Protein Structure

A. Stability
   1. Two-state model
   2. Energetics
   3. Denaturation
   4. Methods to study

B. Protein Folding
   1. Evidence-Anfinsen
   2. Protein Folding
      1. Pathways
      2. Mechanisms
   3. Diseases

Lecture 21 (11/6/23)

- Reading: Ch7; 229-235, 235-241, 251-254, 258-260
- Homework #21

NEXT

- Reading: Ch7; 241-250
- Homework #22

Carbohydrates

A. Definition
B. Roles
C. Monosaccharides-Chemistry
   1. Chirality
      a. One or more asymmetric carbons
      b. Linear and ring forms
   2. Derivatives: the chemistry of carbohydrates
      a. Oxidation
         i. C1
         ii. C6
      b. Reduction
         i. C1/C2
         ii. Other carbons
      c. Ester formation
      d. Amino sugars
   3. Polymerization
      a. The Glycosidic Bond
      b. Non-covalent bonds in macro-molecular structure

D. Oligosaccharides
   1. Glycoproteins & glycolipids
   2. O-linked
   3. N-linked
   4. Sequence determination-ABO

E. Polysaccharides
   1. Polymers of glucose
   2. Polymers of disaccharides

Protein Stability & Protein Folding and Dynamics

Primary structure determines Tertiary structure!

Protein Stability & Protein Folding and Dynamics

How Can Proteins Fold So Fast?

- Proteins fold to the lowest-energy state in the μsec–sec time scales. How can they find the right fold so fast?
- It is mathematically impossible for protein folding to occur by randomly trying every conformation until the lowest-energy one is found (Levinthal's paradox).
- Search for the minimum is therefore not random; there must be a PATHWAY toward the native structure, which is thermodynamically most favorable.

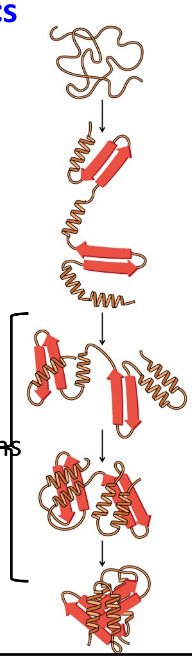
Protein Stability & Protein Folding and Dynamics

Kinetics

How? Protein folding

<10 ms

<5-1000 ms



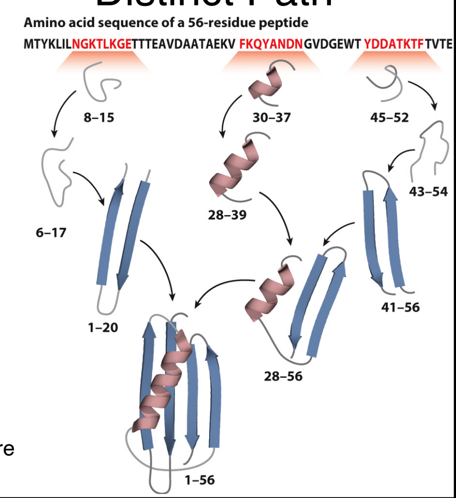
2° Structures form

3° Structural Domains form

Molten Globule

3° Structure forms/4° structure forms

Protein Folding Follows a Distinct Path



Protein Stability & Protein Folding and Dynamics

Computer Simulation of Protein Folding: A 40 residue protein

(video is on Web site under announcements)

Simulating protein folding on the millisecond timescale has been a major challenge for many years. Recently, Folding@home researchers Vincent Voelz, Greg Bowman, Kyle Beauchamp, and Vijay Pande have broken this barrier. This is a movie of one of the trajectories that folded (i.e. started unfolded and ended up in the folded state). See Voelz *et al.* (2010) *J. Am. Chem. Soc.*, **132**:1526 for more details.

Protein Stability & Protein Folding and Dynamics

Protein folding help

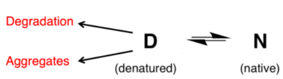
While the primary structure dictates what the fold will be, IN THE CELL, proteins often need help.

Kinetic help

1. Protein disulfide isomerase (PDI)
2. Peptide Prolyl Isomerase (PPI)

***Energetic help (prevent aggregation; maintain equilibria)**

1. Molecular chaperones – heat-shock proteins (Hsp)
   - Hsp70, Hsp90, etc.
2. Molecular chaperonin
   - GroEL-GroES



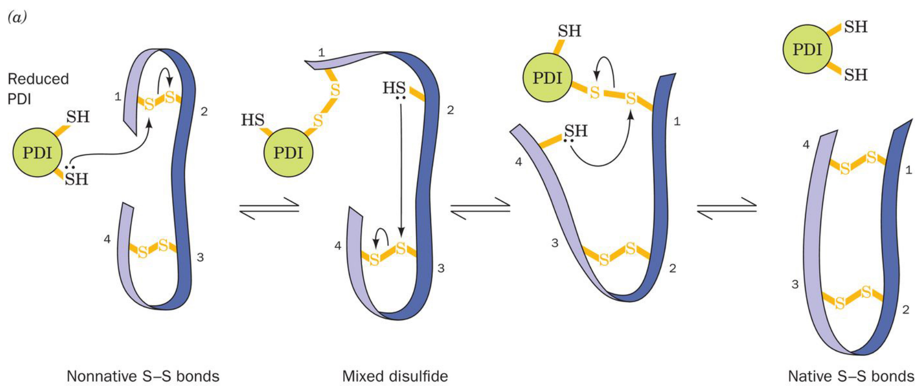
***Thermodynamics**

Protein Stability & Protein Folding and Dynamics

Kinetic help

Protein Disulfide Isomerase (PDI) Catalyzes Disulfide Interchange

Incorrect disulfides



Correct disulfides

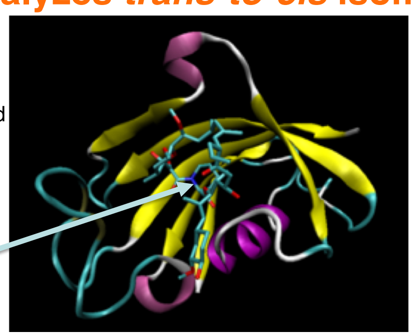
Protein Stability & Protein Folding and Dynamics

Kinetic help

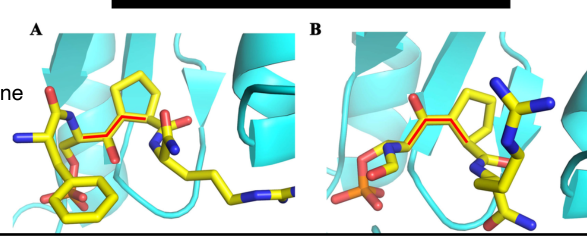
Peptidyl Prolyl Isomerase (PPI) Catalyzes *trans-to-cis* isomerization

Without catalysis this takes 8-10 min!

