**Lecture 6: UV Absorption of Amino Acids & Protein Structure**

Assigned reading: Horton - 3.2B, 3.9-3.10. Nelson 5e - box 3.1, 3.4. **OLI:** Quiz: 2

**Key Terms:**

* UV Absorption
* Peptide/protein nomenclature
* Edman degradation
* Cyanogen bromide cleavage
* Chymotrypsin & trypsin cleavage
* Protein sequence determination.

**6A. UV Absorption Properties of Amino Acids**



Three aromatic amino acids absorb light in the ultraviolet range (UV).



The extinction coefficients (or molar absorption coefficients) of these amino acids are:



|  |  |  |
| --- | --- | --- |
|  | **Amino acid** | **Extinction Coefficient ε(λMAX)** |
| Trp | 5,500 M-1cm-1 (280 nm) |
| Tyr | 1,490 M-1cm-1 (274 nm~280nm) |
| Phe | 220 M-1cm-1 (257 nm) |

The amount of light absorbed by a solution of concentration [X] is given by the Beer-Lambert Law:



where

* A is the absorbance of the sample;
* I0 is the intensity of the incident light;
* I is the intensity of the light that leaves the sample.
* ε is the molar extinction coefficient at a specific wavelength, *e.g.* at λmax;
* [X] is the concentration of the absorbing species
* *l* is the path length (usually 1 cm).

*Therefore, given a known extinction coefficient it is possible to measure the concentration of a protein.*

**Calculation of molar extinction coefficients:** If a molecule contains a mixture of N different chromophores, the molar extinction coefficient can generally be calculated as the sum of the molar extinction coefficient for each absorbing group in the protein:



Therefore, the molar extinction coefficient for a protein can be calculated from its amino acid composition.

**Example**: A protein has two Tryptophan (Trp) residues and one Tyrosine (Tyr) what is its extinction coefficient?

**6B. Protein Structural Hierarchy:**

**1. Primary structure (1°):** The amino acid sequence.

**2. Secondary structure (2°):** Configuration of mainchain.

**3. Tertiary structure (3°):** Side chain packing in the 3-D structure.

**4. Quaternary structure (4°):** Association of subunits.

**6C. Primary Structure**



**Example:** Draw the structure of Gly-Ala-Ala

i) draw mainchain atoms first.

ii) add sidechains

**Sequencing Proteins**



We will focus on N-terminal sequencing using Edman degradation coupled with fragmentation of the peptide in the case of larger proteins.

**Edman Degradation:** Cleaves single amino acids from the amino-terminus. Producing:

i) The PTH derivative of the released amino acid can be identified.

ii) A peptide that is one residue shorter is produced, this can be treated with PITC to obtain the next residue, etc.

iii) Errors accumulate because the release of the PTH-amino acid is not 100%, limiting the length of sequence to ~75 residues:

**Fragmentation**: Because it is not possible to sequence proteins larger than ~75 residues it is necessary to fragment the protein to extend the sequence information. After cleavage, the individual peptide fragments are separated from each other and each is subject to N-terminal sequencing using the Edman degradation method.



**a.** **Cyanogen bromide** (CNBr) cleaves the peptide bond after Methionine residues. You do not need to know the mechanism, just the specificity of the reaction.



**b.** Chymotrypsin hydrolyzes the peptide bonds that follow large hydrophobic residues, *e.g.* Phenylalanine, Tyrosine, Tryptophan. You should remember this cleavage pattern.



**c.** Trypsin hydrolyzes the peptide bonds that follow positively charged residues, *e.g.* Lysine and Arginine. You should remember this cleavage pattern as well.



**Sequence Strategies:**

If only two fragments are produced by the cleavage reaction, then it is straightforward to reconstruct the sequence using the known amino terminal sequence of the original protein to determine which fragment is first.

i) Obtain sequence past cleavage point (X) ii) Cleave, purify fragments, sequence 2nd fragment.

However if the original protein is cleaved into three or more fragments, or they cannot be sequenced in their entire length, then it is not possible to determine the correct order of fragments using a single cleavage agent. Multiple overlapping fragments have to be used to determine the correct ordering:

**An example: Ala-Gly-Met-Ser-Thr-Gly-Val-Val-Lys-Gly-Ser-Ala-Phe-Leu**





Here I have assumed that **7 cycles** of Edman degradation are possible; note ~75 are more typical.

**A.** The first 7 cycles of Edman sequencing on the intact peptide released the following amino acids, in this order: Ala, Gly, Met, Ser, Thr, Gly, Val, therefore , the initial, amino-terminal sequence of the peptide is:



**B.** A **new** sample of the intact peptide was treated with *CNBr*. The two peptides (CNBr-1, CNBr-2) that were produced were isolated and *each* was subject to sequencing, giving the following sequences (*The last three residues (Ala-Phe-Leu )could not be determined due to limitation in the Edman degradation to 7 steps, these are highlighted gray*).



**C.** A new sample of the peptide was treated with *Trypsin*. The two peptides (Trp1, Trp2) that were produced were isolated and *each* was sequenced.



**Sequence alignment strategy:** Find overlaps between fragments obtained with different cleavage reagents & use these overlaps to correctly *order* the peptides obtained from one sequencing reaction.

The overlaps can be readily identified by finding a cleavage site in a peptide that would be cut by another cleavage reagent and then identifying the correct fragment based on the expected amino-terminal sequence.

A) The original amino-terminal sequence contains a Methionine (Met), so I’d look for a CNBr fragment beginning with Ser-Thr-Gly-Val, the information from that CNBr fragment extends the sequence:



B) The lysine indicates that there should be a trypsin fragment beginning with Gly, the sequence of that trypsin fragment allows the completion of the sequence:

