BMB8130 Glycobiology, 2011

Course Organizers: Lance Wells and Michael Tiemeyer

BMB8130 Glycobiology, 2011 Administrative Organization

-- Grade
- 33% Attendance and Participation in Lectures
- 33% EXAM (Learning the Language)
- 33% Clinical Correlation Session
- 1% Instructors’ Discretion

-- Clinical Correlation
- Paper presentations relevant to lecture content; topical papers of basic biochemical or translational biomedical interest
- Teams of 2 – 3 students
- 2 teams per session
- Highlight alternative views or contradictory findings

BMB8130 Glycobiology, 2011 Thematic Content

Bootcamp
- Glycan Structures and Synthesis (Learning the Language)
  - EXAM, Feb. 8.

Compelling Topics
- Glycoanalytics
- Glycans and Cellular Function
- Glycan Macromolecules
- Glycans and Development
- Glycan-based Pathophysiology

Course Content

Double lecture, starting at 8:30 AM

EXAM!

Georgia Glycoscience Symposium
For full program, see: http://www.ccrc.uga.edu/symposium/program.pdf
What is Glycobiology?

A term frequently attributed to Raymond Dwek, et al (circa 1990) to encompass the body of research that contributes to understanding:

The structure, biosynthesis, and biology of saccharides

Or, more explicitly stated (at least more explicit than could have been imagined 30 years ago) -- Glycobiology is the human field of endeavor that strives to characterize and understand:

- The diversity of glycan structures
- The processes by which glycans are synthesized
- The determinants of glycan structure
- The mechanisms of glycan-protein interactions
- The impact of glycans on the structure and function of the molecules to which they are attached
- The contribution of glycans to normal cellular function and tissue development
- The participation of glycans in diverse pathologies

Thus, aspects of glycobiology impact a broad range of disciplines

To name a few --
- organic synthetic chemistry
- protein biochemistry
- enzymology
- analytic chemistry
- structural biochemistry
- cell biology
- developmental biology
- genetics
- genomics
- proteomics
- parasitology
- neurobiology
- reproductive medicine
- endocrinology
- cell signaling
- stem cell biology
- membrane biophysics
- microbiology
- cancer biology
- immunology
- microbiology
- biotechnology

And, in fact, glycobiology impacts life, itself, from conception (sperm-egg interactions) to death (apoptosis, multiple systemic pathologies)
How do glycans impact your work?

Depends on who you are:

– Protein biochemist: Pesky modifications that impart heterogeneity to my otherwise pure protein.
– Structural biochemist: Nasty modifications that make it hard to crystallize some really interesting proteins.
– Synthetic chemist: You want to synthesize what? And you want how much?
– Geneticist: A family of regulated molecules whose expression is not simply related to the activity of a single gene. Painful pleiotropy.
– Molecular Biologist: What’s the big deal—I’ve been doing glycobiology for my whole career? They make a nice backbone structure.

How, then, does Glycobiology fit within the context of modern molecular, genetic, structural, and systems biology?

The Central Dogma, circa 1970

DNA ← RNA → Protein

Organism ← Cell

Rather ignores a role for lipids and carbohydrates, especially at the cell surface

An expanded Dogma

DNA ← RNA → Protein → enzymes

Saccharides and lipids

Modified transcription factors

Organism ← Cell ← glycoconjugates
Predominant cellular distribution of glycans implies role for carbohydrate in the societal interactions of cells in tissues

Electron micrograph of a human lymphocyte (Ruthenium Red staining)

After Yarik, A

Carbohydrates - Basic Terms

- Monosaccharide
  - A carbohydrate that can not be broken down into smaller carbohydrates by treatment with acids- a simple sugar
- Oligosaccharide
  - Approximately 4-12 mono units
- Polysaccharide
  - Usually greater than 12 mono units
  - Often a long linear repeating chain consisting of a single monosaccharide type with or without small side chains

Glycoconjugates

- Glycoproteins
  - Protein + Carbohydrate
  - Most proteins of biological and therapeutic significance
  - e.g. TPA, EPO, LH, Immunoglobulins
- Glycolipids
  - Lipid + Carbohydrate
  - Important tissue and cell type markers
  - e.g. blood group antigens, tumor associated antigens
- Proteoglycans
  - Complex of a protein and one or more polysaccharides known as glycosaminoglycans (GAGs)
  - Immense structural diversity, pleiotropic functions

The Building Blocks (in CFG code)

Symbolic Representations of Common Monosaccharides and Linkages

Galactose (Gal)  Xylose (Xyl)
N-Acetylglactosamine (GalNAc)  N-Acetylgalactosaminic acid (Neu5Ac)
Galactosamine (GalN)  N-Glycolyneuraminic acid (Neu5Gc)
Glucose (Glc)  2-Keto-3-deoxyogalactonic acid (Kdo)
N-Acetylglucosamine (GlcNAc)  Fucose (Fuc)
Glucosamine (GlcN)  Glucuronic acid (GluA)
Mannose (Man)  Isuronic acid (IdoA)
N-Acetylmannosamine (ManNAc)  Galacturonic acid (GalA)
Mannosamine (ManN)  Mannuronic acid (ManA)