Activation energy reduced 20 kcal/mole by PPI → 10^{15}-fold rate increase



Pro



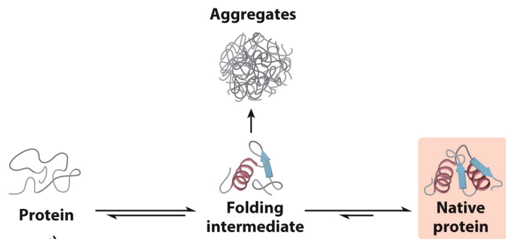
trans-proline

cis-proline

Protein Stability & Protein Folding and Dynamics

Thermodynamic help

Chaperones Prevent Misfolding and Aggregation



Heat-shock proteins (HSP)

Protein Stability & Protein Folding and Dynamics

Thermodynamic help

Chaperones Prevent Misfolding and Aggregation

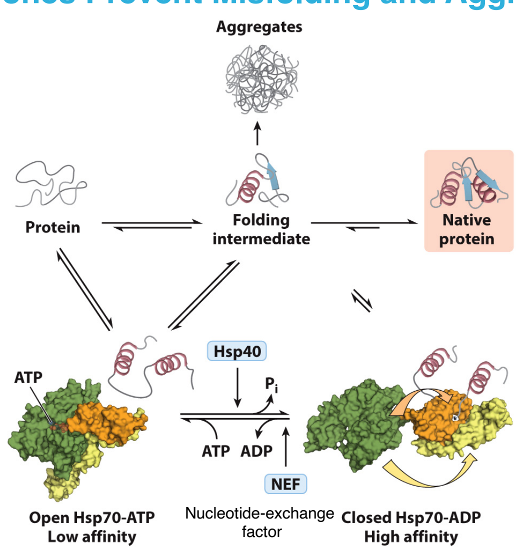


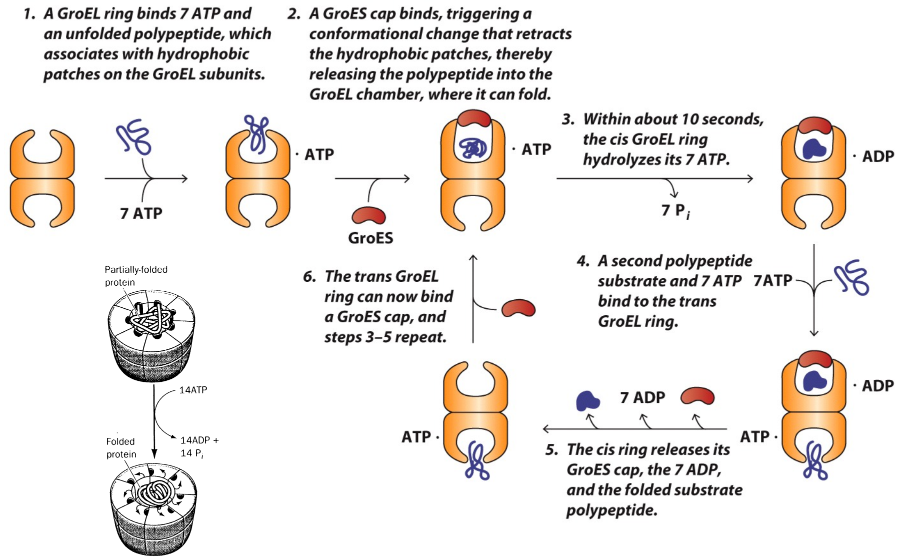
Figure 4-30

Heat-shock proteins (HSP)

Protein Stability & Protein Folding and Dynamics

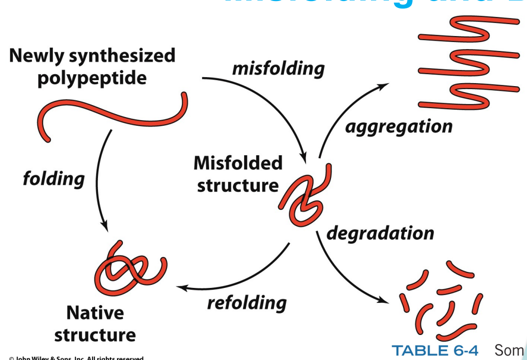
Thermodynamic help

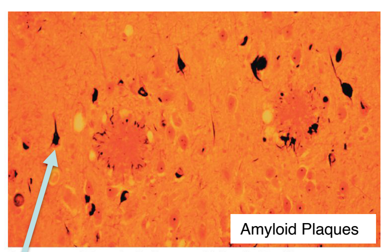
Facilitated folding: Chaperonin



Protein Stability & Folding Disease

Misfolding and Disease





Courtesy of Dennis Selkoe and Marcia Podlisny, Harvard University Medical School

TABLE 6-4 Some Protein Misfolding Diseases

Disease	Defective Protein
Alzheimer's disease	Amyloid-β protein
Amyotrophic lateral sclerosis	Superoxide dismutase
Fibrinogen amyloidosis	Fibrinogen α chain
Huntington's disease	Huntingtin with polyglutamate expansion
Light chain amyloidosis	Immunoglobulin light chain
Lysozyme amyloidosis	Lysozyme
Parkinson's disease	α-Synuclein
Transmissible spongiform encephalopathies (TSEs)	Prion protein

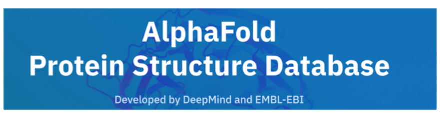
Protein Stability & Protein Folding and Dynamics

Protein Prediction

If the 1° structure is known, but only the 3° structure of a *related* homologous protein is known, a prediction of your protein can be done by "homology modeling" (threading).

