

Chapter 7

Carbohydrates and glycobiology

Problems: 3, 4, 7, 8, 10, 12, 15, 16

Added homework available for extra credit

7.0 Introduction

I think proteins are important because they do most of the cells chemistry
On the other hand, if you go by sheer mass, carbohydrate are the most important because they are the most abundant biomolecule on earth. Each year roughly 100 million tons of CO₂ and water is converted into carbohydrate, primarily cellulose in plants

the carbohydrate sugar and starch are dietary staples in most of the worlds

oxidation of carbohydrate main energy source for most cells

insoluble carbohydrates polymers main structural component of bacterial and plant cell walls and animal connective tissue

carbohydrate polymers used to lubricate joints and connect animal tissues

carbohydrates attached to lipid or protein determine intracellular location or metabolic fate of compounds

Basics

carbohydrates

Empirical formula CH₂O (but can have N P or S as added substituents)

primarily cyclized polyhydroxyl aldehydes or ketones

3 major classes

Monosaccharides

Oligosaccharides

Polysaccharides

Mono - simple sugars - single aldehyde or ketone unit

Most abundant 6C D- glucose or dextrose

Oligo - short chains of monos

Most abundant - disaccharide - 2 units

In cells if more than 2, often attached to lipid or protein to make **Glycoconjugate**

poly- generally more than 20 monomer units

Can be up to thousands

Can be linear or branched

7.1 Mono- and Di-saccharides

already said polyhydroxylated aldehyde or ketone
each OH general makes C a chiral center so many stereo isomers

A. Monosaccharides

two families: Aldoses and ketoses
colorless crystalline solid
soluble water, insoluble nonpolar
sweet taste
backbone is unbranched C chain (that may cyclize)
If end C is aldehyde then aldose
if interior C is ketone then ketose
some simple 3, 5 and 6 C sugars shown **figure 7-1**
3C triose
4C tetrose
5C pentose
6C hexose
7C heptose

B. Chirality

all but dihydroxyacetone have asymmetric centers
so all but DHA are optically active
simplest is glyceraldehyde
As saw a while back, by convention have called one D and the other L (**figure 7-2**)

Have different trivial names for each isomer except isomer for asymmetric C MOST DISTANT from C=O
(Last asymmetric center on bottom)
If OH here is same side on the fisher projection of D-glyceraldehyde,
then compound is called is D
If OH here is same side on the fisher projection of L-glyceraldehyde,
Then compound is considered L
(By figure 7-1& 7-2 right D, left is L)
Show and explain based on figure 7-3

Sugars differ only by 1 optical center called epimers
D and L glucose, D glu and D man, D glu and D gal (**7-3a**)
some L sugars occur in nature

C. Common Monos have cyclic structures

have shown as straight chains to clarify structure

aldotetroses (4C aldehydes)

And all sugars >4 occur primarily in ring structures

result of internal reaction between alcohol and C=O for form

Hemiketal or hemiacetal

reaction shown figure 7-5

note this reaction turns C=O carbon into asymmetric center!

Designate as α or β depending on which way the new OH points

5 member rings called furanoses

6 member rings called pyranoses

isomeric forms differ only in hemi-carbon are called **anomers**

the Hemi atom is called the **anomeric C**

α and β forms interconvert until come to equilibrium

So now our simple D-glucose in solution

1/3 α , 2/3 β and very small amount linear and furanose (5 member rings)

Usually write using structure in Haworth projections

(Figure 7-7)

But remember, rings also have boat and chair conformations

D. Organisms contain a variety of hexose derivatives

lots of derivatives based of sugar structures

Oxidize to COOH

Oxidize aldehyde get aldonic acid

Oxidize C-6 get uronic acids i.e. glucuronic

Add amine

Reduce to methyl

Do more than 1

See figure 7-9

One special N-acetylneuraminic acid (sialic acid) used in many glycoprotein and glycolipids

in metabolism intermediates are often phosphorylated intermediate

serves to trap in cell since no transport protein for phosphorylated sugars

E. Monosaccharide are reducing agents

carbonyl C of many sugars can be oxidized by mild oxidizing reagents, Fe^{3+} Cu^{2+} to COOH

these sugars called **reducing sugars**

basis for Fehlings test used to identify presence of reducing sugars

see figure 7-10

quantitative reaction so can use calculate sugar present

Used to be used to calculate blood or urine glucose levels for diabetics

F. Disaccharides

contain a glycosidic bond

O-glycosidic bond OH of one sugar reacts with anomeric C of another reaction of a hemiacetal to form an acetal