Other Monosaccharides

Use letter designation inside symbol to specify if needed
Potential for high information content is inherent in glycan structure

<table>
<thead>
<tr>
<th>Macromolecule</th>
<th>Building Block</th>
<th>Approximate Mass</th>
<th>Possible Variations in a Trimer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protein</td>
<td>Amino acids</td>
<td>$125 \rightarrow 10^{1.10}$</td>
<td>6</td>
</tr>
<tr>
<td>Nucleic Acid</td>
<td>Nucleotides</td>
<td>$330 \rightarrow 10^{-1.10}$</td>
<td>6</td>
</tr>
<tr>
<td>Carbohydrate</td>
<td>Monosaccharides</td>
<td>$209 \rightarrow 10^{-1.10}$</td>
<td>1,056 to 27,648!</td>
</tr>
</tbody>
</table>

Carbohydrate Synthesis - Secondary gene products

- Oligosaccharide structures are not encoded in the DNA
- Correct synthesis dependent on
  - Glycosyltransferases
  - Glycosidases
  - Availability of monosaccharide and activated forms
  - Organelle distribution and function
- Consequence
  - Glycosylation is very sensitive to subtle shifts in the environment
**Glycoconjugates- Basic Terms**

- **Microheterogeneity**
  - The variation seen in glycosylation at a given glycosylation site. A site may be unoccupied or may be occupied by different sugars. This fact led early researchers to falsely conclude that carbohydrates were not important.

- **Glycoforms**
  - A glycoconjugate may have different glycoforms. That is, the non-carbohydrate portion will remain the same, but variances in the carbohydrate portion will create different glycoforms of a glycoconjugate.

- **Glycone vs. Aglycone**
  - Carbohydrates (glycones) should be considered in the context of what they are linked to (aglycones).

---

**Essentials of Glycobiology**

**Thy-1 Glycoprotein**

- Predicted polypeptide MW = 14 kD
- Observed MW by SDS-PAGE = 25-26 kD
- Glycan accounts for almost 50% of total glycoprotein mass.
Three great waves of Glycobiologic expansion

**The first wave:** The long, grand ascendancy of analytic chemistry, with a little functional analysis in the shadows.
- How are these building blocks put together and who can publish the weirdest structure in JBC this month?
- By the way, there’s some cool function over here.

**The second wave:** Brave, new pioneers bring molecular tools and the promise of deciphering function.
- You mean that genes may actually be relevant to glycans?
- And molecular biology can do what?

Three great waves of Glycobiologic expansion (cont’d)

**The third wave (cresting now):** The great integration, with an eye increasingly focused on developing therapeutics.
- First concerted attempts to generate therapeutics reveal that we don’t know enough to do rational design.
- Leads to expansion of efforts in glycomics to facilitate the identification of relevant glycans.
- Initial applications of genetics to identify glycan function (targeted and accidental).
- Driving forces increasingly come from outside of traditional Glycobiology domains (the help me phone call).
Reasonably well-supported and accepted functions for glycans

- Structural support for cells (extracellular matrix, plant cell wall)
- Signaling molecules (currently only characterized in plants)
- Determinants of protein folding, stability, and half-life
- Mediators of cellular interactions with:
  - Other cells in developing tissues
  - Specific cell types in mature tissues, especially related to immune response and tissue stability
  - Viral, bacterial, parasitic pathogens and associated toxins
- Modulators of signaling activities that drive cellular differentiation

The thrill and the agony of Glycobiology:

Biological roles of oligosaccharides: all of the theories are correct.

Varki A.
Glycobiology Program, UCSD Cancer Center.

Many different theories have been advanced concerning the biological roles of the oligosaccharide units of individual classes of glycoconjugates. Analysis of the evidence indicates that while all of these theories are correct, exceptions to each can also be found. The biological roles of oligosaccharides appear to span the spectrum from those that are trivial, to those that are crucial for the development, growth, function or survival of an organism. Some general principles emerge.

And this is where we begin!
Carbohydrates - Definition and Language

Carbohydrates = “Hydrates” of carbon.

Hexose, e.g.: C\textsubscript{6}H\textsubscript{12}O\textsubscript{6} = (CH\textsubscript{2}O)\textsubscript{6}.

Need to develop a system for talking about and/or representing carbohydrates.

Monosaccharides: single sugars; clear language and numerous pictorial forms.

Oligosaccharides (typically 4-10 sugars): need a more complex language, only one of previous pictorial forms remains tractable.

Polysaccharides: systematic language is accurate but cumbersome, new pictorial representation more useful.

Carbonyl containing compounds: Aldehydes RCHO and Ketones RCOR.

