Carbohydrates – Characterization Overview

Methods for Structure and Property Deduction

- Chemical methods for functional group identification
- Chemical methods for ring size identification
- Methylation analysis
- X-ray crystallography
- Analytical methods for characterization
  - UV-Vis spectroscopy
  - Circular dichroism spectroscopy
  - NMR spectroscopy
  - Mass spectrometry
Monosaccharides

Mutarotation

Optical activity

\[
[\alpha]_D = +112.2^\circ \\
[\alpha]_D = +18.7^\circ
\]

At equilibrium .... \([\alpha]_D = +52.7^\circ\) (Not an average!
But weighted average; 63.6% of the \(\beta\)-anomer)
Simple Tests for Reducing / Non-reducing Sugars

- **Importance of Reducing Sugars**
  - Case of hemoglobin (Hb)
    - Hb is glycated ..... HbA1c levels should be < 6.5%; Reaction of high glucose with NH₂ group at the N-terminus (Val) of Hb; Non-enzymatic; Reflects Glc levels for the past 3 months (typical life of erythrocytes)
  - Can serve as sugar modifying handle
    - Label with chromogenic groups or fluorophores
    - Preparation of sugar-labeled solid matrix

- **Tollen’s Reagent**

  ![Diagram of Tollen’s Reagent]

  D-Glucose $\xrightarrow{\text{Mild Oxidizing Agent}}$ D-Gluconic Acid

  Mild Oxidizing Agent (e.g., Tollen’s reagent) + Metallic Silver

  \[
  \text{D-Glucose} \rightarrow \text{D-Gluconic Acid} \quad \text{(Ag⁺ / NH₃ – H₂O)}
  \]
Simple Tests for Reducing / Non-reducing Sugars

- **Benedict’s reagent**

\[
RCHO + 2Cu^{2+} + 5OH^- \rightarrow RCOO^- + Cu_2O + 3H_2O
\]

- Benedict's reagent is a solution of the citrate complex of CuSO₄ in water. Cu²⁺ is a weak oxidizing agent.
- A positive test is the formation of a red precipitate of Cu₂O.

- **Sugars that exhibit positive Tollen’s / Benedict’s tests**
  - Aldoses
  - Ketoses, if they exhibit good keto-enol tautomerism
  - Glycosides do not show positive test because they cannot exhibit aldehyde ↔ equilibrium
UV-Vis Spectroscopy of Carbohydrates

- Glycans are UV-Vis transparent; groups include –OH, –CHO, –C=O, –OCH₃, –COOH, etc.
- Some glycans have a conjugated double bond; e.g., heparin disaccharide. Such groups absorb at 232 nm, which allows identification and quantitation of oligosaccharides.

- Have to be typically labeled for visualization
- Use reducing end labels such as 2-aminopyridine (2-AP), 2-aminoacridone (AMAC), 7-amino-1,3-naphthalenedisulphonic acid (AGA), p-aminobenzoic acid (PABA), 2-aminobenzamide (2ABA), 4-aminobenzonitrile, 8-aminonaphthalene-1,3,6-trisulfonic acid (ANTs), 1-phenyl-3-methyl-5-pyrazolone (PMP), 1-(4-methoxy)phenyl-3-methyl-5-pyrazolone (PMPMP), 6-aminoquinoline (6-AQ), 1-maltohepaosyl-1,5-diaminonaphthalene, 8-aminopyrene-1,3,6-trisulfonate (APTS), etc.
UV-Vis Spectroscopy of Carbohydrates

Labeling reaction ... typically Schiff base formation followed by reduction to a stable covalent bond

\[
\begin{align*}
\text{Glycosyl donor} & \quad + \quad \text{2-AP} \\
& \quad \xrightarrow{\text{AcOH, } \Delta} \\
& \quad \text{Reduced Schiff base}
\end{align*}
\]
NMR Spectroscopy of Carbohydrates

- Every polysaccharide has a unique 1D NMR spectrum, which contains all of the information about the structure.

