OLLI AT ILLINOIS

Molecular Literacy for All

making sense of the “monstrous and boundless thicket” of everyday chemistry

The building blocks of life:

“As simple as can be?”

ILLINOIS
Beckman Institute for Advanced Science & Technology

Jeff Moore (email: jsmoore@Illinois.edu)

Today's Outline

Biological polymers

Polysaccharides (carbohydrates)

Polynucleotides (RNA and DNA)

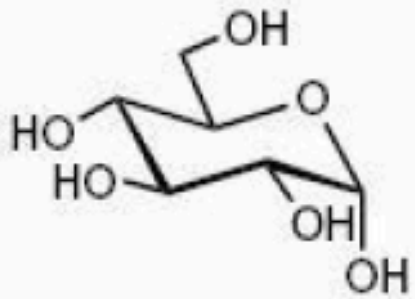
Polypeptides (proteins)

Protein databank

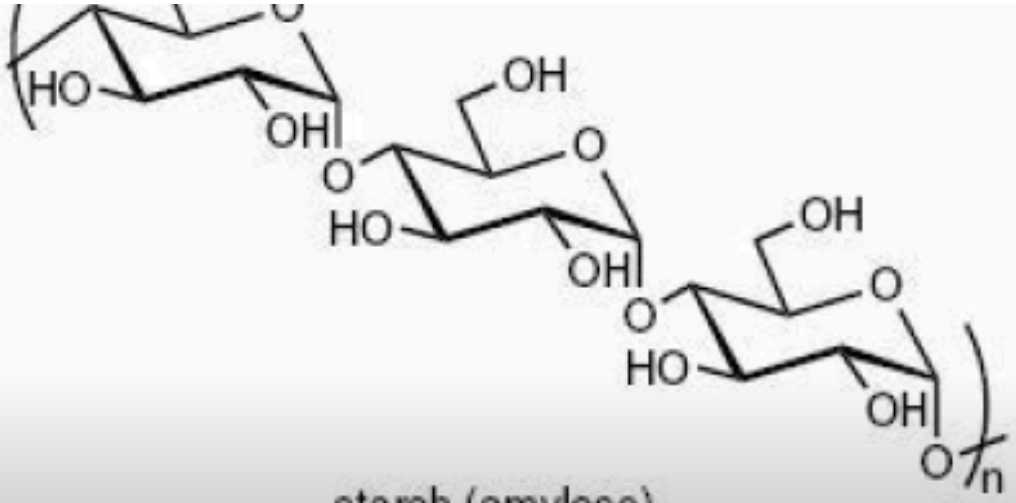
Molecular explorations
through biology and medicine

POLYSACCHARIDES

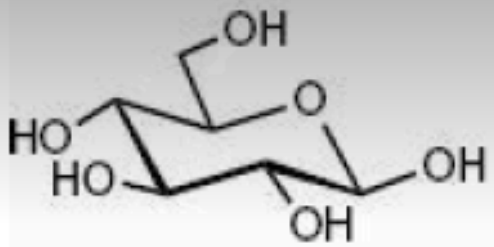
(carbohydrate polymers)



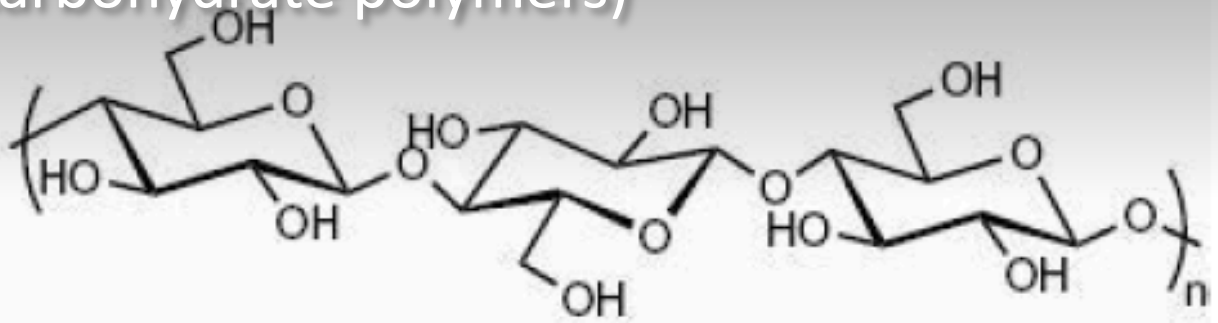
α-D-glucose



starch (amylose)

