INTRODUCTION

Proteomics is something new in the field of biotechnology. It is basically the study of the proteome, the collective body of proteins made y a person's cells and tissues.

Since it is proteins, and to a much lesser extent, other types of biological molecules that are directly involved in both normal and disease-associated biochemical processes, a more complete understanding of the disease may be gained by directly looking at the proteins present within a diseased cell or tissue and this is achieved through the study of the proteome, Proteomics. For, Proteomics, we need 2-D electrophoresis equipment ot separate the proteins, mass spectrometry to identify them and x-ray crystallography to know more of the structure and function of the proteins. These equipments are essential in the study of proteomics.

FROM THE GENOME TO THE PROTEOME

Genomics has provided a vast amount of information linking gene activity with disease. It is now recognized that gene sequence information and pattern of gene activity in a cell do not provide a complete and accurate profile of a protein's abundance or its final structure and state of activity.

The day of spotlight of the human genome is now coming to an end. Researchers are now concentrating on the human proteome, the collective body of all the proteins made by a person's cells and tissues. The genome- the full set of information in the body-contains only the recipes for making proteins; it is the proteins that constitute the bricks and mortar of cells and that do most of the work. Moreover it is the proteins that distinguish the various types of cells: although all cells have essentially the same genome, they can vary in which genes are active and thus in which proteins are made. Likewise diseased cells often produce proteins that healthy cells don't and vice versa.

Proteome research permits the discovery of new protein markers for diagnostic purposes and of novel molecular targets for drug discovery.

PROTEINS

All living things contain proteins. The structure of a cell is largely built of proteins. Proteins are complex, three-dimensional substances composed of one or more long, folded polypeptide chains. These chains, in turn, consist of small chemical units called amino acids. There are twenty kinds of amino acids involved in protein production, and any number of them may be linked in any order to form the polypeptide chain.

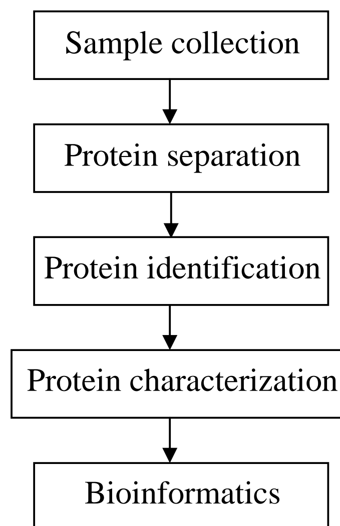
The order of the amino acids in the polypeptide chain is decided by the information contained in DNA structure of the cell's genes. Following this translation, most proteins are chemically changed through post-translation modification (PTM), mainly through the addition of carbohydrate and phosphate groups. Such modification plays an important role in modulating the function of many proteins but the genes do not code it.

As a consequence, the information from a single gene can encode as many as fifty different protein species. It is clear that genomic information often does not provide an accurate profile of protein abundance, structure and activity.

PROTEOMICS: A DESCRIPTION OF THE METHODOLOGY

The exact definition of proteomics varies depending on whom you ask, but most of the scientists agree that it can be broken into three main activities: identifying all the proteins made in a given cell, tissue or organism; determining how these proteins join forces to form networks akin to electrical circuits; and outlining the precise three-dimensional structure of the proteins in an effort to find their Achilles' heels-that is, where drugs might turn their activity on or off. Though the task seems straightforward, it is not as simple as it seems.

The critical pathway of proteome research includes:



These are the major steps involved in the proteome science studies. Each of these are described in detail.

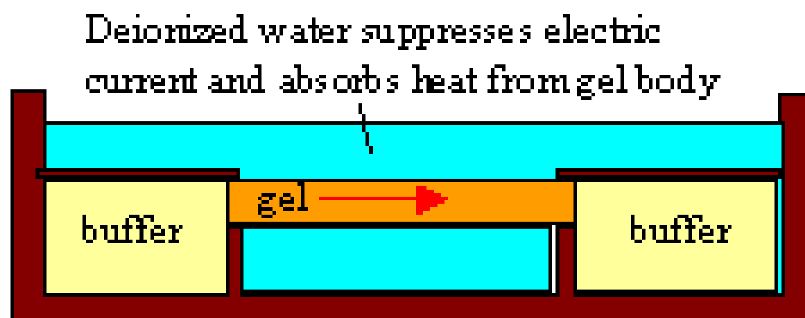
SAMPLE COLLECTION, HANDLING AND STORAGE

Access to relevant body fluid and tissue samples is fundamental to proteome research. Proteome sciences has collected a large bank of clinical samples for its cancer, neurological disease, cardiovascular disease, transplant rejection and diabetes research areas through its ongoing access to the leading hospitals where its collaborative scientists practice. All clinical samples are pathology-authenticated and are accompanied by details medical records to allow the correlation of proteome changes with disease pathology.