BUT, WHAT IF NO STRUCTURE?

- Given all the known 3D structures, predictions of propensities to find residues and/or sequences of residues in certain structures have been effective.
  - e.g., already discussed propensities of residues to be in α-helices, β-sheets, and β-turns.
- Computer programs can now predict to about 90% certainty where these 2° structures will be in a given 1° sequence. And, the overall-fold prediction is pretty good (>80%). See Alpha-Fold:



nature > articles > article

Article | Open access | Published: 15 July 2021

Highly accurate protein structure prediction with AlphaFold

John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, Alex Bridgland, Clemens Meyer, Simon A. A. Kohl, Andrew J. Ballard, Andrew Cowie, Bernardino Romera-Paredes, Stanislav Nikolov, Rishub Jain, Jonas Adler, Trevor Back, Stig Petersen, David Reiman, Ellen Clancy, Michal Zielinski, ... Demis Hassabis + Show authors

Nature **596**, 583–589 (2021) | Cite this article

Protein Stability & Protein Folding and Dynamics

Protein Stability

Key Concepts

- Protein stability depends primarily on hydrophobic effects and secondarily on electrostatic interactions.
- A protein that has been denatured may undergo renaturation.
- Protein structures are flexible and dynamic; may include unfolded regions.

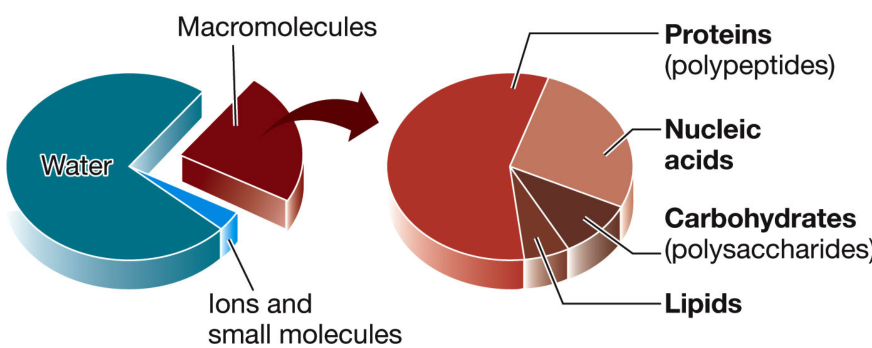
Protein Folding

Key Concepts

- Primary sequence dictates tertiary fold.
- A folding protein follows a pathway from high energy and high entropy to low energy and low entropy.
- Protein disulfide isomerase (PDI) catalyzes disulfide bond formation/exchange and Peptidyl Prolyl Isomerase (PPI) catalyzes *cis*-proline peptide bond formation.
- A variety of molecular chaperones assist protein folding via an ATP-dependent bind-and-release mechanism, including the Chaperonin "Anfinsen Chamber."
- Amyloid diseases result from protein misfolding, with many misfolded proteins forming fibrils with extensive β structure.

Carbohydrates

Carbohydrates



Carbohydrates

Definition

- Carbo-Hydrate: have formula $C_n(H_2O)_n$ (for n ≥ 3)
- The precursor-macromolecule relationship is:
  - Monosaccharide–polysaccharide (or oligosaccharide)
- Carbohydrates are everywhere (ubiquitous) and versatile in function; fulfill a variety of functions.
  - Can be covalently linked with proteins and lipids; are intimately involved in nucleic acids

ROLES	Monosaccharide	Polysaccharide
1. Energy source/storage	glucose, fructose, etc.	Starch, glycogen
2. Structure	glucose, *glycerol*	Cellulose, chitin, lipids & membranes
3. Information	ribose (nucleotides)	Nucleic acids
4. Recognition	many	Glycolipids & glycoproteins

Carbohydrates

The 4 S's

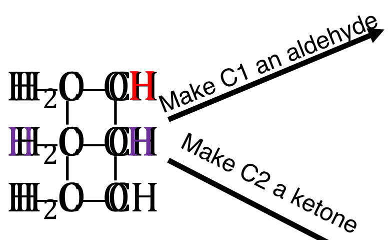
Size	Range from as small as glyceraldehyde (M_w = 90 g/mol) to as large as amylopectin (M_w > 200,000,000 g/mol)
Shape	Depends on size and glycosidic bond
Solubility	Very polar, very soluble, until large polymers
Stability	Stable due to glycosidic bond

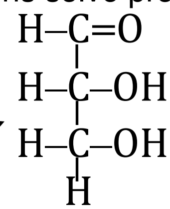
Carbohydrates

Monosaccharides

$C_3(H_2O)_3$

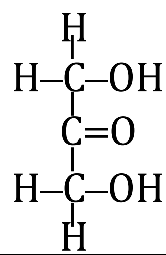
- Basic nomenclature:
  - Use the suffix "-ose" — triose, tetrose, pentose, hexose
  - Aldehyde & ketone functions solve problem of a "carbo-hydrate"





= aldoses

aldotriose (triose), aldotetrose(tetrose), aldopentose, aldohexose



= ketose

Ketotriose (tri**u**lose), ketotetrose (tetr**u**lose)… (pent**u**lose, hex**u**lose).

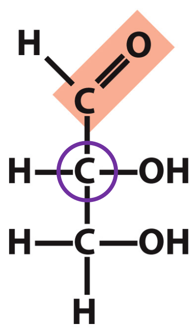
OK, we know enough chemistry to draw a carbon-hydrate!

What is wrong with this structure?

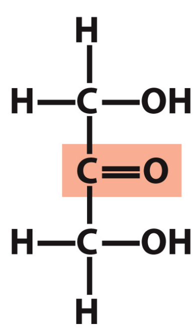
Carbohydrates

Monosaccharides

- An aldose is a carbohydrate with aldehyde functionality.
- A ketose is a carbohydrate with ketone functionality.



Glyceraldehyde, an aldotriose



Dihydroxyacetone, a ketotriose

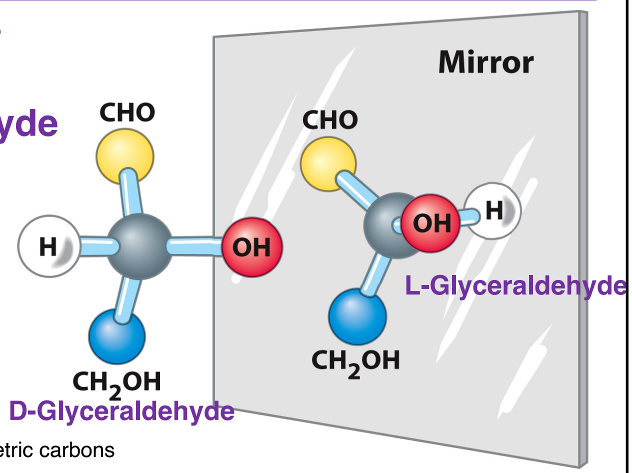
Are there any chiral carbons?

Figure 7-1a
Lehninger Principles of Biochemistry, Seventh Edition

Carbohydrates

Monosaccharides

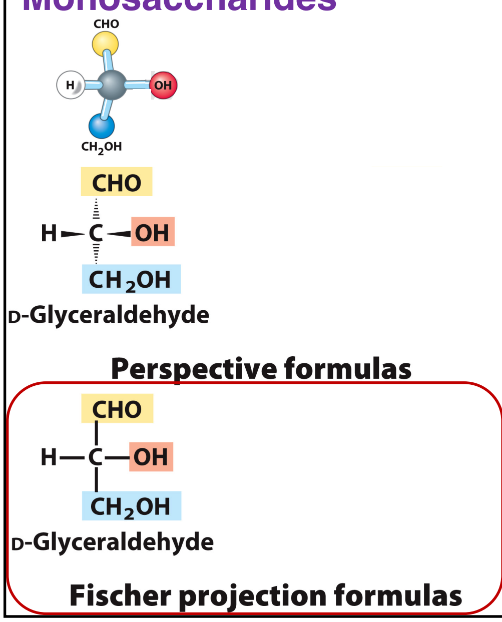
Glyceraldehyde



- Chemical Features:
  - Chirality
    - One or more asymmetric carbons
    - Linear and ring forms
  - Derivatives: the chemistry of carbohydrates
  - Polymerization
    - The Glycosidic Bond
    - Non-covalent bonds in macro-molecular structure

Carbohydrates

Monosaccharides



Fischer projections

- Vertical bonds are between carbons, with highest oxidation state at the top, AND project away from you.
- Horizontal bonds are pointing toward you.
- If hydroxyl is on the left; its **L**
- If hydroxyl is on the right; its **D**
- Here D=R & L=S

It turns out that the L form of glyceraldehyde is called L because it is "levorotary," meaning it will rotate plane-polarized light to the left, or counter-clockwise. D rotates "dextrorotary."

Carbohydrates

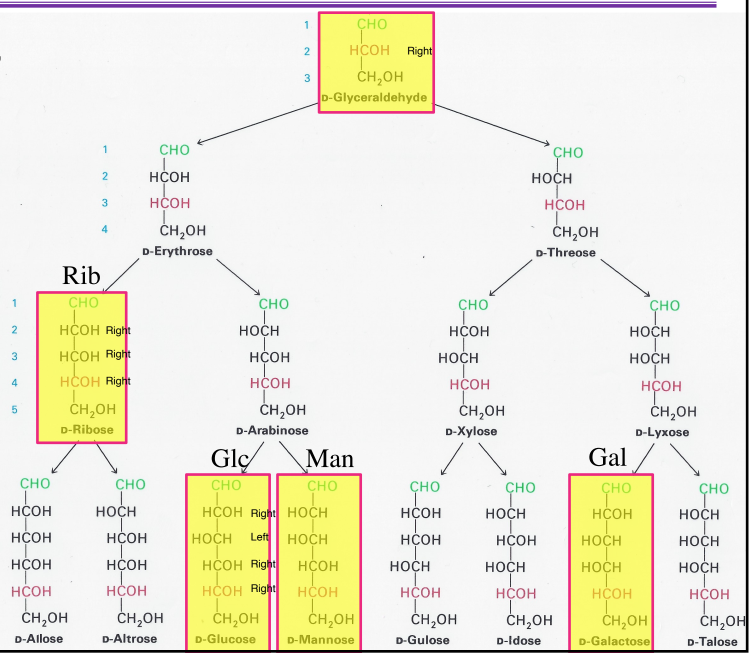
GREEN =functional group
RED=highest numbered chiral carbon

Nearly all sugars in biology are "D" (sort of like most amino acids are "L").

As you go to tetrose, pentose, etc., you are adding more chiral carbons.

A D- or L-sugar is defined by the chirality of the highest numbered carbon.

Another complete set of 15 L-aldoses exist, and are enantiomers of their D-aldose relatives shown here.



Carbohydrates

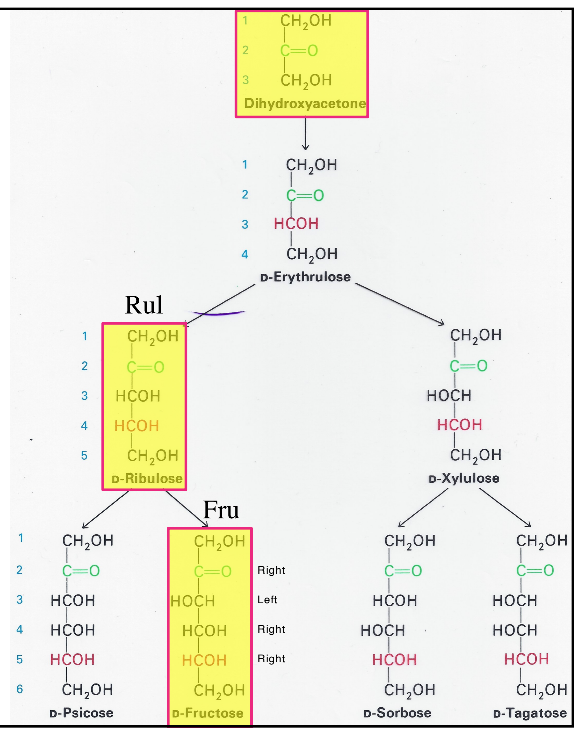
D-ketoses

The first chiral sugar is a tetrulose.