- Additional NMR experiments are performed to assign the resonances, e.g., $^1$H, $^1$H-COSY and $^1$H, $^1$H-TOCSY for $^1$H resonances or $^1$H, $^{13}$C-HETCOR or HSQC for $^{13}$C resonances.

- Sequence determination is performed through experiments such as NOESY ($^1$H, $^1$H through space correlations) or HMBC ($^3$J$_{COCH}$).

- Methods for interpretation of $^1$H NMR spectrum can save much time and effort.
Basic Approach for NMR Structure Determination

1D $^1$H-NMR
+ 1D $^{13}$C-NMR

2D $^1$H,$^1$H-COSY

Basic Approach for NMR Structure Determination

2D $^1$H, $^1$H-NOESY

2D $^1$H, $^{13}$C-HETCOR
Basic Principle

- Through-bond correlation
- Through-space correlation
- Anomeric proton

3J couplings may needed for conformational assignment.
Requires knowledge of saccharide residues.
Additional experiments needed for larger structures.

Fig. 7. The $^{1}C_4$ conformation of $\alpha$-l-galactopyranosyluronic acid residue and a probable global conformation for 4 based on literature reports of the preferred conformations$^{7,8}$. 
NMR Tools on the Internet
(Taken from a lecture on ‘Internet tools for the Interpretation of NMR spectra’
by
Dr. Roland Stenutz (roland@organ.su.se), University of Stockholm, Sweden)

Polysaccharide structure:
Components | ”type” | Hex or HexNAc
relative configuration | Glc or Man
absolute configuration | D- or L-
ring size | -p or -f
Linkages | position | →4) or →6)
stereochemistry | α- or β-
Sequence | | →4)Glc(→4)Gal(→
or
| | →4)Gal(→4)Glc(→

NMR can be used to perform each of these steps, except for the
determination of the absolute configuration
Most saccharides are typically present in one form in nature, which serves to identify the configuration
Internet-based Approaches to NMR Analysis

- comparison with a database (SugaBase)
  simple and accurate but limited to known structures or sub-structures.

- comparison with simulated NMR spectra (CASPER)
  requires information about the components and linkages to limit the number of possible structures.

- Just a few examples of tools available in polysaccharide analysis

Taken from ‘Internet tools for the Interpretation of NMR spectra’ by Dr. Roland Stenutz (roland@organ.su.se), University of Stockholm, Sweden)
SugaBase

http://www.boc.chem.uu.nl/sugabase/databases.html

Carbohydrate Structure

Help: [Search level] [Strict linkages] [Nomenclature]

Search level: Full Residue
Strict linkage check: □

Carbohydrate Structure:
b-L-Fucp-(1-3)-a-D-Glcp-(1-1)-Methyl

C#: P-0201-A00404
CC: CC5P:AC0404
MHz 75
Temp 333
Solv D2O
Original Reference: Acetone
Reference Value: 31.45
Correction Applied: -0.31

<table>
<thead>
<tr>
<th>Residue</th>
<th>Linkage</th>
<th>Carbon PPM</th>
<th>J</th>
<th>Hz</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>a-D-Glcp</td>
<td></td>
<td>C-1 99.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>C-2 70.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>C-3 83.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>C-4 69.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>C-5 72.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>C-6 61.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>C-1 104.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>C-2 71.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>C-3 73.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>C-4 72.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>C-5 71.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>C-6 16.1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Taken from ‘Internet tools for the Interpretation of NMR spectra’ by Dr. Roland Stenutz (roland@organ.su.se), University of Stockholm, Sweden)
Sweet-DB

- http://glycosciences.de hosts several tools and databases for the carbohydrate chemist

- Sweet-DB is an interface to CarbBank/SugaBase

Taken from ‘Internet tools for the Interpretation of NMR spectra’ by Dr. Roland Stenutz (roland@organ.su.se), University of Stockholm, Sweden)
Searching With NMR Data on Sugabase