PROTEIN SEPARATION

The sample we have collected will have many proteins included in it. We need to separate these in order to study them. One of the main technologies used is two-dimensional gel electrophoresis.

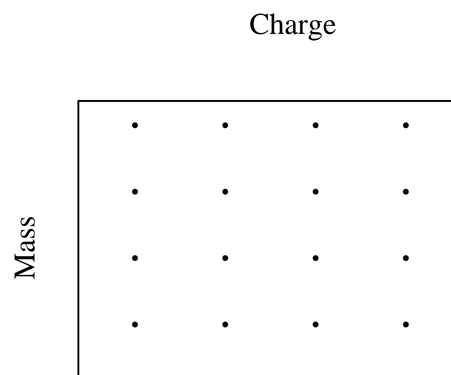
Electrophoresis is a technique used in laboratory that results in separation of charged particles and proteins in general are charged particles. Electrophoresis may be in general defined as the movement of a solid phase (the proteins in sample) with respect to a liquid (the buffer solution). The main function of the buffer is to carry the current and to keep the pH of the solution constant during migration. A solid substance called the medium supports buffer solution.



Here we use a gel as the substrate. A common material is agarose which is prepared from common seaweed. Purified agarose is in powdered form, and insoluble in water at room temperature, but is soluble in boiling water. When it starts to cool, it undergoes polymerization. The polymers crosslink and form the gel. If more agarose is added, the gel will become more firm.

While solution is still hot, we pour it into a mould called casting tray so that it will assume the shape we want. For setting, gel body is immersed in deionised water. Deionised water, an insulator, prevents massive heat generation. Much higher voltage, such as 280 volts can be applied to derive rapid sample migration.

Scientists add a mixture of proteins to an edge of the gel. An electric field is applied across the gel. The gel is in the form of a mesh network. In two-dimensional electrophoresis, separation is done according to mass in one direction an according to electric field in the perpendicular direction. Each protein has an individual mass and charge. So they will separate out as individual dots in the gel. Researchers can then isolate each of these proteins for further analysis.

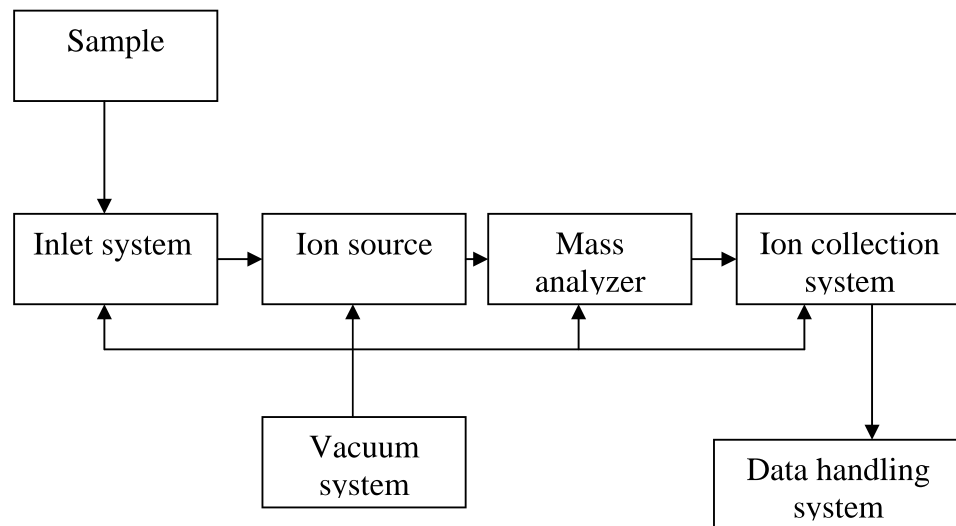


Two-Dimensional gels

Gels separate proteins by charge and mass

PROTEIN IDENTIFICATION

Mass spectrometry is a method of protein identification. The instrument used is a mass spectrometer. A mass spectrometer is an apparatus that produces a stream of charged particles from the substance being analyzed, separates the ions into a spectrum according to their mass-to-charge ratios, and determines the relative abundance of each type of ion present.