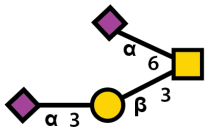


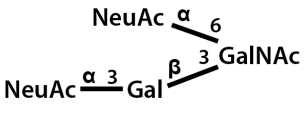
β-D-glucose

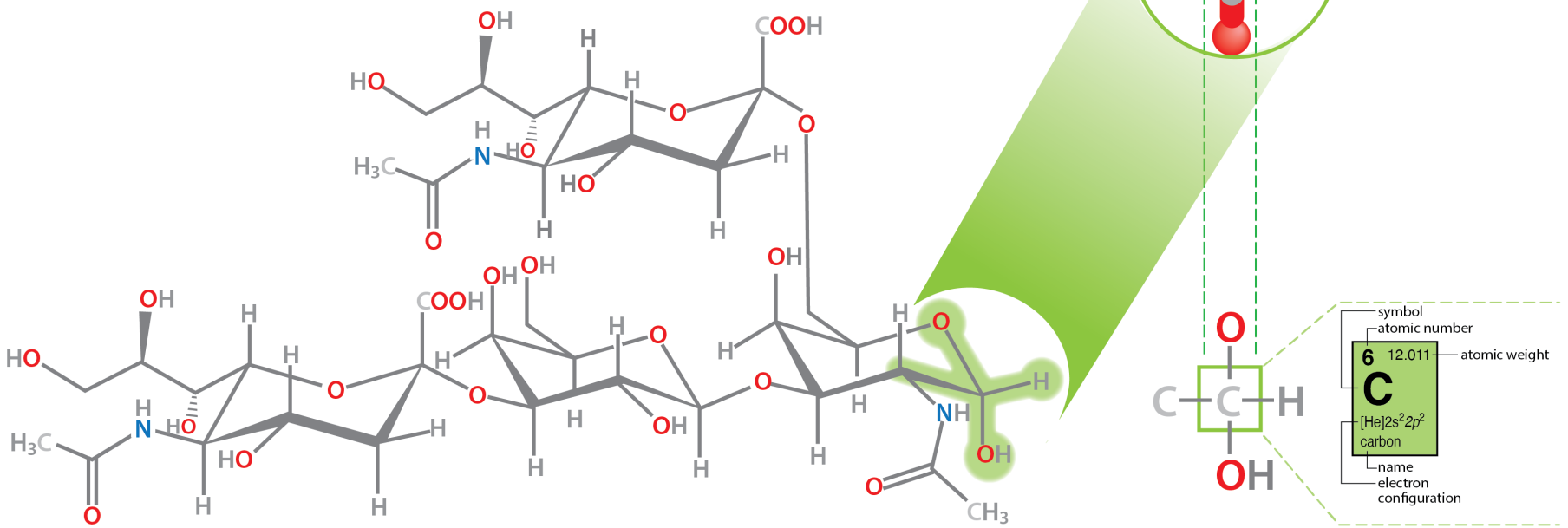


cellulose

SNFG	Abbr.	Systematic name
	GalNAc	2-Acetamido-2-deoxy-D-galactopyranose
	Gal	D-Galactopyranose
	NeuAc	5-Acetamido-3,5-dideoxy-D-glycero-D-galacto-non-2-ulopyranosonic acid

