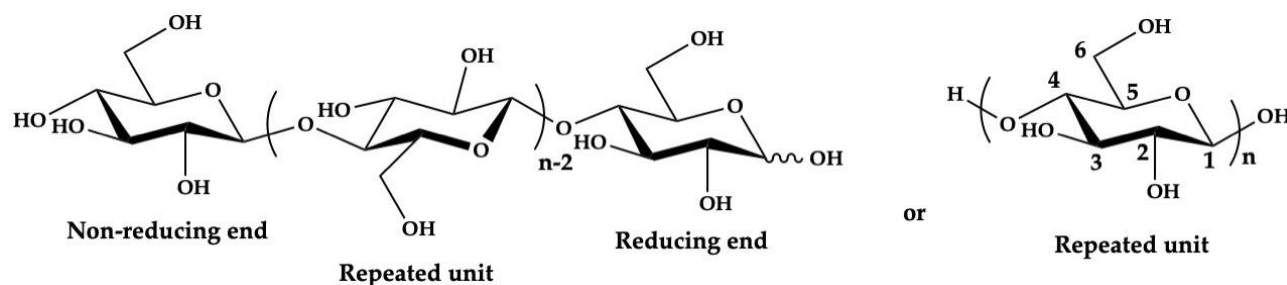
Depicting polysaccharides should follow certain rules; it is beyond the scope of this document to depict them all. Chair depictions are often useful for display both structure and stereochemistry.

A comprehensive library of monosaccharide residues is freely available at:

<https://glycopedia.eu/resources/library-of-monosaccharide/article/presentation>

One important rule that is often ignored by authors is that the OH at the terminal anomeric position should be outside of the right repeat unit bracket; the linkage oxygen of the repeat unit (e.g. O-4 in cellulose, below) should be inside of the left bracket, while the terminal H should be outside the left bracket. This reflects the fact that the linkage oxygen (O-4 for cellulose) is considered part of the repeated unit.

Cellulose



Adapted from French, A. D. (2017). Glucose, not cellobiose, is the repeating unit of cellulose and why that is important. *Cellulose*, **24**, 4605–4609. DOI 10.1007/s10570-017-1450-3.

Please note also that a newer, pictorial-based system has gained increasing popularity in recent years for depicting natural polysaccharides (Varki, A. et al., *Essentials of Glycobiology* **2017**, 3rd edition, Plainview NY, Cold Spring Harbor University Press, ISBN 978-1-621821-32-8). This pictorial form is now widely adopted by the carbohydrate community, it conveys a lot of information in a relatively simple form and is recommended for use in Carbohydrate Polymers manuscripts. Keep in mind that it may be more complicated to use for properly depicting certain semisynthetic polysaccharide derivatives.

Symbols and the rules to be used are available at

<https://www.ncbi.nlm.nih.gov/glycans/snfg.html#Examples>