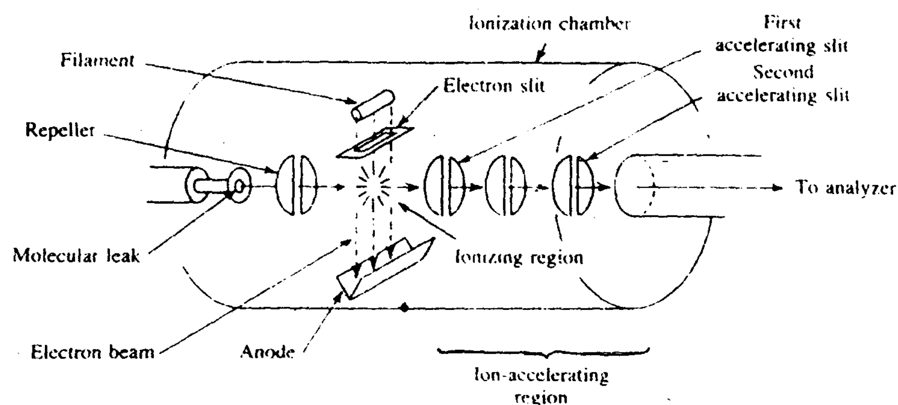


Components of a mass spectrometer

Functionally, all mass spectrometers perform three basic tasks: (1) creating ion fragments from sample, (2) sorting ions according to mass-to-charge ratio (3) measuring relative abundance of ion fragments of each mass.

Once the ions are formed, they can be sorted based on their energy, momentum or velocity. A measurement of any two of these gives mass-to-charge ratio. Conventionally the method is to use energy and momentum: accelerating ions in an electric field and dispersing them in a magnetic field based on their momentum.

There are various ionization methods of which electron-impact ionization is most highly developed and used.



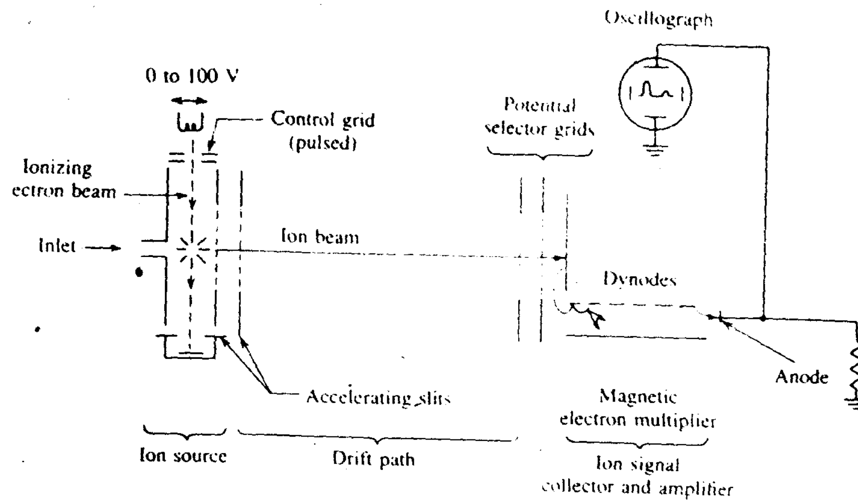
The function of mass analyzer is o separate ions according to their different mass-to-charge ratio. Ion collection systems too are of different types. One of them is a Faraday cup collector.

It is essential that the whole system must be maintained in vacuum conditions. Most systems use a combination of oil diffusion pumps to maintain a high vacuum.

Mass spectrometers are of many types. One of the most prominent ones is the time-of-flight mass spectrometer.