There is also another complete set of 7 L-ketoses exist, and are enantiomers of their D-ketose relatives shown here.

Fructose is the ketose form of glucose

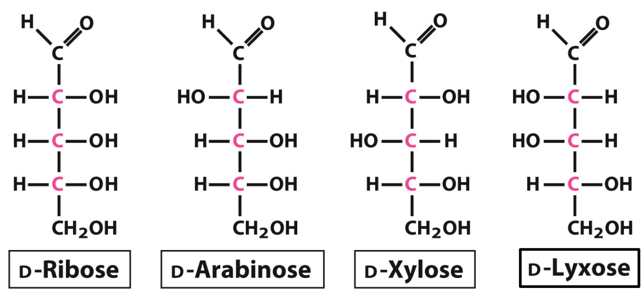
Ribulose is the ketose form of ribose



Carbohydrates

Monosaccharides: Stereoisomer Nomenclature

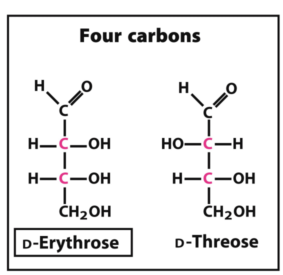
- Enantiomers
  - stereoisomers that are nonsuperimposable **complete** mirror images
  - Example: D-sugars & L-sugars
- Diastereomers
  - stereoisomers that are not complete mirror images
  - Diastereomers have different physical properties (e.g., water solubility)
  - Example:



Carbohydrates

Monosaccharides: Stereoisomer Nomenclature

- Epimers
  - Epimers are stereoisomers that differ at only **one** chiral center
  - Epimers are diastereomers; diastereomers have different physical properties (i.e., water solubility, melting temp)
  - Example:

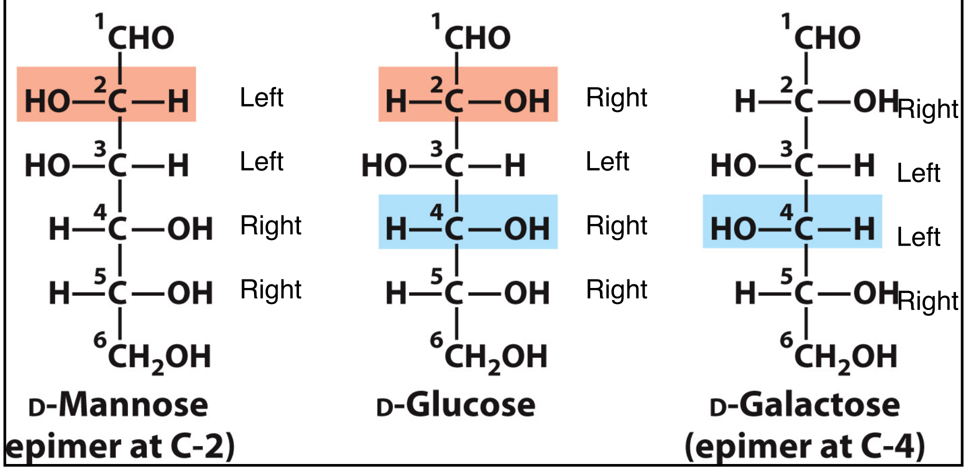


  - Glc & Man (2 epimer); Glc & Gal (4 epimer); Rib & Ara; Rib & Xyl
- Anomers
  - Anomers have different chirality at carbon involved in ring formation

Carbohydrates

Monosaccharides: Stereoisomer Nomenclature

- D-Mannose and D-galactose are both epimers of D-glucose.
- D-Mannose and D-galactose vary at more than one chiral center and are diastereomers, but <u>not</u> epimers.



Carbohydrates

Monosaccharides: The most important sugars

- Glyceraldehyde and dihydroxyacetone are the simplest (3 carbon) aldose and ketose, respectively.
- Ribose (Rib) is the standard five-carbon sugar.
- Glucose (Glc) is the standard six-carbon sugar.
- Galactose (Gal) is an C4-epimer of glucose.
- Mannose (Man) is an C2-epimer of glucose.
- Fructose (Fru) is the ketose form of glucose.
- Ribulose (Rul) is the ketose form of ribose.

Need to know, recognize, draw Fisher Projection, name, abbreviate

Carbohydrates

Monosaccharides: Stereoisomer Nomenclature

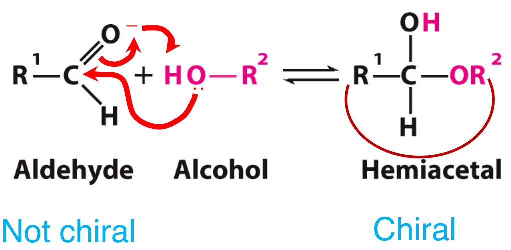
- Enantiomers
  - stereoisomers that are nonsuperimposable **complete** mirror images
  - Example: D-sugars & L-sugars
- Diastereomers
  - stereoisomers that are not complete mirror images
  - Diastereomers have different physical properties (e.g., water solubility)
  - Example: ribose & lyxose
- Epimers
  - Epimers are stereoisomers that differ at only **one** chiral center
  - Epimers are diastereomers; diastereomers have different physical properties (i.e., water solubility, melting temp)
  - Example:
  - Glc & Man (2 epimer); Glc & Gal (4 epimer); Rib & Ara; Rib & Xyl
- Anomers
  - Anomers have different chirality at carbon involved in ring formation

Carbohydrates

Monosaccharides: What are these "ring" forms?

- SUGARS WITH ≥5 CARBONS RAPIDLY AND STABLY FORM **RINGS** THROUGH HEMIACETAL (ALDOSES) AND HEMIKETAL (KETOSES) BONDS.