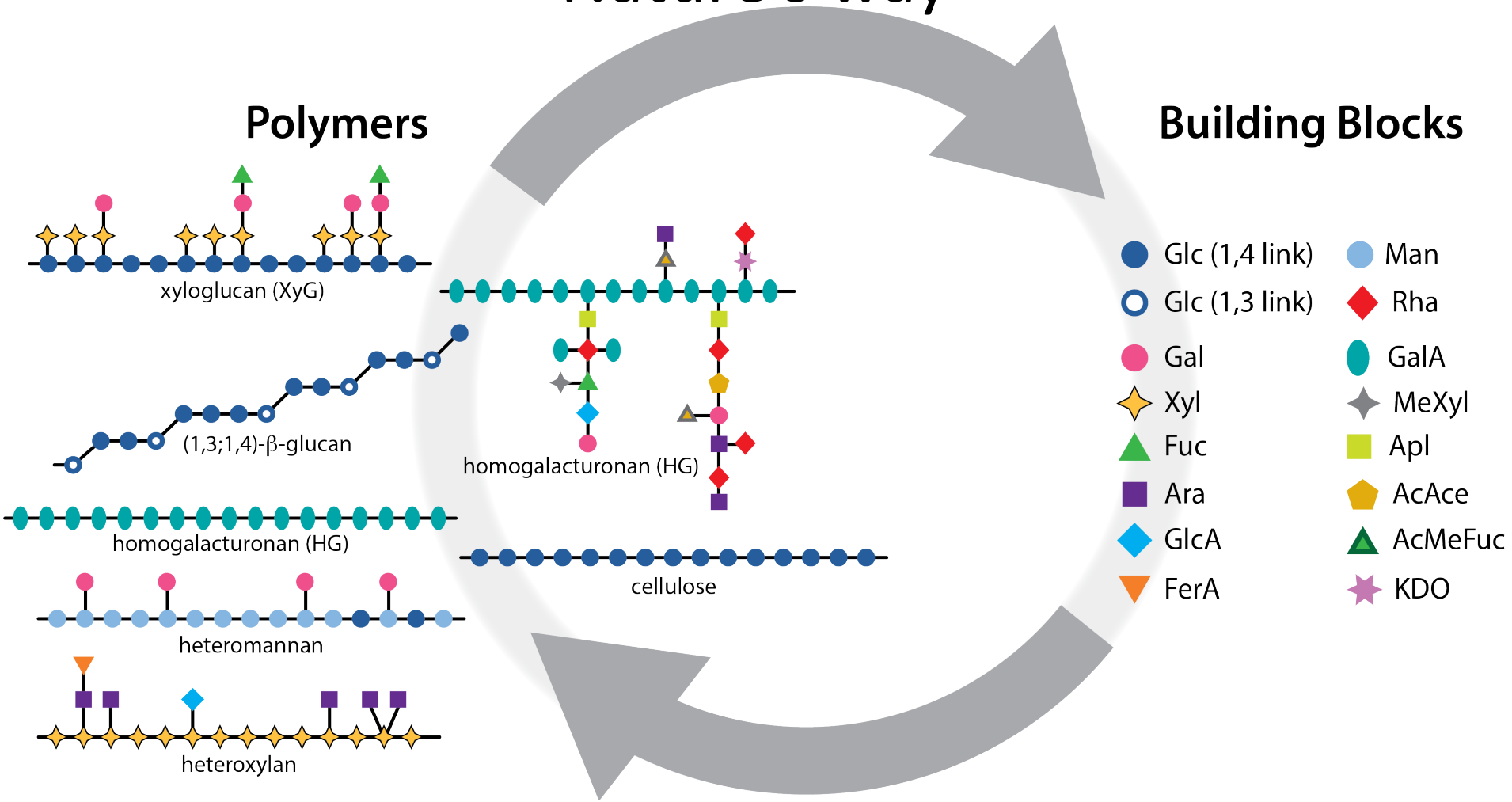


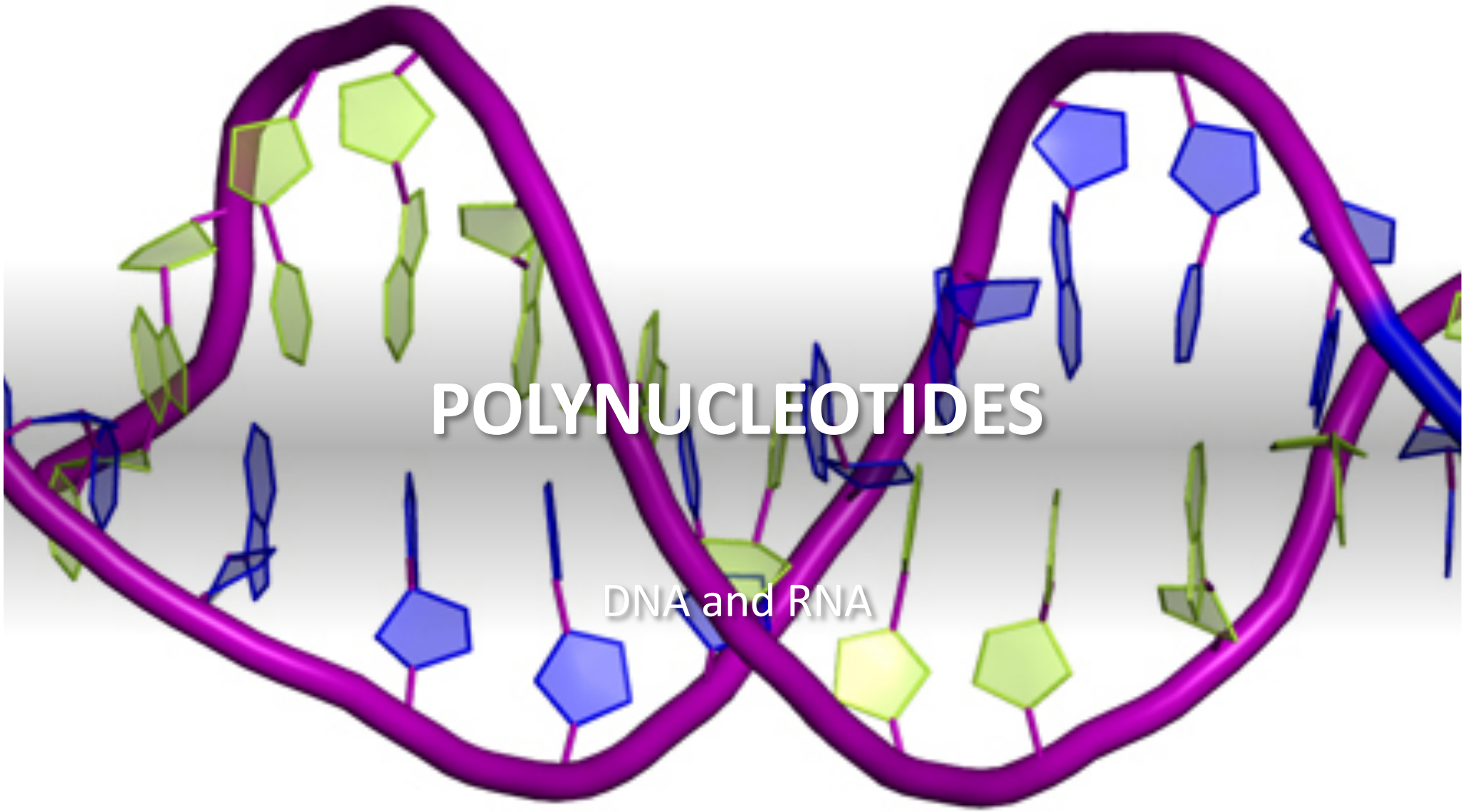




https://www.ncbi.nlm.nih.gov/glycans/snfg.html

Nature's way

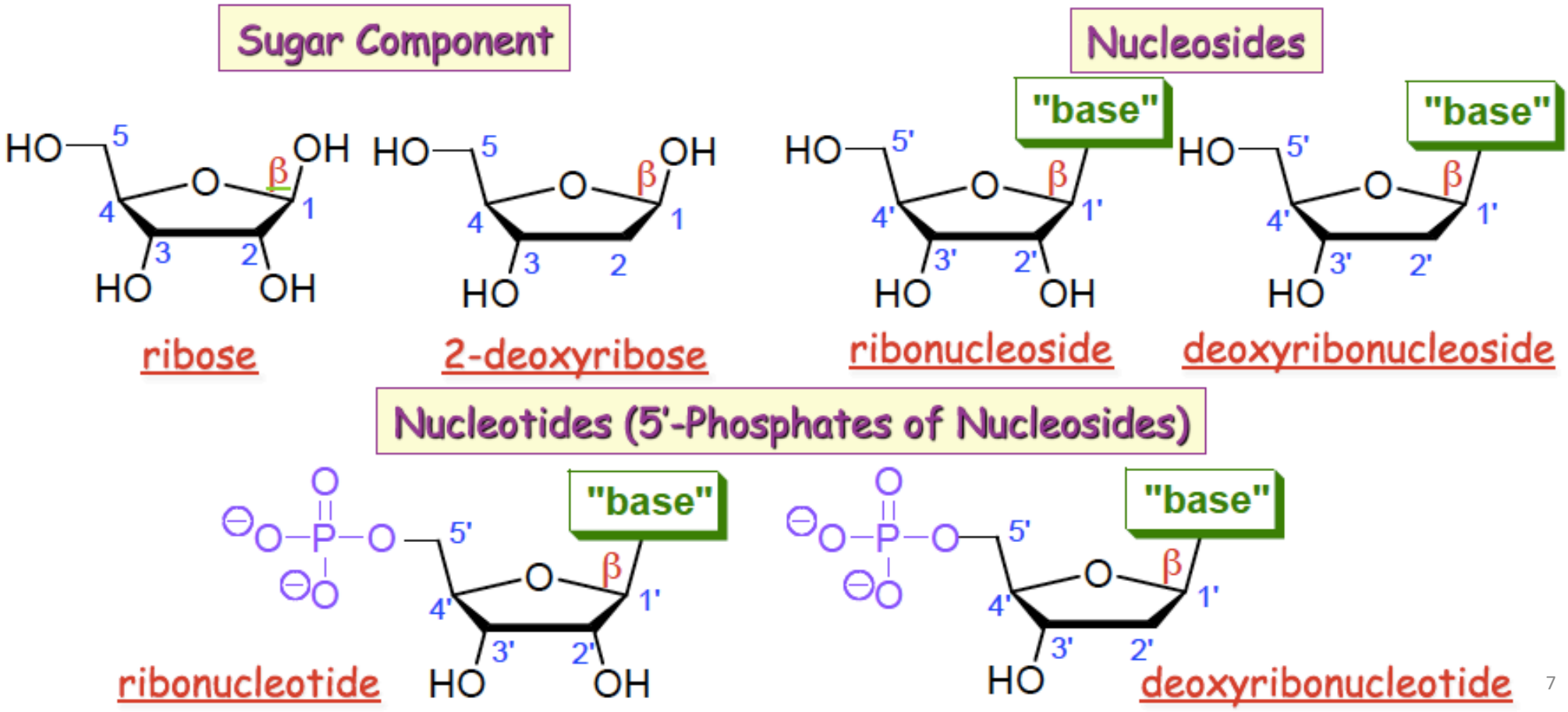

Polymers
xyloglucan (XyG)
(1,3;1,4)-β-glucan
homogalacturonan (HG)
homogalacturonan (HG)
cellulose
heteromannan
heteroxylan
Building Blocks
Glc (1,4 link)
Glc (1,3 link)
Gal
Xyl
Fuc
Ara
GlcA
FerA
Man
Rha
GalA
MeXyl
Apl
AcAce
AcMeFuc
KDO



POLYNUCLEOTIDES

DNA and RNA

Nucleosides and Nucleotides

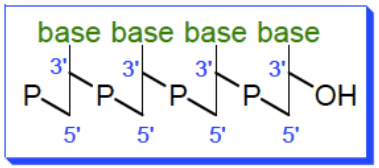


The Nucleic Acids, DNA and RNA, are Polymers

The Dickerson Dodecamer

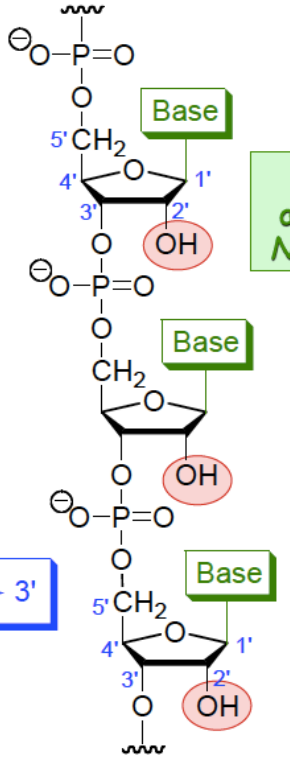
