IIT Guwahati

Lecture 7

Course BT 631

Protein Structure function and Crystallography

Prof. Arun Goyal

Dept. of Biosciences and Bioengineering

Structure of Protein

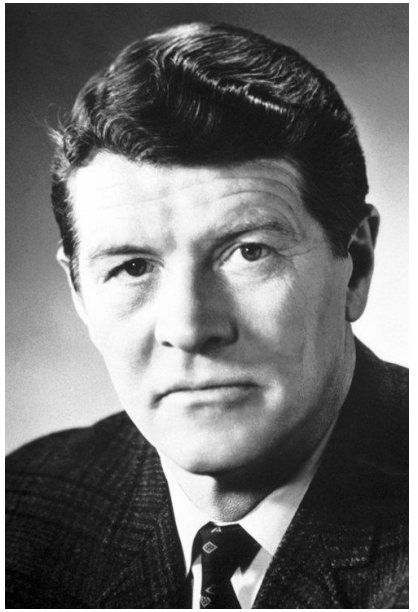
Protein exists in four different forms

- **A) Primary Structure**
- **B) Secondary Structure**
- **C) Tertiary Structure**
- **D) Quaternary Structure**

Structure of Protein

Primary Structure

- Role of primary structure in the process of protein folding was first determined by Christian Anfinsen on the Ribonuclease in 1961.
- It is defined as simple sequential arrangement of amino acid in a poly-peptide backbone linked through peptide bond.
- Primary structure acts as determinant of protein folding. e.g. In sickle cell anemia Glumate to Valine transition changes the shape of haemoglobin molecule.



Christian Anfinsen

Asn-Gly-Phe-Glu-Gln-Ala-Arg-Asp-Cys-Leu-Ile-Trp-Pro-Tyr-Ser-Met-Lys-Val-His

(Primary structure of protein)

How do you determine the primary structure of a protein?

1. **Determining amino acid composition of a protein.**

Steps

i. Hydrolysis (heat at 100°-110°C in 6M HCl for 24 h or longer)

ii. Separation by chromatography

iii. Quantitative analysis (color producing reagents e.g. ninhydrin)

2. **N-terminal amino acid analysis**

Steps

i. React the peptide with a reagent that selectively labels the terminal amino acid.
e.g. DFNB (2,4-Dinitro-1-fluorobenzene)

ii. Hydrolyze the protein.

iii. Determine the amino acid by chromatography and compare with standards.

How do you determine the primary structure of a protein?

3. C-terminal amino acid analysis (Carboxypeptidases)

- Carboxypeptidases are proteolytic enzymes that remove L-amino acids, one residue at a time, from the carboxyl-terminus of polypeptide chains, i.e., they are exoproteases.
- To determine the C-terminal amino acid sequence, the protein or peptide being analyzed is digested with carboxypeptidase and aliquots removed at timed intervals, and analyzed for the presence of free amino acids.
- The amount of each amino acid released is plotted against time, and the C-terminal sequence deduced from the relative rate of release of each amino acid.

How do you determine the primary structure of a protein?

Four carboxypeptidases have been used extensively to provide peptide and protein sequence data. These are:
carboxypeptidase A (EC 3.4.17.1) from bovine pancreas
carboxypeptidase B (EC 3.4.17.2) from porcine pancreas
carboxypeptidase C (EC 3.4.12.1) from orange leaves
carboxypeptidase Y (EC 3.4.16.1) from yeast

Historically, carboxypeptidases A and B were the first to be discovered and used for sequence determination.

Carboxypeptidase A releases most C-terminal amino acids, but will not cleave at Arginine, Proline, Hydroxyproline or Lysine residues.

The specificity of carboxypeptidase B is far more restricted, cleaving only C-terminal arginine and lysine residues.

Carboxypeptidases A and B, therefore, tended to be used together, but even so, exopeptidase activity was blocked when a proline was reached.

How do you determine the primary structure of a protein?

The isolation of carboxypeptidase C provided the combined specificity of carboxypeptidases A and B, but also cleaved at Proline residues *i.e.* it cleaves all C-terminal amino acids.

Carboxypeptidase Y (CPY, isolated from baker's yeast) has the same broad specificity as carboxypeptidase C, but because of its strong action on protein substrates and its ability to work in the presence of urea and detergents, it is nowadays the enzyme of choice for C-terminal sequence analysis.

Carboxypeptidase Y cleaves all L-amino acids one residue at a time from the C-terminal of polypeptide chains.