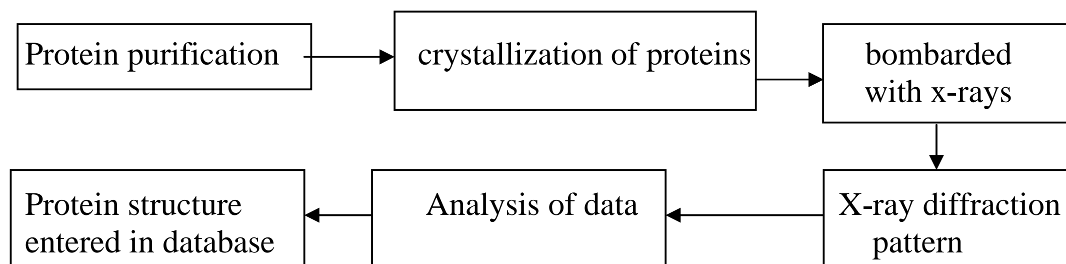
Here the energy and velocity of the ionized particles are made use of to separate them. The time-of-flight (TOF) mass spectrometer works in a pulsed rather than a continuous mode. Here the measurement of mass/charge ratio is determined by measuring energy and velocity of the particles.

The essential principle of TOF mass spectrometry is that if ions with different masses are accelerated to the same kinetic energy from the same point, each ion acquires a characteristic velocity which depends on its m/z ratio.



PROTEIN CHARACTERISATION

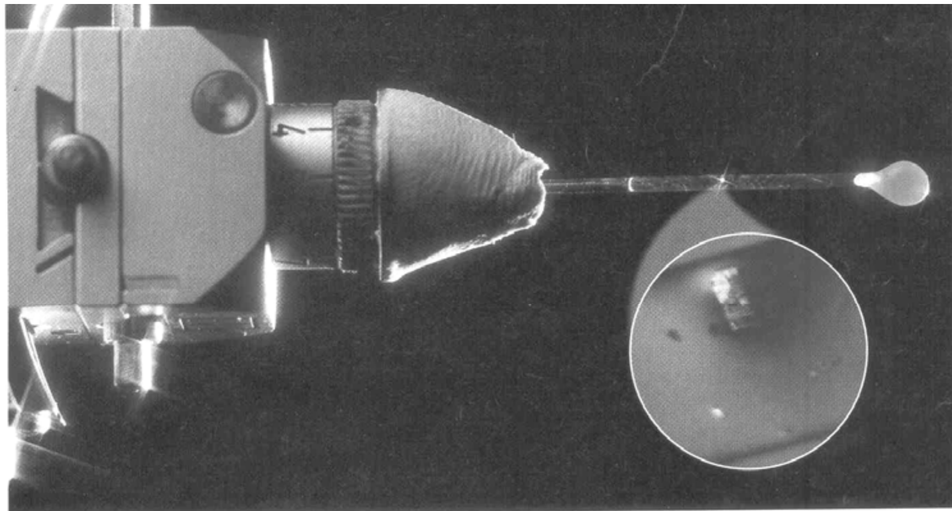
In this stage we need to define the function of each protein. This is a huge task. For this, we need to know about the structure of each protein and how these relate with other proteins. We can only obtain data by trial and error method.



The proteins have to be purified. Each protein molecule associates with others and forms a supersaturated solution, which differ from protein to protein. Once crystals are formed, they are frozen by immersing in cryogenic liquid or exposure to cryogenic gas to prevent ice lattice formation.

These crystals are used for x-ray crystallization. X-rays have played an important role in science and in our lives in the past century. X-ray crystallography allows us to investigate the structure of matter at atomic and molecular levels. X-rays on hitting the crystal of the protein get reflected back based on the bond lengths and bond angles of the atoms that make up the protein structure.

This may have to be repeated till we get an exact structure. By knowing all the protein structures and thus the amino acid sequence, we can know what the function of the protein is. We can also determine whether it is harmful or not.



X-RAY CRYSTALLOGRAPHY requires growing a pure crystal (*inset*) of the protein under study. Here a crystal of CD4, the protein that serves as a gateway for the AIDS virus to infect immune cells, is held in a tiny tube sealed with a ball of wax. The tube will be bombarded with x-rays to yield a pattern that scientists can interpret to determine the three-dimensional structure of an individual molecule of the protein.

BIOINFORMATICS

Biological research is becoming increasingly database driven, motivated, in part, by the advent of large scale functional genomic and proteomics experiments such as those comprehensively measuring gene expression. These provide a wealth of information on each of the thousands of proteins encoded by a genome. A challenge in bioinformatics is integrating databases to connect this information as well as performing large-scale studies to collectively analyze many different data sets.

Proteome data generated is entered in Proteome Science's proprietary databases, and in some cases, public databases. This makes future references much easier. Analysis of the data and searches of proprietary and public protein and gene databases are coordinated using bioinformatics technology.