In the polymers (the nucleic acids DNA and RNA), the sugars are linked to each other through phosphoric acid groups attached to the 5′ position of one sugar and the 3′ position of the other.

This defines a direction to the backbone. The backbone bears a negative charge at physiological pH.



5′ ⟶ 3′

shorthand notation (5′ on left!)

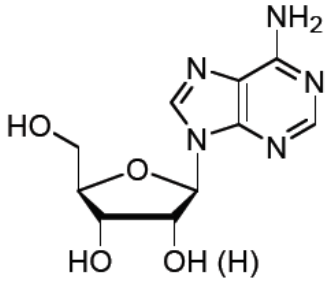


The "base" is attached via a β-*N*-glycosidic bond.

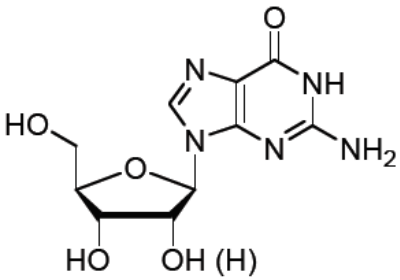
2′-OH
Present in RNA
Absent in DNA

The Nucleosides

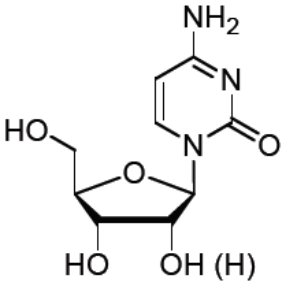
The Dickerson Dodecamer



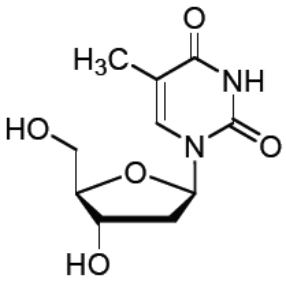
Adenosine (A)



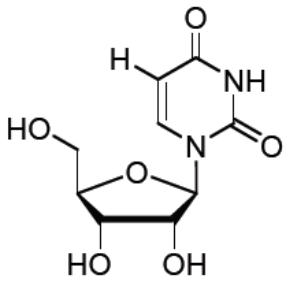
Guanosine (G)



Cytidine (C)



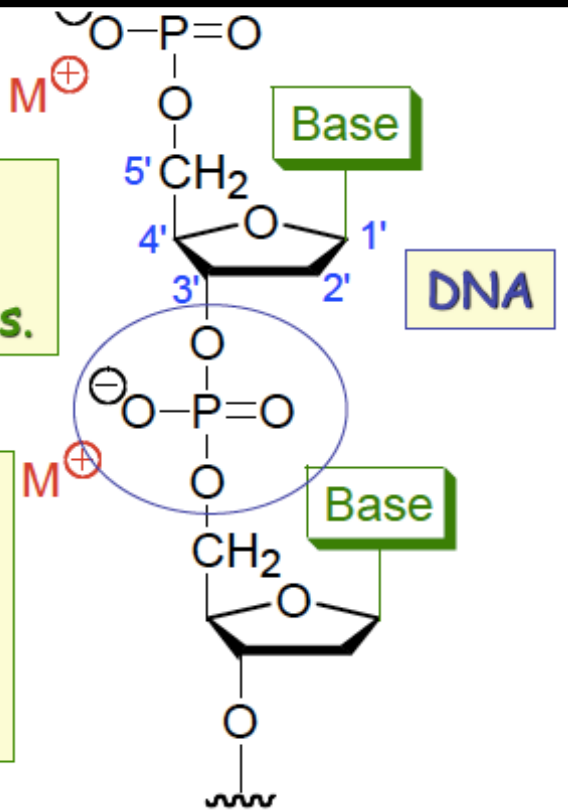
Thymidine (T)
(DNA only)



