Session 2

BUILDING BLOCKS
Atoms

- Six basic elements
  - Hydrogen (H), Carbon (C), Nitrogen (N), Oxygen (O), Phosphorous (P), Sulfur (S)

<table>
<thead>
<tr>
<th>Atoms</th>
<th>H</th>
<th>C</th>
<th>N</th>
<th>O</th>
<th>P</th>
<th>S</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valency</td>
<td>1</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>Electronegativity</td>
<td>2.1</td>
<td>2.5</td>
<td>3.0</td>
<td>3.5</td>
<td>2.1</td>
<td>2.5</td>
</tr>
<tr>
<td>H. Sapien (%)</td>
<td>9.3</td>
<td>19.4</td>
<td>0.8</td>
<td>62.8</td>
<td>0.6</td>
<td>0.6</td>
</tr>
<tr>
<td>Bacteria (%)</td>
<td>9.9</td>
<td>12.14</td>
<td>3.0</td>
<td>73.7</td>
<td>0.6</td>
<td>0.3</td>
</tr>
<tr>
<td>Earth (%)</td>
<td>0.9</td>
<td>0.2</td>
<td>0.9</td>
<td>50</td>
<td>0.12</td>
<td>0.12</td>
</tr>
</tbody>
</table>

- Valence electrons determine the bonding
- Electronegativity determines the polarity
- Composition is similar between living organisms
Molecular Bonds

- Ionic bonds
- Covalent bonds
- Hydrogen bonds
- Van der Waal bonds
- Hydrophobic interactions
- Catch bonds

Water is polar because of unequal e⁻ sharing:

\[ \text{O} = 3.5 \text{ vs. } \text{H} = 2.1 \text{ (e⁻ negativity)} \]

Solvents and hydrocarbon chains are nonpolar:

\[ \text{C} = 2.5 \text{ vs. } \text{H} = 2.1 \]
All life is from 68 molecular building blocks. 

- Nucleic Acids: 8 nucleosides
  - Deoxyadenosine, Deoxycytidine, Deoxyguanosine, Deoxythymidine, Adenosine, Cytidine, Guanosine, Uridine

- Proteins: 20 amino acids
  - Fuc, Gal, Glc, GlcA, Man, GalNAc, GlcNAc, NeuAc, Xyl, Kdn, Kdo, Ara, Araf, Col, Frc, Galf, GalUA, GlcLA, Hep, Leg, ManUA, FucNAc, GalNAcUA, ManNAc, ManNAcUA, MurNAc, PerNAc, QuilNAc, Per, Pse, Rha, Tel, Fa, Gl, Gipl, Pk, Pi, Sc, Sphi, Stf

- Glycans: 32+ sugars
  - Fucose, Galactose, Glucose, Glucuronic Acid, Mannose, N-Acetylgalactosamine, N-Acetylgulosamine, Neuraminic Acid, Xylose, Noronic Acid, Octulosonic Acid, Arabinose, Arabinofuranose, Colltose, Fructose, Galactofuranose, Galacturonic Acid, Glucolactilic Acid, Heptose, Legiomanonic Acid, Mannuronic Acid, N-Acetylglucosamine, N-Acetylgalacturonic Acid, N-Acetylmannosamine, N-Acetylmannosaminuronic Acid, N-Acetylmuramic Acid, N-Acetylperosamine, N-Acetylquinovosamine, Porphamate, Pseudaminic Acid, Rhamnose, Talose

- Lipids: 8 types
  - Fatty Acyls, Glycerolipids, Glycerophospholipids, Polyketides, Prenol Lipids, Saccharolipids, Sphingolipids, Sterol Lipids

“From the construction, modification, and interaction of these components, the cell develops and functions.” – James Marth

J. Marth “A Unified Vision of the Building Blocks of Life” Nature Cell Biology, 2008, 10(9):1015-16
Nucleic Acids

DNA – genetic code
- Adenine, Guanine, Cytosine, Thymine

RNA – translation into proteins
- Adenine, Guanine, Cytosine, Uracil
- Messenger RNA (mRNA)
- Transfer RNA (tRNA)

Others...
- Ribosomal RNA (rRNA)
- Small interfering RNA (siRNA)
- Micro RNA (microRNA)
- Small nuclear RNA (snRNA)
Nucleoside Triphosphates

- **ATP – energy currency**
  - Adenosine triphosphate
  - ATP + H₂O → ADP + PO₄³⁻ + 7.3 kcal/mol
  - Actin & Myosin

- **GTP - regulatory**
  - Guanosine triphosphate
  - G-protein Signal Transduction
  - Microtubules
Proteins

- Individual amino acids are translated into long chains called polypeptides.
- **Peptide bond**: carboxyl + amino $\rightarrow$ CO–NH + H$_2$O

- **Residue** (R): each amino acid in a polypeptide.
- Residue sequence read from N-terminal to C-terminal.
- Final sequence will “fold up” into a 3-dimensional structure.
- Substitution of just one residue can change a protein’s structure-function relationship.
Central carbon plus amino group (-NH$_2$), carboxyl group (-COOH), hydrogen atom (-H), and distinct side chain (20 in total)

- Side chains determine shape and function

Met, M
Start of all proteins

Gly, G
Simplest Hydrophobic

Pro, P
Forms a rigid kink
### Hydrophobic Amino Acids

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Structure</th>
<th>Abbreviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alanine</td>
<td><img src="image1" alt="Alanine" /></td>
<td>Ala or A</td>
</tr>
<tr>
<td>Valine</td>
<td><img src="image2" alt="Valine" /></td>
<td>Val or V</td>
</tr>
<tr>
<td>Isoleucine</td>
<td><img src="image3" alt="Isoleucine" /></td>
<td>Ile or I</td>
</tr>
<tr>
<td>Leucine</td>
<td><img src="image4" alt="Leucine" /></td>
<td>Leu or L</td>
</tr>
<tr>
<td>Methionine</td>
<td><img src="image5" alt="Methionine" /></td>
<td>Met or M</td>
</tr>
<tr>
<td>Phenylalanine</td>
<td><img src="image6" alt="Phenylalanine" /></td>
<td>Phe or F</td>
</tr>
<tr>
<td>Tyrosine</td>
<td><img src="image7" alt="Tyrosine" /></td>
<td>Tyr or Y</td>
</tr>
<tr>
<td>Tryptophan</td>
<td><img src="image8" alt="Tryptophan" /></td>
<td>Trp or W</td>
</tr>
</tbody>
</table>

### Hydrophilic Amino Acids

#### Basic Amino Acids

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Structure</th>
<th>Abbreviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lysine</td>
<td><img src="image9" alt="Lysine" /></td>
<td>Lys or K</td>
</tr>
<tr>
<td>Arginine</td>
<td><img src="image10" alt="Arginine" /></td>
<td>Arg or R</td>
</tr>
<tr>
<td>Histidine</td>
<td><img src="image11" alt="Histidine" /></td>
<td>His or H</td>
</tr>
</tbody>
</table>

#### Acidic Amino Acids

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Structure</th>
<th>Abbreviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aspartate</td>
<td><img src="image12" alt="Aspartate" /></td>
<td>Asp or D</td>
</tr>
<tr>
<td>Serine</td>
<td><img src="image13" alt="Serine" /></td>
<td>Ser or S</td>
</tr>
<tr>
<td>Threonine</td>
<td><img src="image14" alt="Threonine" /></td>
<td>Thr or T</td>
</tr>
</tbody>
</table>

#### Polar Amino Acids with Uncharged R Groups

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Structure</th>
<th>Abbreviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glutamate</td>
<td><img src="image15" alt="Glutamate" /></td>
<td>Glu or E</td>
</tr>
<tr>
<td>Asparagine</td>
<td><img src="image16" alt="Asparagine" /></td>
<td>Asn or N</td>
</tr>
<tr>
<td>Glutamine</td>
<td><img src="image17" alt="Glutamine" /></td>
<td>Gln or Q</td>
</tr>
</tbody>
</table>

### Special Amino Acids

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Structure</th>
<th>Abbreviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cysteine</td>
<td><img src="image18" alt="Cysteine" /></td>
<td>Cys or C</td>
</tr>
<tr>
<td>Glycine</td>
<td><img src="image19" alt="Glycine" /></td>
<td>Gly or G</td>
</tr>
<tr>
<td>Proline</td>
<td><img src="image20" alt="Proline" /></td>
<td>Pro or P</td>
</tr>
</tbody>
</table>
**Amino Acid Sequence**

