Lecture 7
Thursday, January 17, 2008

1. Exploring Proteins

- Purification
- Sequencing
- Identifying proteins in vivo.
  - Proteomics
  - Laser Scanning Confocal Microscopy
  - Fluorescent proteins

2. Central Dogma

- DNA, RNA structures.
- Replication, transcription, and translation

Chapter Problems:

Chapter 2: 1, 2, 5, 6, 7, 8, 9, 11, 13. For problem 9, assuming the membrane is 25 Å thick, how many amino acid residues, forming a perfect $\alpha$-helix, are required to span that distance?

Chapter 3: 1, 4, 10, 11, 14, 15, 17, 18.

Chapter 4: 1, 3 (assume B-form helix), 5, 7, 10 (why would U be used in RNA but T in DNA, given this chemistry?), 11, 12, 13, 20, 22, 23 (why might the protein sequence be more conserved than the DNA sequence?)

Note that many of the problems in Chapters 3 and 4 are “thought” problems, meaning you will probably have no idea how to approach them at first. Still, you should think about them a bit and then check the answer in the back of the book. I am not interested so much in that you get them right without checking, but that you understand the thinking behind the answer.
1. Protein Purification
   separate proteins:
   - size
   - charge
   - polarity
   - function.
   - if have recombinant protein.
   - over-express our protein.
   - get 10's mg of protein from 1 L culture of bacteria.

- because we have gene, can add affinity tags.
  protease
  gene for my + "tag" protein.

  tags in common use
  His tag  His-H-H-H-H-H.
  - bind metal very tightly
- Use Ni^{2+} on column
  my prof. + His tag.

- FLAG binds to antibody
- sugar-binding tag = maltose binding protein (MBP)

2. Protein sequencing.
   after purification, know linear sequence of amino acids.
   - esp. if unknown.
   - also to characterize your protein.

- do in 2 stages.

1. Determine amino acid composition.
   - take purified protein
   - acid digest. place in 6 M HCl
     - boil / heat
     - 24 hrs.

   - breaks every peptide bond
- separate, identify & quantify each a.a. present.

- make phenyl hydantoin derivative.

\[ \begin{align*}
&\text{phenyl isothiocyanate} \\
\downarrow \\
&\text{phenyl hydantoin derivative} \quad \text{PTH aa}
\end{align*} \]

- advantage of derivative:
  - makes amino acid visible by UV/VIS spectroscopy; sensitive detection.

- makes same derivative with all a.a.s.; the only difference is R
- separate PTH aa's on reversed-phase column.

Chromatogram

- each peak = 1 amino acid.
  - length which one based on elution time
  - gives composition by integrating each peak.

- problem: composition is not uniquely diagnostic.

  TIM ITM MIT

- sequence is.
b. sequence the peptide/protein.

Edman degradation; gives us

AGD FGR...

1 PTH aq at a time from N-terminus.

\[
\begin{align*}
\text{A} & \quad \text{G} \quad \text{length } n^\prime \\
+ & \quad \text{N=C=S} \\
\downarrow \text{dilute HCl}
\end{align*}
\]

\[
\begin{align*}
\text{PTH-Ala} \quad + \quad & \quad \text{H}_2\text{N} \quad \text{length } n-1 \quad \text{length} \\
\end{align*}
\]

repeat

PTH-Gly

repeat

PTH-Asp, etc.
- get 1 amino acid at a time for ~5-7 repeats (rounds) of Edman

- side reactions occur that limit use of Edman

- can't sequence entire protein

- to get entire sequence, fragment peptide x sequence each fragment

  tools for fragmenting peptides

  1. cyanogen bromide CNBr

  fragments after Met

  ![Fragmentation Diagram]

  sequence N-terminal 5-7 res. each fragment.
2. iodozobenzene: fragments after Trp (W)

3. proteases
   - proteins that catalyze breaking peptide bonds.
     - can have high specificity — only cleave specific aa. sequence
     - or they can be promiscuous
       - pepsin → very broad

example:

  trypsin          cleaves after
  chymotrypsin     K, R
  aromatics (F, Y, W)
  also L, M
  pepsin           broad specificity
  V8 protease      

4. C-terminal peptidase.  cleaves 1 res.
  Carboxy peptidase A.  at C-term
Edman

Case A:
fragment: break in middle.

more modern techniques:
   - new paradigm in biochemistry.
   - identify & quantify simultaneously all proteins expressed in cell.
   - genomics.
   - transcriptomics.
   - glycomics.
   - metabolomics

What is normal?
How change with disease?
4. in vivo protein tracking.
   - often fuse protein gene to fluorescent protein.
   - green fluorescent protein (GFP)
   - reddish.
   - Geoff Stouse.

Quantum Dots.
Gold nanoparticles.

Chapt. 4. DNA, RNA & Flow of genetic information.
Examine Central Dogma.

DNA → RNA → Protein.
transcription translation.

1. nucleic acid structure.
   - already introduced 4 bases in DNA
- **Compare DNA and RNA.**
  a. RNA lacks T instead, it has Uracil (U)
  b. Sugars differ in DNA and RNA
2'-OH in RNA has huge effect on structure & chemistry
- DNA sugar more reduced than RNA sugar
- Structurally, alters the double helix

DNA
- B-form
  - 10 bases / 360°
  - 34 Å per turn
  - 10 Å / turn

RNA
- A-form
  - 11 bases / 360°
  - 25 Å

- Less pronounced major / minor grooves.
  - Narrower major, wider minor

Figure on p 785 text.
- in both cases, bases are inside sugars & phosphates on outside.

- chemically, 2'-OH good nucleophile.

- DNA & RNA can form other structures.
  - more for RNA:
    - i. hairpins in single strands.
    - ii. cruciform if 2 strands
    - iii. circular, twisted, esp. in DNA.

- DNA lacks 2'-OH → more "bendy"
FIGURE 11.10 DNA forms.