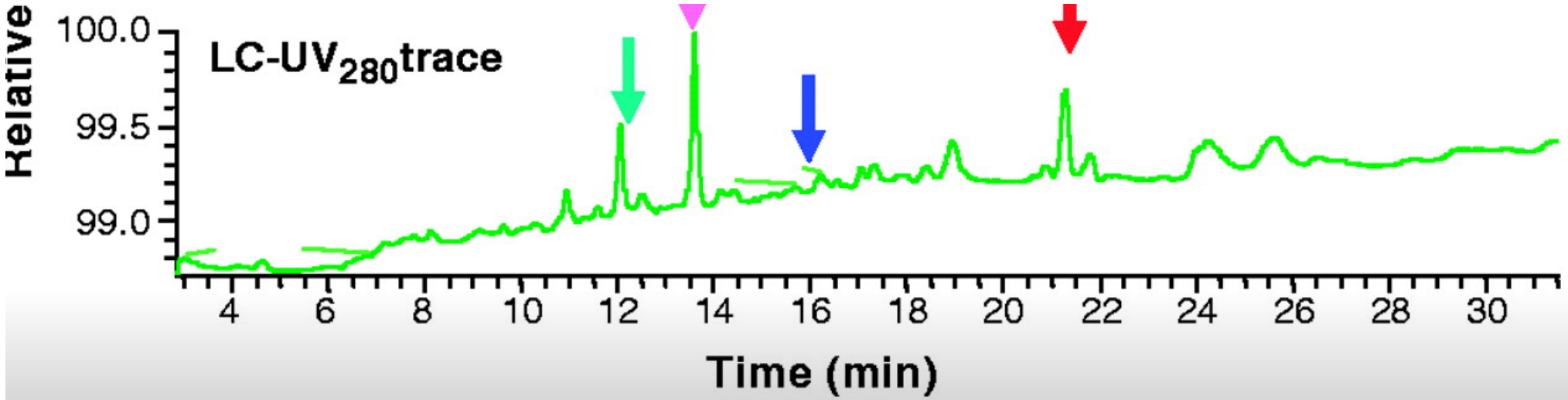
Uridine (U)
(RNA only)

Phospho mono esters can be hydrolyzed fairly easily under basic conditions, whereas Phospho di esters are resistant to basic hydrolysis.

Consequently, the DNA backbone has remarkable stability toward hydrolysis. (It has been estimated that the half-life for the hydrolysis of a phosphodiester at pH 7 and 37°C is 80×10^6 years!). DNA is stable, even on a prehistoric time scale!



The Phosphodiester Linkage Has Special Stability

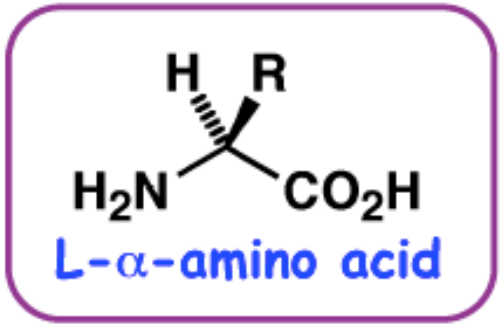


POLYPEPTIDES

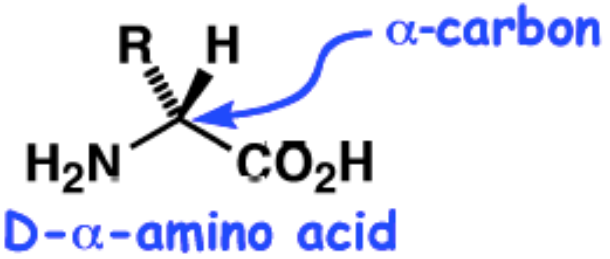
Proteins

MVRVPVPQLQPQNPSQQQPQEQVPLVQ
QQQFPGQQQPFPPQQPYPQPQPFPSQQ
PYLQLQPFPQPQLPYPQPQLPYPQPQL
PYPQPQPFRPQQPYPQSQPQYSQPQQP
ISQQQQQQQQQQQKQQQQQQQQILQQ
ILQQQLIPCRDVVLQQHSIAYGSSQVL
QQSTYQLVQQLCCQQLWQIPEQSRCQA

Proteins are polymers (i.e., made of many parts).
Protein backbones are poly(amino acids): a linear sequence of α-amino acid building blocks that are joined together by peptide (amide) linkages. All amino acids (except glycine) have a stereogenic carbon (the α-carbon). Proteins are composed of [D or L] amino acids.



L-α-amino acid

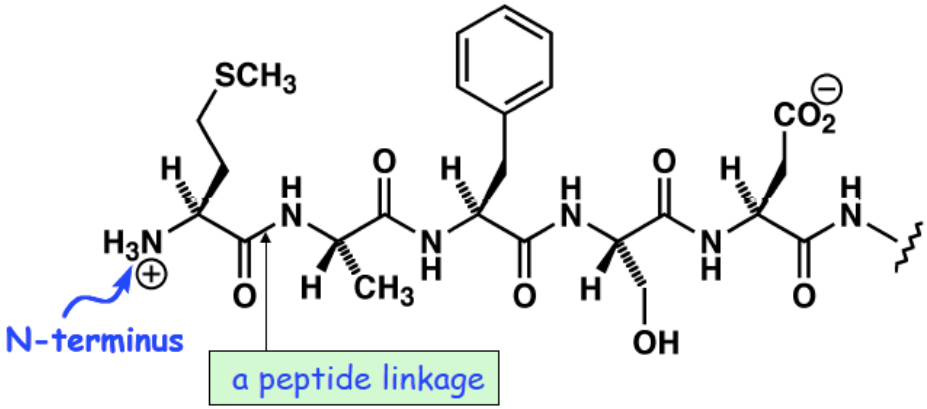


D-α-amino acid

Amino Acids and Proteins

A Representative Peptide Sequence

Here's a portion of the N-terminal part of a linear polypeptide chain. Notice that the sequence is directional i.,e., "N to *C*". To be more specific, the "N to *C*" direction of a peptide is found by locating the atom sequence nitrogen→C_α→carbonyl carbon.