Hemiacetals



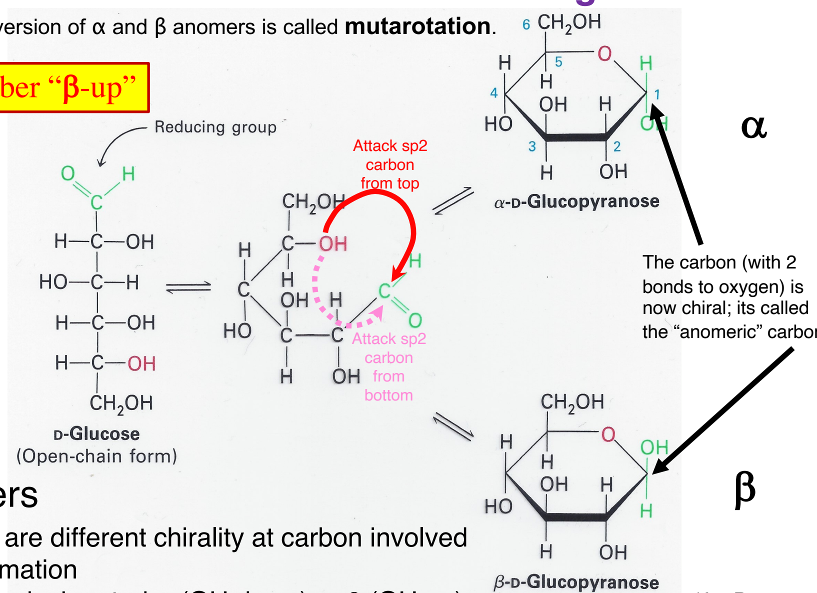
- If the aldehyde (R^1) and the alcohol (R^2) are on the same molecule, you have a RING!
- Due to the oxygen, the ring is heterocyclic

Carbohydrates

Monosaccharides: What are these "ring" forms?

The interconversion of α and β anomers is called **mutarotation**.

Remember "β-up"



- Anomers
  - Anomers are different chirality at carbon involved in ring formation
  - These are designated α (OH down) or β (OH up)

*for D-sugars

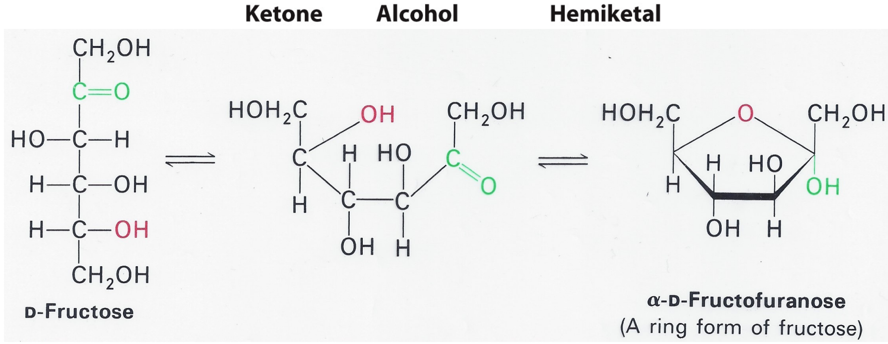
Carbohydrates

Monosaccharides: What are these "ring" forms?

Hemiketals

$$R^1-\overset{\displaystyle R^2}{C}=O + HO-R^3 \rightleftharpoons R^1-C(OH)(R^2)-OR^3$$

Ketone **Alcohol** **Hemiketal**



α (OH down) or β (OH up)

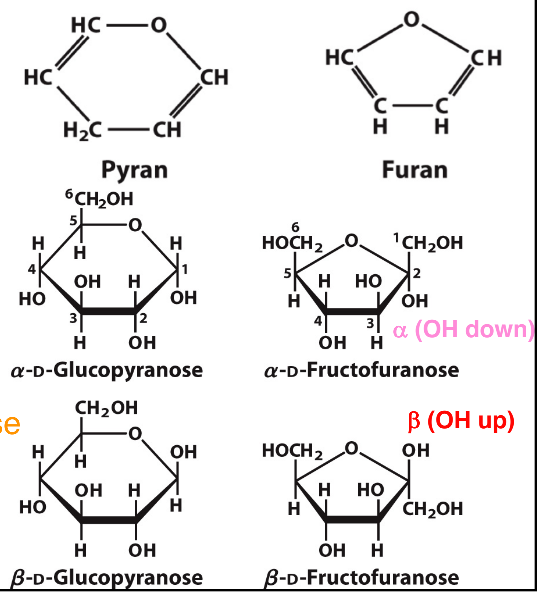
Besides the α & β, what is this pyran and furan?
Its all about which alcohols are used in these reactions.

Carbohydrates

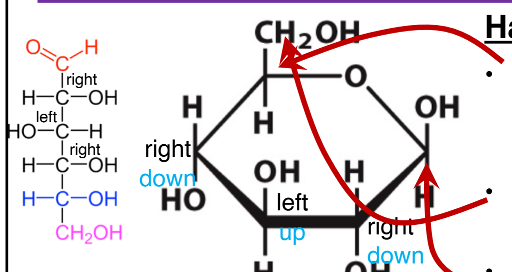
Monosaccharides: What are these "ring" forms?

- Six-membered oxygen-containing rings are called pyranoses after the pyran ring structure.
- Five-membered oxygen-containing rings are called furanoses after the furan ring structure.

The way we are drawing these sugars is called a Haworth projection…..