FUTURE SCOPE

Proteomics is a developing field. The whole proteome of our body is not yet catalogued. New machines can be made which can sort out the proteins in them right when they are given the sample. These can be made to be done with more accuracy than now available.

Proteomics will be vital to the future of the pharmaceutical industry. We can make more and better drugs using this technology.

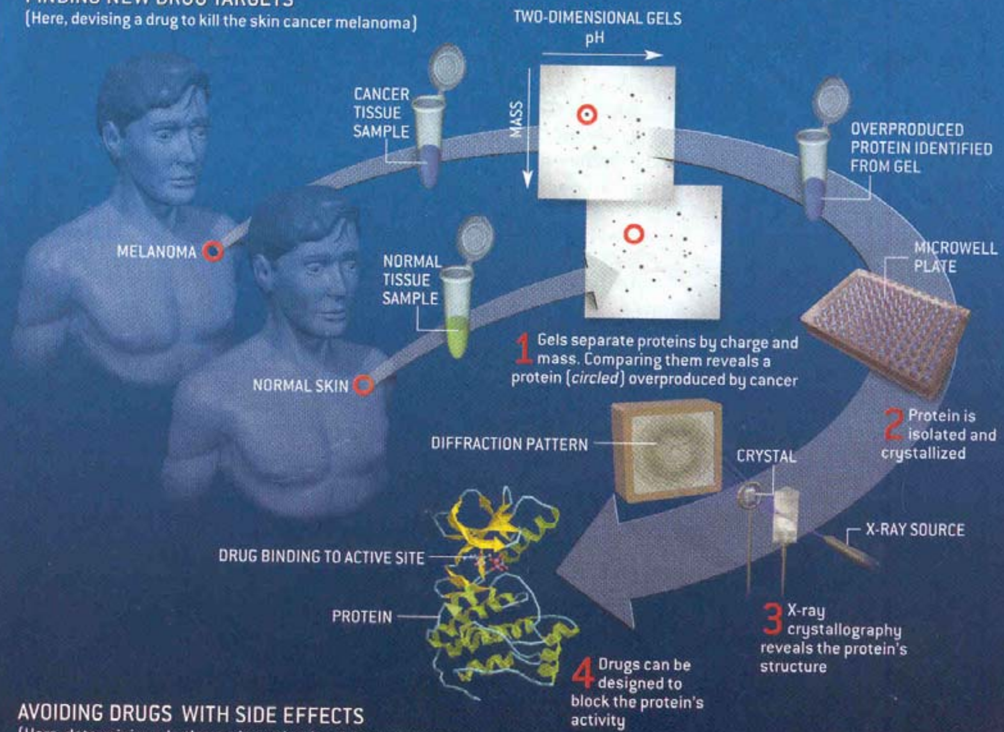
APPLICATIONS

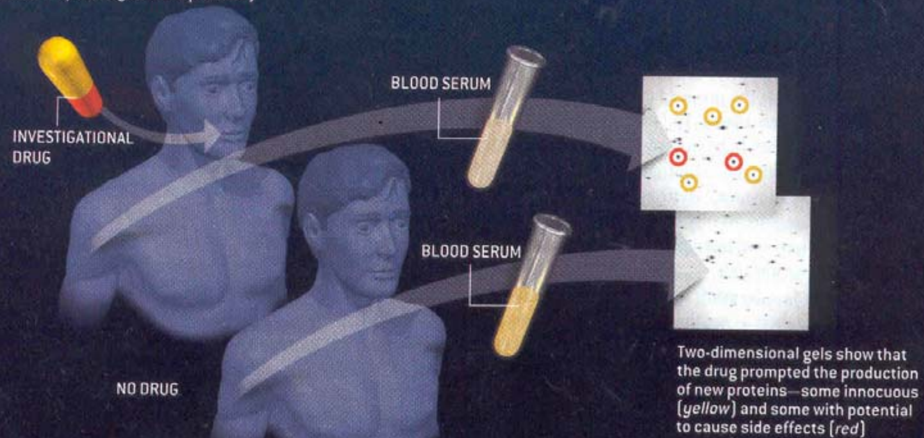
Proteomics is vital in the field of medicines.

- It can be used for diagnostic purposes.
- It can also be used to monitor the effect of treatment.
- Exact location of disease causing protein can be found.
- Drugs can be made more effective.
- Medicines with lesser side effects can be manufactured.

This can be illustrated with a figure.

HOW PROTEOMICS CAN HELP DRUG DEVELOPMENT





ADVANTAGES

- More and better medicines can be manufactured.

  As the exact protein, which causes the disease, is known, better drugs can be made.

  Such medicines will take effect soon.

  The speed of clinical tests can be accelerated.

  Efficiency of tests can be enhanced.

- Industrialization and use of computers.

  Automation is going on in this field.

  Computers are used to store data.

  Tests are done with the help of machines for accuracy.

DISADVANTAGES

- The study of proteins is very complex

  It is hard to determine the complete structure of proteins with the currently used equipments.

  It is impossible to determine the entire structure of the proteins by just knowing the sequence of the amino acids forming it.

  The three dimensional structure of the proteins are difficult to estimate with accuracy as they can fold into any shape. Moreover, there can be the addition of any phosphate or carbohydrate group to these amino acid groups which may change its function.

- There are a large number of proteins in our body and all of these have to be catalogued.

  A typical cell makes hundreds of thousands of distinct proteins. It will be difficult to catalogue all these proteins and their functions.

- There are variations in the proteins from cell to cell.

  The set of proteins produced in each cell differs from the other. For instance, the pancreas makes a different set of proteins than the brain does.
  Nature of proteins at any instant too can vary.

CONCLUSION

Proteome research permits the discovery of new protein markers for diagnostic purposes and of novel molecular targets for drug discovery. It is obvious that without the help of instruments like mass spectrometers, electrophoresis equipments and x-ray crystallographs , proteomics would be impractical.

The abundance of information provided by proteome research is entirely complementary with the genetic information being generated by genomic research. Proteomics will make a key contribution to the development of functional genomics. The combination of proteomics and genomics will play a major role in biomedical research and will have a significant impact on the development of diagnostic and therapeutic products of the future.

BIBLIOGRAPHY

1. “Proteins Rule”, Scientific American, pp 26-33, April, 2002.
2. Williard, Merrit, Dean&Seattle, “Instrumental methods of analysis”, CBS, pp 465-488.
3. L.A.Yeddes, L.E.Baker, Principles of Applied Biomedical Instrumentation, pp 304-307
4. John.G.Webster, Medical Instrumentation-Application and Design, pp 505, 418-420.
5. Joseph.D.Bronzino, Handbook of Biomedical Instrumentation, pp 1492-1499.
6. Paul Bertone and Mark Gerstein, “Integrative data mining: the new dirction in bioinformatics”, IEEE Engineering in medicine and Biology, pp33-40, July 2001.
7. www.proteomelab.org
8. www.sciam.com

ABSTARCT

Proteomics is biotech's "new new thing." It is basically concerned with the proteome, which is the collection of all the proteins in made by a person's cells and tissues.

Proteomics seems to overcome the shortcomings of genomics. The enthusiasts of proteomics are racing to catalogue the proteins in our body to figure out how they network with one another. Proteomics can be broken down into three main activities: identifying all the proteins made in a given cell, tissue, or organism; determining how those proteins join forces to form networks; and outlining the precise three dimensional structures for the proteins. Without the help of instruments, the idea of proteomics would not work out. Proteomics would still remain a theory. Proteomics could lead to more and better drugs.

ACKNOWLEDGEMENT

I extend my sincere gratitude towards **Prof . P.Sukumaran** Head of Department for giving us his invaluable knowledge and wonderful technical guidance

I express my thanks to **Mr. Muhammed kutty** our group tutor and also to our staff advisor **Ms. Biji Paul** for their kind co-operation and guidance for preparing and presenting this seminar.

I also thank all the other faculty members of AEI department and my friends for their help and support.

CONTENTS

- INTRODUCTION
- FROM THE GENOME TO THE PROTEOME
- PROTEINS
- PROTEOMICS: A DESCRIPTION OF THE METHODOLOGY
  - SAMPLE COLLECTION, HANDLING AND STORAGE
  - PROTEIN SEPARATION BY 2D ELECTROPHORESIS
  - PROTEIN IDENTIFICATION BY MASS SPECTROMETRY
  - PROTEIN CHARACTERISATION BY X-RAY CRYSTALLOGRAPHY
  - BIOINFORMATICS
- FUTURE SCOPE
- APPLICATIONS
- ADVANTAGES
- DISADVANTAGES
- CONCLUSION
- BIBLIOGRAPHY