Met-Ala-Phe-Ser-Asp-

M A F S D

The Aliphatic Amino Acids

Proteins are composed of 20 different α-amino acids. All of the structural and functional diversity among proteins can be attributed to the chemical differences found in this set of building blocks. Except for proline, all naturally occurring α-amino acids are [1° 2° 3°] amines. What makes them different is the side chains (-R) they bear.

The aliphatic α-amino acids

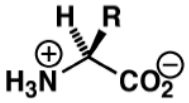
Glycine (Gly, G) —H

Alanine (Ala, A) —CH_3

Valine (Val, V) —$CH(CH_3)_2$

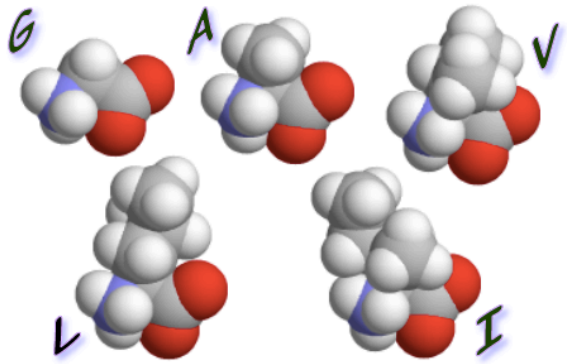
Leucine (Leu, L) —$CH_2CH(CH_3)_2$

Isoleucine (Ile, I) —$CH(CH_3)CH_2CH_3$



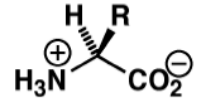
- ✓Branched
- ✓Compact
- ✓Bulky
- ✓Hydrophobic

Ideal suited for: packing & filling space

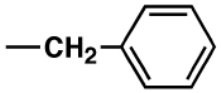


The Aromatic & Heterocyclic Amino Acids

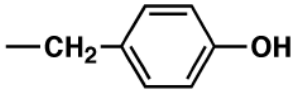
The aromatic α-amino acids



Phenylalanine (Phe, F)

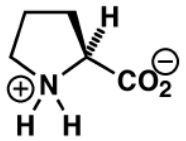


Tyrosine (Tyr, Y)

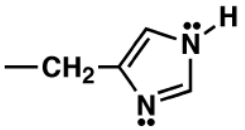


The heterocyclic α-amino acids

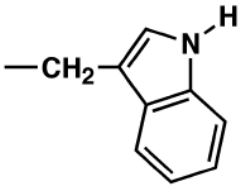
Proline (Pro, P)



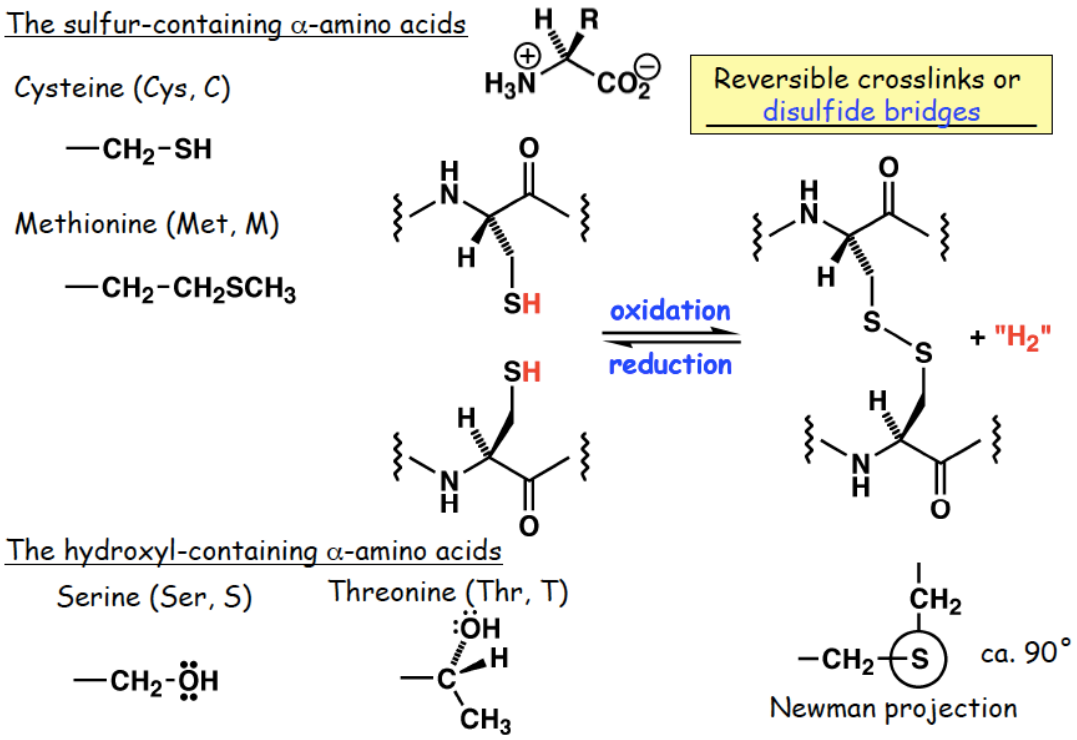
Histidine (His, H)