Example 1: Proton Search
Example 2: Carbon Search

Taken from ‘Internet tools for the Interpretation of NMR spectra’ by Dr. Roland Stenutz (roland@organ.su.se), University of Stockholm, Sweden)
Taken from ‘Internet tools for the Interpretation of NMR spectra’ by Dr. Roland Stenutz (roland@organ.su.se), University of Stockholm, Sweden)
Best Search Result

Searched for nmr information. Results: 1 - 10 of 174

Structure matches query with 95.00%

\[
\begin{align*}
\text{b-L-Fucp-(1-3)} & + \\
\quad | \\
\text{b-D-Glcp-(1-1) -Methyl} & + \\
\quad | \\
\text{a-D-Manp-(1-2)} & +
\end{align*}
\]

Taken from ‘Internet tools for the Interpretation of NMR spectra’ by Dr. Roland Stenutz (roland@organ.su.se), University of Stockholm, Sweden)
CASPER
(http://www.casper.organ.su.se)

- **Help using CASPER**
- **Sequence determination**
  Determine the sequence of a poly- or oligo-saccharide from NMR chemical shifts and the results of methylation and sugar analysis.
  - Example 1 - α(1→6)glucan
  - Example 2 - a Shigella LPS
- **Structure simulation**
  Build a structure, simulate its 1H- and 13C-NMR spectra, and, optionally assign experimental spectra.
  - Example 1 - Assignment of the 13C-spectrum of methyl b-cellobioside
- **Literature** List of references.
- **Guide to methylation analysis**
# CASPER

<table>
<thead>
<tr>
<th>Title</th>
<th>b-L-Fucp-(1-3)-α-D-Glcp-(1-1)-Methyl</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Residue</th>
<th>Linkage</th>
<th>‘Reducing’ end</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>D-GlcpOMe</td>
<td>not linked</td>
</tr>
<tr>
<td>b</td>
<td>L-Fucp</td>
<td>(1→3)</td>
</tr>
<tr>
<td>a</td>
<td>none</td>
<td>(1→2) residue 1</td>
</tr>
<tr>
<td>a</td>
<td>none</td>
<td>(1→2) residue 1</td>
</tr>
<tr>
<td>a</td>
<td>none</td>
<td>(1→2) residue 1</td>
</tr>
<tr>
<td>a</td>
<td>none</td>
<td>(1→2) residue 1</td>
</tr>
</tbody>
</table>

### 13C-Chemical shifts

<table>
<thead>
<tr>
<th>99.7</th>
<th>70.8</th>
<th>83.5</th>
<th>69.7</th>
<th>72.2</th>
<th>61.3</th>
</tr>
</thead>
<tbody>
<tr>
<td>104.0</td>
<td>71.8</td>
<td>73.7</td>
<td>72.1</td>
<td>71.8</td>
<td>16.1</td>
</tr>
</tbody>
</table>

Load shifts from [ ] Bläddra...
Correct by subtracting [ ] ppm

---

Taken from ‘Internet tools for the Interpretation of NMR spectra’ by Dr. Roland Stenutz (roland@organ.su.se), University of Stockholm, Sweden)
Calculated shifts

Assignment

Experimental shifts

Taken from ‘Internet tools for the Interpretation of NMR spectra’ by Dr. Roland Stenutz (roland@organ.su.se), University of Stockholm, Sweden)
CASPER also allows incorporation of glycosylation and substituent shifts, positional shifts.

Sequences can also be derived using CASPER.