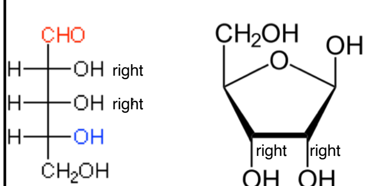
Carbohydrates



D-Glucose β-D-Glucopyranose

Haworth projections

- Highest carbon with hemiacetal/hemiketal alcohol in back; oxygen to right.
- If chiral (as for D-sugars), next highest carbon up if D, or down if L.
- The anomeric carbon is usually drawn on the right side.
- Other chiral carbons are down if <u>right</u> in Fisher projection and up if <u>left</u> in Fisher



D-ribose β-D-Ribofuranose

- β-D-Glc pyranose goes up, down, up down, up as you go from C5 to C1
- β-D-Rib furanose goes up, down, down, up as you go from C4 to C1

Carbohydrates

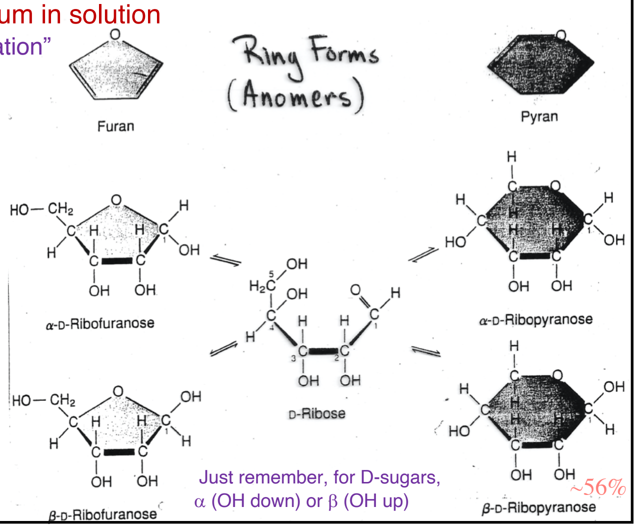
- Pentoses and hexoses readily undergo intramolecular cyclization.
- The former carbonyl carbon becomes a new chiral center, called the anomeric carbon.
- ALL forms are in equilibrium in solution

  Process is called "mutarotation"

– When the former carbonyl oxygen becomes a hydroxyl group, the position of this group determines if the anomer is *α* or *β*.

– If the hydroxyl group is on the opposite side (trans) of the ring as the CH_2OH moiety, the configuration is *α*.

– If the hydroxyl group is on the same side (cis) of the ring as the CH_2OH moiety, the configuration is *β*.



Just remember, for D-sugars, α (OH down) or β (OH up)

Carbohydrates

Relative amounts of tautomeric forms for some monosaccharide sugars at equilibrium in water at 40°C[a]

Monosaccharide	Relative Amount (%)				
	α-Pyranose	β-Pyranose	α-Furanose	β-Furanose	Total Furanose
➡ Ribose	20	56	6	18	24
Lyxose	71	29	[b]	[b]	<1
Altrose	27	40	20	13	33
➡ Glucose	36	64	[b]	[b]	<1
➡ Mannose	67	33	[b]	[b]	<1
➡ Fructose	3	57	9	31	40

[a] In all cases, the open-chain form is much less than 1%. For data on other sugars, see S. J. Angyal, The composition and conformation of sugars in solution, *Angew. Chem.* 8:157–226(1969).

[b] Much less than 1%.

Why is the β-Glc more stable than α-Glc, and *visa versa* for Man?

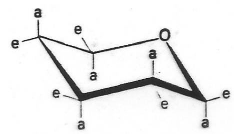
to answer this, we need to look at **ACTUAL** structures.

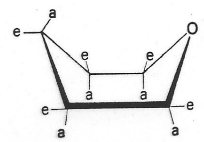
These Haworth Projections are not the actual conformation…. What is?

Carbohydrates

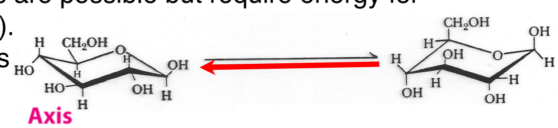
Actual Conformations of Cyclized Monosaccharides

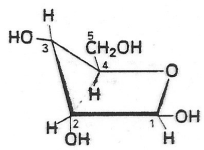
- Cyclohexane rings have “chair” or “boat” conformations.



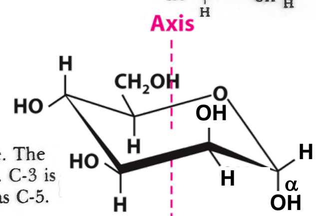


- Pyranose rings favor “chair” conformations.
- Multiple “chair” conformations are possible but require energy for interconversion (~46 kJ/mole).
- Furanose rings are envelopes





An envelope form of β-D-ribose. The $C_{3'}$-*endo* conformation is shown. C-3 is out of plane on the same side as C-5. Also termed “pucker”



Two possible chair forms of β-D-glucopyranose

We can understand why there is a 2:1 ratio of β:α glucose….. steric hindrance

AND why α–mannose is favored (2:1 ratio of α:β)

Carbohydrates

Monosaccharides: Chemistry

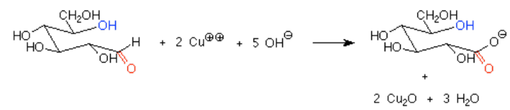
- **Chemical Features:**
  - Chirality
    - One or more asymmetric carbons
    - Linear and ring forms
  - Derivatives: the chemistry of carbohydrates
    - ① Oxidation
      - C1
      - C6
    - ② Reduction
      - C1/C2
      - Other carbons
    - ③ Ester formation
    - ④ Amino sugars
  - Polymerization
    - The Glycosidic Bond
    - Non-covalent bonds in macro-molecular structure

Carbohydrates ①

Monosaccharides: Derivatives

Oxidation: These make “sugar acids” or “acid sugars”

- Oxidation of aldehyde/ketone to acid (2e⁻ loss); reaction from the C1 of aldoses
  - – named as “onic” acids (“onate” for conjugate base)
  - – these sugars can reduce Cu^{2+} to Cu^{+} (Fehling’s/Benedict’s test) or Ag^{+} to Ag^{0} (Tollens’ test)
  - – Example: gluconic acid



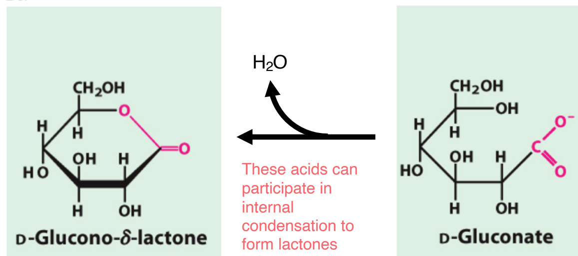
$CuSO_4 \rightarrow Cu^{2+} + SO_4^{2-}$
Blue

Cu^{2+} + Reducing sugar $\rightarrow Cu^{+}$

This reaction is called Benedict’s test

$Cu^{+} \rightarrow Cu_2O$
Red (ppt)

The copper is reduced and the sugar is oxidized. So, the sugar is called a **Reducing Sugar**. Reducing sugars have a free anomeric carbon.