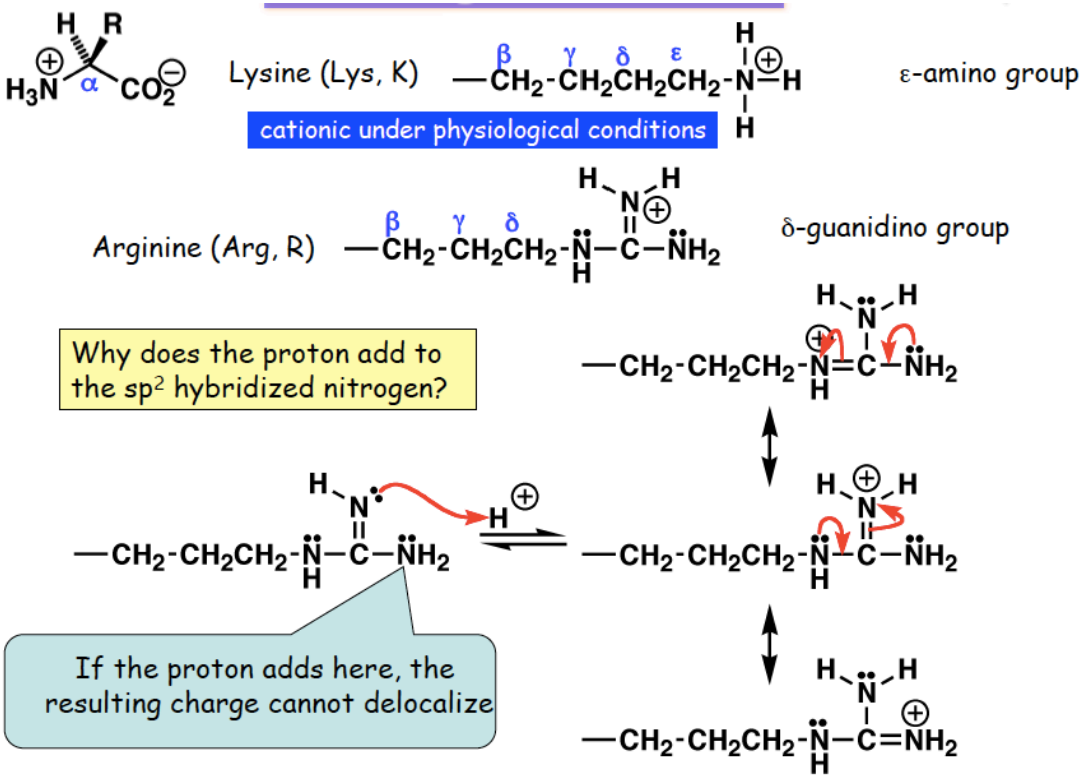
Tryptophan (Trp, W)



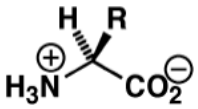
Sulfur and Hydroxyl-Containing Amino Acids



Amino Acids with Positively Charged Groups

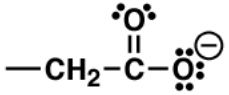


Amino Acids with Negative Charges and Their Derivatives



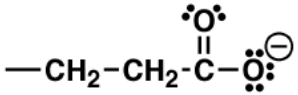
The acidic α-amino acids

Aspartate (Asp, D)



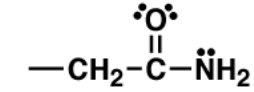
anionic under physiological conditions

Glutamate (Glu, E)

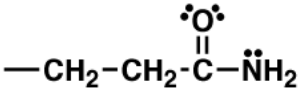


α-amino acids bearing primary amides

Asparagine (Asn, N)

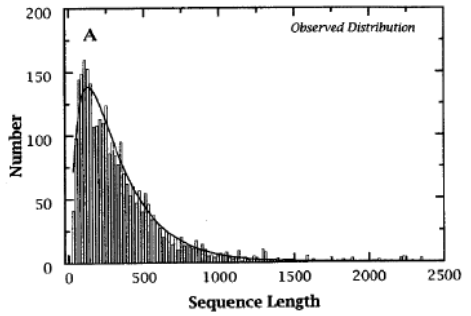


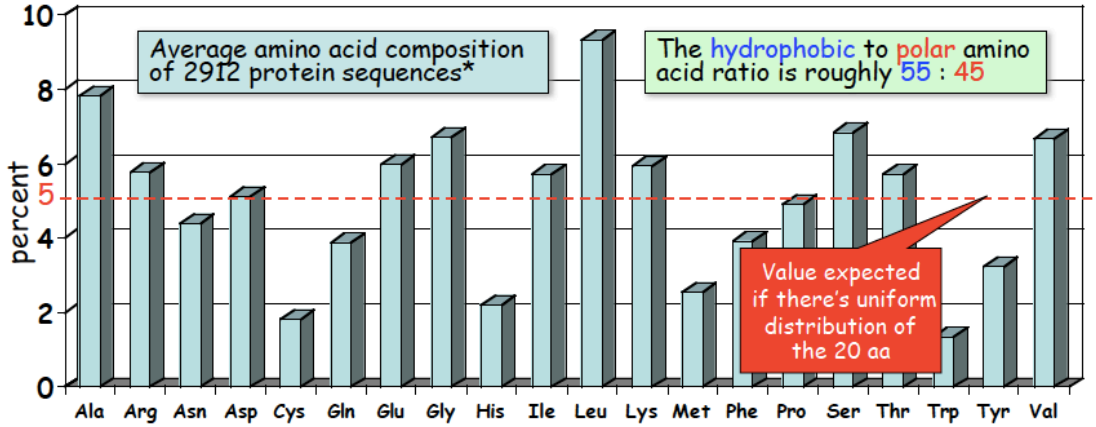
Glutamine (Gln, Q)



The Protein Recipe of Life

How many amino acids make up a typical protein polymer? The plot at the right shows the size distribution of a sample of 2912 protein sequences.* The mean sequence length for this sample is 319 amino acids.