Figure 7-11

when anomeric (C=O) C of one sugar is tied up on this linkage it cannot be oxidized

in polysaccharides the end of the chain with a free anomeric C

Is called **reducing end**, since it can be oxidized, so it is a reducing agent

glycosidic bonds easily hydrolyzed by acid

Resist base cleavage

Reducing end orientation also used in naming systems

for di or larger sugars

Start naming from non reducing end of sugar

Give α or β to name of first sugar to describe sugar and linkage to next sugar also include in name furano- or pyrano- to distinguish 5 or 6 member ring

of C used in linkage indicated in parentheses

Name the next sugar

Continue as necessary

Many use abbreviations shown in **table 7-1** and skip furano/pyrano for larger more complex sugars

See how Maltose (**figure 7-11** is named)

Now try names **figure 7-12**

Note sucrose - how have linked BOTH reducing ends

Is a non-reducing sugar

Use a double headed arrow

N-glycosidic bonds used to link anomeric C to N of bases

Found in nucleotides and glycoproteins

7.2 Polysaccharides

polymers of medium to high molecular weight

also called glycans

will differ in sugars, sequence, length, linkages, degree of branching

Homopolysaccharides - single sugar

heteropolysaccharides two or more sugars

Homopolysaccharides - often used for energy storage (glycogen or starch) or structural elements in plant cell walls or animal exoskeleton (cellulose, chitin)

Heteropolysaccharides - used in extracellular support for cells from bacteria on up

not built as carefully as proteins - no specific template is used
just start enzymes polymerizing
so have a range of MW
have a range of sequence, branching, etc

A. Homopolysaccharides used for E storage

Starch and Glycogen - stored fuels
Both are glucose homopolymers
both are stored in cells as clusters or granules
both are heavily hydrated

I. Starch **Figure 7-13, 7-14, 7-20**

Actually 2 polymers
Amylose
Linear $\text{glc}(\alpha 1 \rightarrow 4)\text{glc}$
MW 1000s to 10^6

Amylopectin
 $\text{glc}(\alpha 1 \rightarrow 4)\text{glc}$
But about 1 in 25-30 $\text{glc}(\alpha 1 \rightarrow 6)\text{glc}$ additional branch

Tend to roll into helices or double helices
Tries to tangle with each other
?? This is why jellies and jams jell?? Pectin

II. Glycogen

Similar to amylopectin, but branches about every 10 (8 to 12)
More compact
Can be up to 7% of liver wt
Also found in muscle
Many non reducing ends, 1 reducing end
When using as fuel take from the many nonreducing ends so faster

Why glycogen and starch

1 lower osmotic pressure
2 if have high \square in cell will take more E to transport in

B. Homopolysaccharides used for Structure

Cellulose and Chitin

i. Cellulose 7-15

fibrous, tough water insoluble found in plant cell walls
 most of the mass of wood
 cotton is almost pure cellulose
 linear unbranched homopolymer of glucose
 β 1-4 linkage
 change of α to β makes big difference in structure
 (Will discuss in a minute)

α linkage (starch glycogen) hydrolyzed by α amylases from saliva and intestine
 most animal lack β amylase needed to break down cellulose
 Microorganism *Trichonympha* found in gut of termite is how it gets it done
 cows and other ruminants have extra stomach compartment with bacteria that can do this

ii. Chitin

linear homopolymer of N-acetylglucosamine
 Figure 7-17
 β 1-4 linkage
 looks like cellulose with C2 OH replaced with NHCOCH_3
 not digestible by vertebrates
 hard exoskeleton of arthropods (insects, lobsters, crabs)