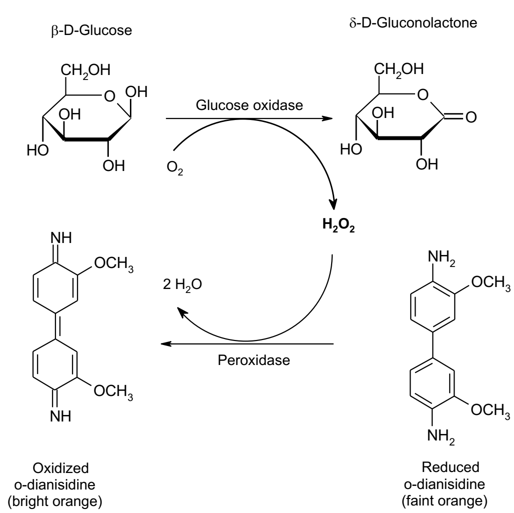


Carbohydrates ①

Monosaccharides: Derivatives

Oxidation: These make “sugar acids” or “acid sugars”

Colorimetric Glucose Analysis



- Enzymatic methods are used to quantify reducing sugars such as glucose.
  - – The enzyme glucose oxidase catalyzes the conversion of glucose to glucono-δ-lactone and hydrogen peroxide.
  - – Hydrogen peroxide oxidizes organic molecules into highly colored compounds.
  - – Concentrations of such compounds is measured colorimetrically.
- Electrochemical detection is used in portable glucose sensors.

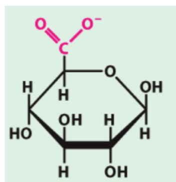
Carbohydrates ①

Monosaccharides: Derivatives

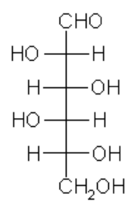
Oxidation: These make "sugar acids" or "acid sugars"

Oxidation of alcohol to acid (4e⁻ loss)

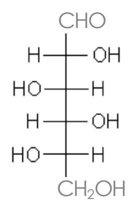
- Reaction for C6 groups; like C1 oxidation, normally on aldoses
- Named as **ur**onic acids ("uronate" for conjugate base)
- Many L-sugars are Uronic acids
- Examples: D-glucuronic acid, L-iduronic acid



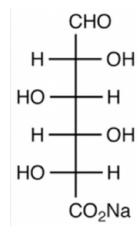
β-D-Glucuronate



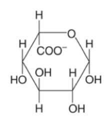
D-Idose



L-Idose



L-Iduronic acid



β-L-Iduronic acid (IdoA)

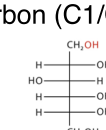
Carbohydrates ②

Monosaccharides: Derivatives

Reduction: These make "sugar alcohols" or "deoxysugars"

Reduction of aldehyde/ketone to alcohol (2e⁻ gain)

- Only carbon not already an alcohol is the anomeric carbon (C1/C2)
- Named as sugar **"ol"** or "itol"
- Examples: glycerol, mannitol, glucitol (sorbitol)

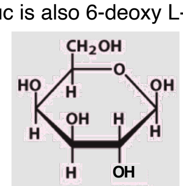


Sorbitol

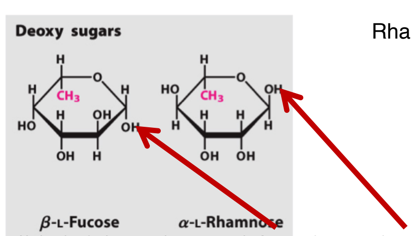
Reduction of alcohol carbon to methyl/methylene (2e⁻ gain)

- Can react at any except the anomeric carbon
- Named as "x-deoxy" sugar with x being the reduced carbon
- Many are L-sugars, and have specific trivial names
- Examples: 2-deoxyribose, L-Fucose (Fuc) and L-Rhamnose (Rha)

Fuc is also 6-deoxy L-Gal

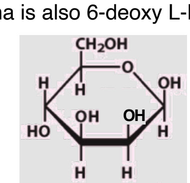


β-D-Gal



Notice that for L-sugars the β-anomer is **down** and α-anomer is **up**

Rha is also 6-deoxy L-Man



β-D-Man

Carbohydrates

Monosaccharides: Chemistry

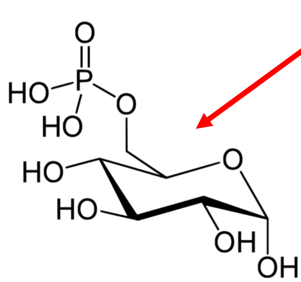
- **Chemical Features:**
  - – Chirality
    - One or more asymmetric carbons
    - Linear and ring forms
  - – Derivatives: the chemistry of carbohydrates
    - ① Oxidation
      - – C1 -onic
      - – C6 -uronic
    - ② Reduction
      - – C1/C2 -tol
      - – Other carbons deoxy-
    - ③ Ester formation
    - ④ Amino sugars
  - – Polymerization
    - The Glycosidic Bond
    - Non-covalent bonds in macro-molecular structure

Carbohydrates ③

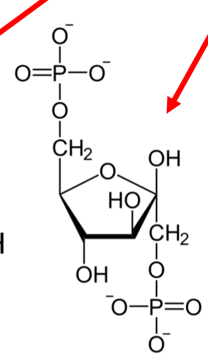
Monosaccharides: Derivatives

Esters: condensation of alcohol (sugar) and an acid

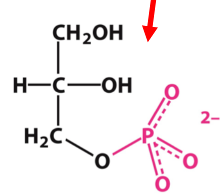
- Most important sugar esters use a phosphoric acid
  - These are called phospho-sugars or sugar phosphates
  - Examples: nucleotides, Glc 6-P, Fru 1,6-P_2, glycerol 3-phosphate



α–D–Glucose 6–phosphate



β–D–Fructose 1,6–bisphosphate



D-Glycerol 3-phosphate