*see: S. H. White "Global Statistics of Protein Sequences" Annu. Rev. Biophys. Biomol. Struct. (1994) 23, 407-39

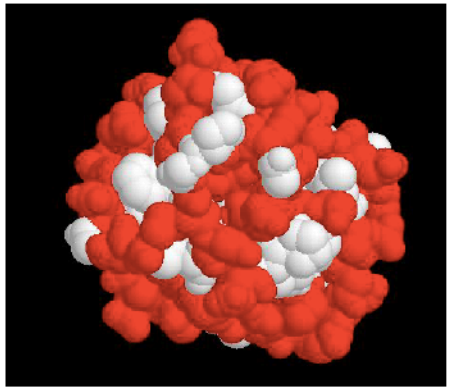
Protein Chains Fold into Unique 3D Structures

Factors that drive protein organization:

- The hydrophobic amino acids are buried in the core, excluding water.
- The surface is predominately hydrophilic amino acids that help maintain solubility.
- In the interior, the chain is arranged into secondary structures to "neutralize" the polar atoms through H-bonds.

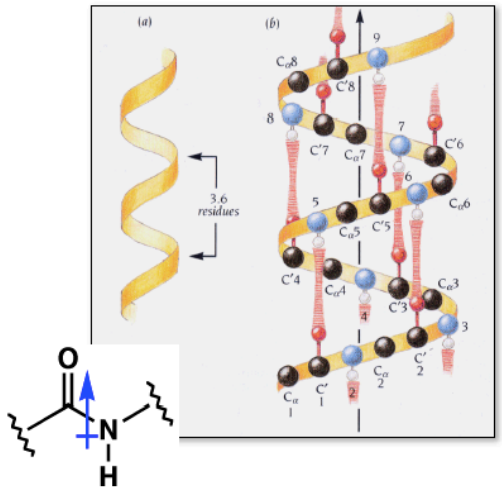
globular proteins are compact and densely packed

amphiphilic molecules

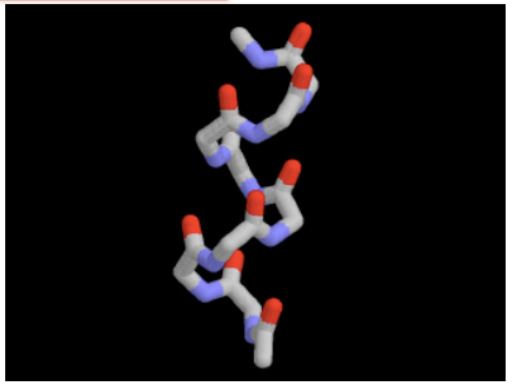


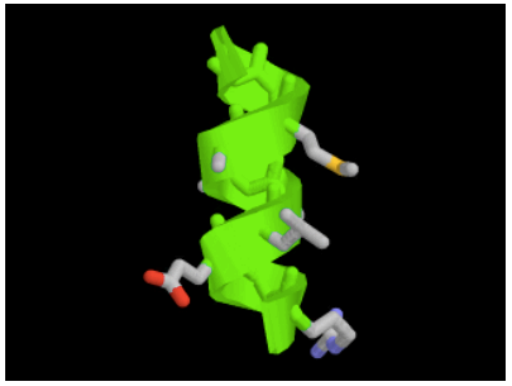
Two renderings of the protein Lysozyme

Secondary Structures – The alpha-Helix

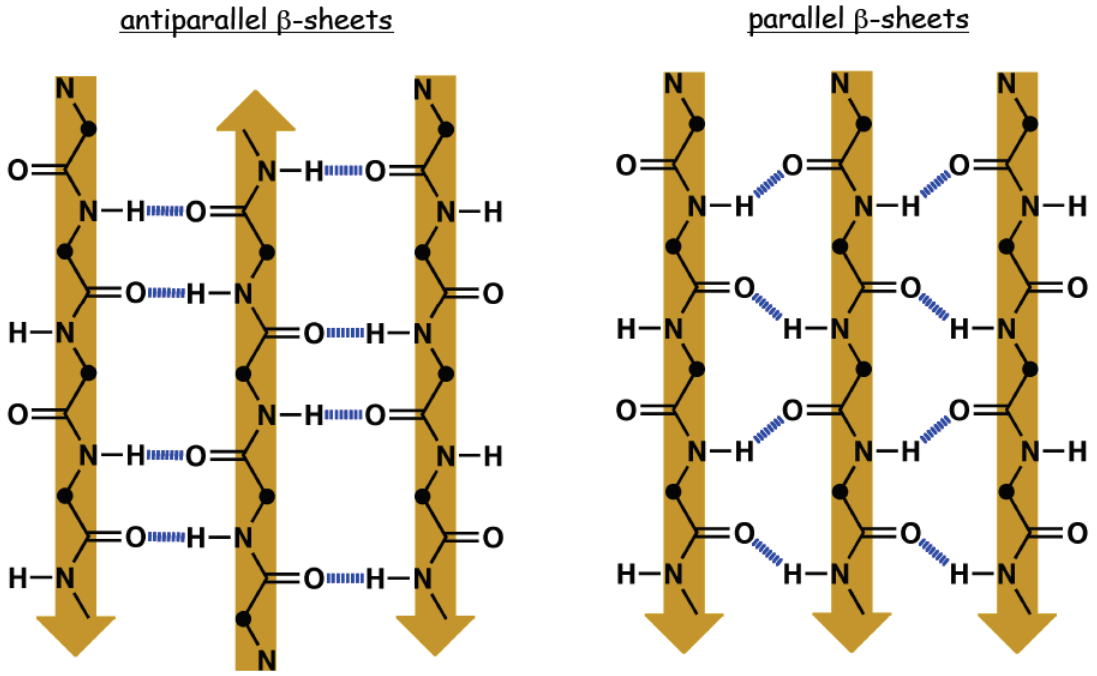


The α helix is one of the major elements of secondary structure in proteins. Backbone N and O atoms are hydrogen bonded to one another. The helix has a large dipole moment aligned parallel to the helical axis along the N to C direction. The side chains are oriented radially outward.





Secondary Structures – The Beta Sheet



The Protein Data Bank