C. Steric factors and H-bonding in polysaccharide folding

3D structure is of carbohydrate, like proteins is determined by 100's of low energy noncovalent interactions that hold the polymer in a given conformation

Carbohydrates polar and lot of H-bonds

So hydrophobic interaction is less important

Won't see a hydrophobic core

Maximizing H bonds will be dominate force

Will also see charge/charge and charge/H-bond interactions in derivatives of carbohydrates

As in proteins, not all conformations are allowed because of steric interactions

Sugar ring is rigid, rotation occurs between sugars Figure 7-18

Can make Ψ Φ plot like Ramachandran plot

Did for peptides

Figure 7-19

See 2 main conformations

Flat out

Bent

(Compare 7-15 to 7-21 or bring in models)

In α linkage sugars bents linear, can't get close enough to H bond unless roll in helix (7-18a 7-20)

Has lots of H- bonds to solvent so very water soluble

This is structure you see in starch and glycogen

In β linkage line up to get H bonds from one sugar to next to make rigid if flip 180 each sugar (7-18b, not amylose but cellulose! 7-16)

This is structure you get for cellulose and chitin

Further, when lie side-by-side get H bonds from to chain to another

All H bonds satisfied, no need for water so virtually insoluble in water

D. Bacterial and algal cell wall of peptidoglycans

Bacterial

Figure 20-31

rigid part of cell wall

heteropolymer of N-acetylglucosamine (β 1-4) N acetylmuramic

linear polymer lie side by side, crosslinked by peptides

peptides vary on bacterial species

Lysozyme hydrolyzes β 1-4 linkage between GlcNAc and Mur2Ac

this is why used in Mol Bio to lyse bacterial cells

found in tears, Snot? - Eye defense against bacteria

Saw in last chapter tie penicillin

Algal

Certain marine red algae have compound called agar in cell wall

(Yes, it's the same agar you use in molecular biology)

Figure 7-21

D-Galactose linked to a modified L galatose

agar is complex mixture based on above backbone, but varying amount sulfonate and pyruvate

Agarose a component of agar with fewest charged groups, with great gel forming properties

Two molecules from a double helix, trapping water inside the helix

Then helices associate trapping more waters to make a rigid but very hydrated gel

E. Glycosaminoglycans in the extracellular matrix

Space between cells in multicellular animals filled with gel called **extracellular matrix** or ground substance

Holds cells together, give porous pathway for diffusion of metabolites
interlocking meshwork of heteropolysaccharides and fibrous proteins

Proteins included collagen, elastin, fibronectin, laminin

Heteropolysaccharide

Glycoaminoglycans

A family of linear polymers, with different units

Figure 7-22

Sulfated sugars have high negative charge density

Used for recognition and binding of protein ligands

Glycoaminoglycan and attached protein called
proteoglycans

(More details next section)

Hyaluronates

Highly viscous solutions

Used for lubricant in joint fluid

Used for gel in eye

Important in cartilage and tendon

Chondroitin sulfate

Cartilage, tendon, ligaments

Keratin Sulfates cornea, cartilage, bone, horny structures

7.3 Glycoconjugates : Proteinglycans, glycoproteins, glycolipids

Sugars used as information carriers

destination labels for proteins

mediate cell-cell interactions

Cell-matrix interactions

cell-cell recognition and adhesion

cell migration

blood clotting

immune response

most of the above done as **glycoconjugate** sugar joined to something else

i. Proteoglycans

Macromolecules at cell surface or extracellular matrix

glycoaminoglycan covalently joined to membrane protein or

secreted protein

sugar is usually bulk of the molecule and main site of activity
major component of connective tissue

ii. Glycoproteins

one or several oligosaccharides covalently linked to a protein
found on outer face of PM, extracellular matrix, in blood
inside cells found in organelles like

Golgi

Secretory granules

Lysosomes

Sugars usually smaller than proteins

Sugar structure more detailed, usually more information for binding
and recognition

iii. Glycolipids

membrane lipids with sugars for polar head

also used for specific recognition

This is a hot area

A. Proteoglycans of the cell surface and extracellular matrix

Mammalian cells up to 40 different types of proteoglycans

Major influence on cell-cell interaction, growth factor activation & cellular
adhesion

part of basal lamina - base matrix surface that epithelial cells grow on
usually a core protein and covalently attached glycoaminoglycan