Taken from ‘Internet tools for the Interpretation of NMR spectra’ by Dr. Roland Stenutz (roland@organ.su.se), University of Stockholm, Sweden)
Mass Spectrometry in the Structure Determination of Glycans

- The premier structure determination tool today
- Highly sensitivity (µg quantities)
- Highly information rich ... full sequence with detailed stereochemical structure possible
- Widely applicable ... N- & O-glycans, sialylated glycans, GAGs, etc.
- May be used in tandem with chemical or enzymatic pre-treatment
- Various MS technologies are available for ionization
  - MALDI (Matrix assisted laser desorption ionization)
  - ESI or nESI (Electrospray or nanoElectrospray ionization)
  - FAB (Fast atom bombardment)
- Various approaches are available for structure elucidation
  - HR TOF (high resolution time of flight)
  - MS/MS or MS/MS^n (tandem MS)
  - Ion Mobility
A Strategy for MS Sequencing of N-Glycans

Fig. 1. N-Glycan analysis strategy. The entries in rounded rectangles represent the flow and chemical state of the analyte. The entries in parallelograms represent chemical or enzymatic treatment steps. The entries in hexagons represent instrumental analysis methods. * If more than 500 µg of glycoprotein are available, methylation analysis (panel C) can be performed in addition to MS/MS (panel B), in order to confirm the linkages obtained by sequential exoglycosidase digestions (panel A). In this case, use only three fifths for exoglycosidase treatments, otherwise use four fifths. * The N-glycans are subjected to sequential digestions with various exoglycosidases. After each digestion one fifth of the initial sample amount (i.e. the portion that was taken for exoglycosidase treatments) is separated, permethylated, and analyzed by MALDI-TOF mass spectrometry.

Exoglycosidases in Sequencing of Glycans

Figure 1. Exoglycosidase enzymes commonly used to determine the structure of N-linked glycans by sequential degradation.

Permethylation/Peracetylation and Linkage Analysis

- Permethylation advantages include easier determination of linkages and branching
- Allows simultaneous analysis of sialic acid and neutral glycans
- Better ionization in MS chamber
- Makes analysis possible with pM quantities of sample

Chemical reactions:

\[ \text{CH}_3\text{I, dry DMSO} \quad \text{CH}_2\text{Cl}_2 \quad 2\text{M TFA, } >100 \, ^\circ\text{C} \quad 1\text{M NH}_4\text{OH, NaBD}_4 \quad \text{Ac}_2\text{O, base, heat} \]
The Mass Spectrometer

- Molecules can be imparted charge (+ve or −ve)
- Charged molecules can be made to move (accelerate)
- A moving object (e.g., a molecule) can be made to deflect
- Deflection is related to mass
The Tandem Mass Spectrometer
Nomenclature of Product Ions

- Fragment ions containing the non-reducing terminus are labeled A, B, C
- Fragment ions containing the reducing terminus are labeled X, Y and Z
- Subscripts 1, 2, 3, …. refer to the number of residue from either the reducing or the non-reducing terminus
- Subscript 0 refers to cleavage of the bond with the aglycon
- Fragment labels A and X are reserved for intra-ring cleavage
- Superscripts refer to intra-ring bonds cleaved (clockwise numbering)
- Products ions can be \([M+nH]^{n+}, [M-nH]^{n-}, [M+nNa+mH]^{(n+m)+}\),

**FIGURE 1.** Nomenclature for glycoconjugate product ions generated by tandem MS (Modified from Domon & Costello, 1988b).
Formation of Product Ions

FIGURE 2. Fragmentation of (a) protonated and (b) alkali-cationized glycosidic bonds (Modified from Cancilla et al., 1996).
Formation of Product Ions

**FIGURE 3.** Three possible fragmentation pathways for metal-cationized oligosaccharides. Reprinted with permission of Cancilla et al. (1999). Copyright 1999 American Chemical Society.

Formation of Product Ions

Glycan Characterization Using SimGlycan™

Fig. 19.1. Glycan Prediction Schema for SimGlycan™: MS/MS data collection mechanism and uploading it into SimGlycan™ along with search parameters.

Fig. 19.2. Glycan Prediction Schema for SimGlycan™: MS/MS database searching, scoring and ranking of the probable glycans using SimGlycan™.

Taken from Apte and Meitei, Methods Mol. Biol. 2010, 600, 269-281.