MCEEETTALVCDNGSGLCKAG
FAGDDAPRAVFPSIVGRPRHQ
GVMVGMGQDSYVGDEAQSK
RGILTLKYPIEHGIITNWDDME
KIWHHSFYNELRVAPEEHNPTLL
TEAPINPKANREKMTQIMFET
FNVPAMYVAIQAVLSLYASGRT
TGIVLDSGDGVTHNVPIYEGYA
LPHAIMRLDLAGRDLTDYLMKI
LTERGYSFVTTAEREIVRDIKEK
LCYVALDFENEMATAASSSSL
EKSYELPDGQVITIGNERFRCP
ETLFQPSFIGMESAGIHETTYN
SIMKCDIDIRKDLYANNVLSGG
TTMYPGIADRMQKEITALAPS
TMKIKIIAPPERKYSVWIGGSIL
ASLSTFQQMQWISKPEYDEAGP
SIVHRKCF

**SMOOTH MUSCLE ACTIN (ACTG2)**
Translation

- Start with Methionine (AUG)
- Build codon by codon
- End with UAA, UAG, UGA
Protein Folding

- Hydrophobic residues form inner core of protein
- Denaturing is loss of protein structure
- Post-translational modifications can change protein folding (...and change function)
Protein Structure

- **Primary** – amino acid sequence
- **Secondary** – sub-structures through hydrogen bonds
  - Alpha helix
  - Beta sheet
- **Tertiary** – overall shape of a single protein unit
- **Quaternary** – union of more than one protein units
FoldIt (Hw1)

- Developed at UW by Seth Cooper from David Baker’s Rosetta program
Glycans

- “Carbohydrates” – hydrated carbons, (C, H, O)
- Combinatorial:
  - Glycoproteins – transmembrane proteins, e.g. integrins
  - Proteoglycans – protein core with polysaccharide chains, e.g. ECM’s aggregan
  - Glycolipids – located in extracellular layer of the cell membrane
**Lipids**

- Phospholipids have nonpolar hydrocarbon tails and negatively charged phosphate heads
- In water, phospholipids cluster together to form bilayers
- At air-water interface, tails point out and you have soap bubbles
All Together Now...

Nucleic Acids (DNA and RNA)
- Deoxyadenosine, Deoxycytidine, Deoxyguanosine, Deoxythymidine,
- Adenosine, Cytidine, Guanosine, Uridine

Glycans
- Fucose, Galactose, Glucose, Glucuronic Acid, Mannose,
- N-Acetylgalactosamine, N-Acetylgalcosamine, Neuraminic Acid,
- Xylose, Nononic Acid, Octulosonic Acid, Arabinose,
- Arabinofuranose, Kolitose, Fructose, Galactofuranose,
- Galacturonic Acid, Gluconactilic Acid, Heptose,
- Legionaminic Acid, Mannuronic Acid, N-Acetylfucosamine,
- N-Acetylgalacturonic Acid, N-Acetylmannosamine,
- N-Acetylmannosaminuronic Acid, N-Acetylmuramic Acid,
- N-Acetylpersamine, N-Acetyquinovosamine,
- Perosamine, Pseudaminic Acid, Rhamnose, Talose

Proteins
- dA, dC, dG, dT, rA, rC, rG, rU
- Fuc, Gal, Glc, GlcA, Man, GalNAc, GlcNAc, NeuAc, Xyl, Kdo, Ara, Araf, Col, Frc, GaI, GalUA, GlcLA, Hep, Leg, ManUA, FucNAc, GalNAcUA, ManNAc, ManNAcUA, MurNAc, PerNAc, QuiNAc, Per, Pse, Rha, Tal
- Fa, Gl, GlpI, Pk, Pl, Scl, Sphi, Stl

Lipids
- Fatty Acyls, Glycerolipids, Glycerophospholipids,
- Polyketides, Frenoi Lipids, Saccharolipids,
- Sphingolipids, Sterol Lipids

Alanine, Arginine, Aspartic Acid, Asparagine,
- Cysteine, Glutamic Acid, Glutamine,
- Glycine, Histidine, Isoleucine, Leucine, Lysine,
- Methionine, Phenylalanine, Proline, Serine,
- Threonine, Tryptophan, Tyrosine, Valine
Questions?