Family of core proteins 20-40,000 MW

Each has several covalently attached heparin sulfate chains

Heparin sulfate similar to heparin but with lower sulfate conc

Structure heparin given V&V page 258

Point of attachment is usually a ser in Ser-Gly-X-Gly sequence

Linkage through a tetrasaccharide bridge (**see figure 7-24**)

Structure of heparin

Many proteoglycans secreted into matrix

some are integral membrane proteins (**see figure 7-25**)

Two major families of membrane heparin sulfate proteoglycans

Syndecans -

protein has one domain in membrane and second outside
membrane

3-5 heparin sulfate molecules

Can have additional chondroitin sulfate molecules

Syndecans can be found in extracellular matrix rather than on
cell surface

A specific protease cleaves protein near cell surface

Glypicans

Protein is outside cell but anchored to membrane by a lipid anchor

Glypicans can also be found in extracellular matrix rather than on cell surface

A specific lipase cleaves lipid from protein near cell surface

Chains can bind a variety of ligands, so modulate ligand-cell interaction

Can see structural domains within heparin sulfate molecule

Domain 3-8 disaccharides

Differ in sequence and ability to bind specific proteins

Highly sulfonated domains (called NS)

specifically to extracellular proteins to alter their activities through several different mechanisms

Alternate with GlcNac-GlcA domains (called NA)

Exact pattern varies with particular proteoglycan

Same core protein may have different heparin sulfate structures in different cell types

NS domains bind specifically to extracellular proteins to alter their activities

Proteoglycan aggregates

Enormous assemblies See figure 7-27

Core-single hyaluronan core up to 50,000 disaccharide

Attached to this are 100 or more Aggrecan core proteins (Mass ~ 250,000)

Attached to each core protein are many chondroitin sulfate and keratan sulfate molecules

Net molar mass $>2 \times 10^8$

Volume of single molecule and associated water the size of a bacterial cell!!!

Aggrecan core protein interact strongly with collagen (remember what that is?) To make connective tissue

Interwoven between the above are the fibrous proteins

(Like collagen, elastin, fibronectin)

Often contains domains to bind to the various saccharides

Hold everything together with various levels of rigidity

The 3-D matrix of interaction is what holds you and me together!

Note also interwoven are the glycosaminoglycans we talked about in the last section, pure heteropolysaccharides with no proteins.

Net of all of the above- Extracellular matrix **Figure 7-28**

B. Glycoproteins

About ½ of all mammalian proteins are glycosylated
 about 1% of the genes are involved in synthesis and attachment of sugars to proteins

carbohydrates smaller but more structurally diverse
 attached through O of ser or thr

Peptide chain tends to be rich in Gly, Val & Pro
 attached through N of asn

Peptide chain has consensus of N-not prolein- S or T-
see figure 7-29

may be single carbohydrate or multiple

carbohydrate anywhere between 1 and 70% of mass

primary sequence of many sugars are known, a few are shown in figure

One class of glucoproteins called mucins

Typically a protein with many O-linked polysaccharides

Present in most secretions

Guess what mucus is made of?

One key feature is slipperyness

Another class in cytoplasm and nucleus

Single N-acetyl glucosamine O linked to Ser

Reversible and usually same site that is eventually
 phosphorylated/dephosphorylated

Glycomics - characterization of all carbohydrates in a cell

Many secreted proteins are glycoproteins

This includes antibodies, many peptide hormones, many milk proteins, many pancrease (digestive) proteins

Reason for adding carbohydrate not entirely clear

Very hydrophillic so change polarity and solubility

May be added during synthesis to influence folding events

May make more resistant to proteolysis

C. Glycolipids

carbohydrates attached to lipids

i. Gangliosides - membrane lipids of eukaryotic cells - polar head is

carbohydrate

Complex carbohydrate lots of sialic acid

Some carbohydrate are identical with ones found in glycoproteins

For instance the one used for blood typing

ii. Lipopolysaccharides- dominant surface feature in gram negative bacteria like e coli

Prime targets for antibodies used in defense

Example figure 7-30

7.4 Carbohydrates as informational signal

Carbohydrates decorate cell surface

Seems to me major mediator in cell/cell interactions

because have at least 20 different monosaccharides

can link a or b

can link 1-4 or 1-6

can have branches

with just six sugars estimate 1.44×10^5 different structures

(6 AA would give you 20^6 or 6.4×10^7 structures)

(6 nucleotides would give you 4^6 or 4,096 structure)

so is extremely rich in information

A. Lectin

Lectins protein that bind carbohydrates with high affinity and specificity found in all organisms

protein uses hydrogen bonding to identify sugars and sequence

can distinguish between closely related sugars

Used in wide variety of cell-cell recognition and adhesion processes

can be used in the lab for separating different glycoproteins

How lectins are used in your body

many plasma glycoproteins have sialic acid (Neu5AC)

Fig 10-11 page 252 Voet&Voet

At ends of oligosaccharides

body removes sialic acid with sialidase to mark protein as 'old'

Protein then removed and destroyed

Mediated by a lectin for asialooligosaccharides

Similar mech for removal of old red blood cells

if treat fresh RBC with sialidase then put back in body, removed in a few hours

Similar mech. used for some peptide hormones although different sugar is used at ends

Remove sugar, hormone taken out of circulation

Selectins **Figure 7-31**

Family of lectins found in cell PMs mediate cell-cell recognition
1 case

Near infection P selectin expressed on surface of capillary endothelial cells

Interacts with oligosaccharide on surface of T lymphocyte

Added integrin molecules on lymphocyte and endothelial cell surface

Makes cell stop and exit capillary to go to infection site

Some pathogens use lectin to bind to host cells to start their attack

H pylori, responsible for many ulcers, binds to cell on inner surface of stomach this way

Cholera toxin molecules binds to intestinal cells via oligosaccharide of GM₁

Pertussis toxin same way

several animal viruses binds to oligo's on host cell surface

Lectins are used intracellularly for sorting

An oligosaccharide containing mannose-6-P marks newly made protein for transfer from golgi to lysosome through a lectin mediated process

Lectins in ER serve to 'proof-read' for folding

All protein synthesized in ER have oligosaccharides attached

Oligo's can be bound by calnexin (membrane bound lectin)

Or calreticulin (soluble lectin)

Both lectins used to bring newly synthesized protein to a protein disulfide isomerase (PDI)

In correctly folded protein, carbo is trimmed one way and another lectin draws protein to Golgi complex for further maturation

If protein not folded corrected carbois trimmed differently, and a different lectin starts process of kicking mis-folded protein into cytosol for recycling!

B. Lectin-Carbohydrate Interactions **Figure 7-33**

Have been able to crystalize lectin bound carbohydrate so can analyze interactions

Thanks to structure and H-bonds can build very strong, highly specific sites to bind every sugar, but more than that, can also use slat bridges and hydrophobic interactions, so just like protein-protein interactions

Binding of an individual carbohydrate and single carbohydrate binding

domain of lectin usually has a weak interaction micro to millimolar K_D

But since lectin uses multiple binding sites and looks for multiple interactions (cooperativity again) the lectin-carbo interaction can become very strong and very specific

7.5 Analysis of carbohydrates

tougher than protein or DNA because branches and many different linkages generally remove from protein or lipid

then stepwise degradation to find each bond

use mass spec

use NMR for structure

releasing oligo's from proteins

use glycosidases specific for N- or O- linkages

use lipases to remove from lipids

After that purification similar to protein purification

Analysis follows logic [Figure 7-36](#)

hydrolyze complexly to find the individual sugars

Methylate all OH's, hydrolyze again

Any free OH's were in sugar-sugar bonds

use exo- and endoglycosidases to clip into pieces for analysis

use mass spec and NMR to sequence

Once you know what the pieces look like, you can start doing mass spec of a sample to see all the oligo in that sample [Figure 7-37](#)