

Methods involved in protein sequencing

H Marc*

Department of Bioengineering, University of California, San Diego, USA

*Corresponding author. E-mail: hongmarc123@gmail.com

Received: 21-Nov-2022, Manuscript No: IJBB-22-84244; **Editorial assigned:** 23-Nov-2022, Pre QC No: IJBB-22-84244 (PQ); **Reviewed:** 07-Dec-2022, QC No: IJBB-22-84244; **Revised:** 16-Dec-2022, Manuscript No: IJBB-22-84244 (R); **Published:** 26-Dec-2022, DOI: 10.15651/IJBB.22.9.012

DECSRIPTION

Proteins are found in every cell and are required for every biological process; however, protein structure is extremely complex. Determining a protein's structure entails first protein sequencing, determining the amino acid sequences of its constituent peptides, as well as determining what conformation it adopts and whether it is complexed with any non-peptide molecules. Discovering the structures and functions of proteins in living organisms is an important tool for understanding cellular processes and makes it easier to develop drugs that target specific metabolic pathways. Mass spectrometry and the Edman degradation reaction are the two most common direct methods of protein sequencing. If the DNA or mRNA sequence encoding the protein is known, it is also possible to generate an amino acid sequence from it.

Hydrolysis

A sample of the protein is hydrolyzed by heating it in hydrochloric acid for 24 hours or longer to 100-110 degrees Celsius. Proteins with a large number of bulky hydrophobic groups may require longer heating time. These conditions, however, are so intense that some amino acids (serine, threonine, tyrosine, tryptophan, glutamine, and cystine) are degraded. To get around this, Biochemistry Online suggests heating different samples for different amounts of time, analysing each resulting solution, and extrapolating back to zero hydrolysis time a number of reagents to prevent or reduce degradation, including thiol reagents or phenol to protect tryptophan and tyrosine from chlorine attack, as well as pre-oxidising cysteine and determining the extent of amide hydrolysis by measuring the amount of ammonia produced.

Separation

Ion-exchange chromatography or hydrophobic interaction chromatography can be used to separate the amino acids. The former is demonstrated by using sulfonated polystyrene as a matrix, adding amino acids in an acid

solution, and passing a buffer of steadily increasing pH through the column. When the pH reaches their respective isoelectric points, amino acids will be eluted. The latter technique can be used by employing reversed-phase chromatography. Many commercially available C8 and C18 silica columns have successfully separated amino acids in solution in under 40 minutes using an optimised elution gradient. After the amino acids have been separated, their quantities are calculated by adding a reagent that produces a coloured derivative. If the amount of amino acids is greater than 10 nmol, ninhydrin can be used; it produces a yellow colour when reacted with proline and a vivid blue when reacted with other amino acids. The amino acid concentration is proportional to the absorbance of the resulting solution. Fluorescamine can be used as a marker in very small amounts, as low as 10 pmol. When it reacts with an amino acid, it forms a fluorescent derivative.

Protein Sequence Prediction Using DNA and RNA Sequences

The amino acid sequence of a protein can also be indirectly deduced from its mRNA or, in organisms lacking introns (e.g., prokaryotes), from the DNA that codes for the protein. If the gene sequence is already known, this is all very simple. However, the DNA sequence of a newly isolated protein is rarely known, so if this method is to be used, it must be discovered in some way. One method is to sequence a short section of the protein, perhaps 15 amino acids long, using one of the methods described above, and then use this sequence to generate a complementary marker for the protein's RNA. This can then be used to isolate the mRNA coding for the protein, which can then be replicated in a polymerase chain reaction to yield a significant amount of DNA, which can then be easily sequenced. The protein's amino acid sequence can then be deduced from this. However, the possibility of amino acid removal after the mRNA has been translated must be considered.