Glyceraldehyde and “Fischer Projections”

Glyceraldehyde (aldotriose), a 3 carbon aldehyde sugar or “aldotriose,” exists as 2 mirror image isomers.

\[ \text{d-glyceraldehyde} \quad \text{l-glyceraldehyde} \]

The origin of D vs. L nomenclature for sugars - does stereocenter farthest from the aldehyde terminus have the configuration of D- or L-glyceraldehyde?

Dihydroxyacetone (ketotriose): No chiral carbon.

\[ \text{Dihydroxyacetone (ketotriose)} \]

All monosaccharides with one exception have at least one chiral carbon with the total number (k) being equal to the number of internal (CHOH) groups; that is k = 2 for Aldoses and k = 3 for Ketoses with n = number of carbon atoms in the monosaccharide.

Possible Stereoisomers = 2 raised to the power of k (how many for hexose?)
1/18/11

**D-”Aldoses” with 4 Carbons**

Two possible isomers at each new carbon center. Mentally insert new carbon center between aldehyde terminus (C1) and what was previously C2. Note that D configuration is retained. Two sugars that differ only in the configuration around a single chiral carbon are called **EPIMERS**.

**D-”Aldoses” with 5 Carbons**

Which 5 carbon sugars are epimers of D-ribose?

**D-”Aldoses” with 6 Carbons**

**special attention to Gal, Man, and Glc abbreviations**

6 carbon ketose: fructose

**Chapter 2, Figure 3**

**Essentials of Glycobiology**

Second Edition

16 Hexoses
A Few of the Common non-(CH₂O)n Building Blocks of Oligosaccharides
Products of esterification, oxidation, reduction, acetylation, etc.

- lactones
- deoxy
- Amino
- N-acetyl
- GlcA & GalNAc
- sialic acids
- Sia

Fischer Structures Does Not Take Into Account that 5 and 6 Carbon Sugars Tend to Cyclize:
RING CONFIRMATION

Cyclization Can Produce Multiple Isomers

Fischer Projection is NOT good enough
From Fischer to Haworth to abbreviated Haworth diagrams.
Looking down the carbon spine from C1 down of the Fischer projection, noting whether hydroxyl groups are to the right (down) or left (up). Now imagine around the periphery of a flat hexagon.
“A new asymmetric carbon” is formed so we need more language (great!)
New asymmetric carbon is termed the anomeric carbon and $\alpha$ is used to denote the anomer where the absolute stereochemistry of the anomeric position and the most remote sterocenter in the sugar chain are the same, $\beta$ opposite

For Hexoses $\alpha$ =anti (anomeric hydroxyl to C6); $\beta$ =beside

Where did this pyranose ($p$) term come from?

Pyranose ($p$) and Furanose ($f$)
both 5 and 6 membered rings are stable so:

A Last Bit of Nomenclature for Chair Structures (the most accurate way of drawing the sugars) Haworth suggests that sugars are flat but neither furanose or pyranose rings are actually flat in their lowest energy confirmations

Although $D$-glucose has a strong preference for one chair conformation, this is not true for all sugars.
Chapter 2, Figure 8

**The Glycosidic Bond**

How do we describe molecules containing sugars that are attached to one another? (We’ll limit our discussion to cases where the anomeric center of one sugar is attached to an oxygen atom of another sugar - that is, we’ll discuss only “glycosides.”)

There are basically 2 things we need to keep track of: 1) the anomeric configuration of the “glycosidic linkage,” and 2) the identity of the carbon on the next sugar that shares the bridging oxygen.

**Mutatrotation**

In solution anomers can introconvert

<table>
<thead>
<tr>
<th>% Distribution at 20°C</th>
<th>Pyranose</th>
<th>Furanose</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>α</td>
<td>β</td>
</tr>
<tr>
<td>D-Glc</td>
<td>34</td>
<td>65</td>
</tr>
<tr>
<td>D-Gal</td>
<td>32</td>
<td>64</td>
</tr>
<tr>
<td>D-Man</td>
<td>66</td>
<td>34</td>
</tr>
<tr>
<td>D-Fru</td>
<td>4</td>
<td>68</td>
</tr>
</tbody>
</table>

**Naming of Oligosaccharides**

Here’s a simple example: maltose, or D-Glc-α(1-4)-D-Glc.

How did we come up with that name, how do we know which sugar to name 1st?

We always begin naming with the sugar furthest from the “reducing terminus” of the oligosaccharide. The “reducing sugar” is the sugar that still contains a free anomeric carbon. Branches are put in parenthesis. Notice the α or β is now no longer “free” but is fixed in one orientation.

Many “common name” disaccharides (sucrose, lactose, etc.) see Ess. Glycobiol.

Note: we even keep this nomenclature when the oligosaccharide is attached to an aglycone.
Example of Naming a Branched Oligosaccharide

Pictorial system is becoming “the standard” representation (especially for complex structures)