Lecture 8 - Protein Analysis Tools and Methods

1.1 Protein Analysis Tools and Methods

Notes:

Welcome to Lecture 8 – Protein Analysis Tools and Methods. This lecture was written by Dr. Leah Gillis of the Florida Department of Health, Bureau of Laboratories in Miami. Dr. Roberts is providing lecture audio by request. This lecture was reviewed in August 2022 by Dr. Gillis and was reviewed and updated in August 2023 by Dr. Marty Soehnlen and by Dr. Kara Mitchell.

1.2 Introduction

Introduction

- The study of protein characterization tools and methods is important historically and presently.
- Amino acids, peptides, and proteins are critical in all biological \bullet processes.
- Understanding them is key to host and infectious agent molecular \bullet interactions.
- Recent studies have used multiple tools and methods to determine \bullet biomarkers of disease which provide improved diagnosis of infectious organisms and disease states.

Notes:

Introduction

For this lecture we will study protein characterization tools and methods. The lecture requires the reading of papers that provide perspectives into the historical significance and the renewed importance of studying and learning about the details of proteins. As amino acids, peptides, and proteins are critical in all biological processes, the knowledge gained yields insights into host and infectious agent molecular interactions. As we will discover, recent studies have used these tools and methods to develop biomarkers of disease, with the overall goal to provide significant improvements in the diagnosis of infectious organisms and disease states.

1.3 Learning Objectives

Learning Objectives

1. Describe original protein sequencing methods

2. Appraise protein characterization techniques and tools including:

- A. Quantification
- B. Chromatography
- C. Electrophoresis
- D. Mass spectrometer analysis

3. Evaluate current protein analysis methods in determining new biomarkers of infectious disease diagnosis

Notes:

Learning Objectives

- 1. Describe original protein sequencing methods
- 2. Appraise protein characterization techniques and tools including:
	- A. Quantification
	- B. Chromatography
	- C. Electrophoresis
	- D. Mass spectrometer analysis

3. Evaluate current protein analysis methods in determining new biomarkers of infectious disease diagnosis

1.4 Lecture Journal Article #1

Notes:

To get us thinking about amino acids and their importance in a molecular biology course, we will first read a paper from 2002 to commemorate Fred Sanger and his sequencing of the Beta chain of insulin in 1951. Narrator's Note: This paper is posted in Canvas and is assigned reading for this lecture. This essay was written 50 years following Sanger's momentous paper and celebrates Sanger's completion of the first complete amino acid sequence of a protein. Keep in mind that this occurred before the 1953 publication of the structure of DNA by Watson and Crick and before DNA was shown to be the encoding molecule for proteins, our phenotypic calling cards.

Remarkably, Fred Sanger also developed methods to sequence the other two molecules (DNA, RNA) in the paradigm of Biology [DNA to RNA to Amino acid (protein)] and was the recipient of two Nobel Prizes.

1.5 Knowledge Check Question 1

Notes:

Knowledge Question 1

Why did Sanger use the FDNB method instead of the previously published Edman degradation method?

Answer:

Although the Edman sequential degradation method was the preferred method for many years (talking decades here) Sanger preferred the yellow DNP-derivatives of the cleavage products, which were easier to follow as yellow bands moving down the column. (Built in quality control for the lab experiment is good!)

1.6 FDNB and Edman Degradation Process

Notes:

FDNB and Edman Degradation Process

Steps for sequencing a polypeptide:

1)Shown here are Sanger's steps for identifying the amino acid terminal residue of a polypeptide – see b.

2)For shorter peptides, the Edman degradation process will identify the entire amino acid sequence and steps (a) and (b) are omitted.

Note that unless otherwise stated, the illustrations used are from Lehninger Principles of Biochemistry, 3rd ed, David L Nelson and Michael M. Cox, Worth Publishers, New York, NY.

1.7 Protein Sequencing

Notes:

Protein Sequencing: Cleaving and Ordering

Process of cleaving proteins, then sequencing and ordering the peptide fragments:

1)Determine the amino acid composition and amino terminal residue of an intact sample. Next, the disulfide bonds are broken before peptide fragmentation enabling efficient sequencing. For this example there are only two Cys (C) residues, thus only one loacation for the disulfide bond.

2)Peptide fragmentation proceeds as described in figure, using enzymes (proteases) and/or chemicals.

3)The shorter polypeptides are sequenced using the Edman degradation method on a machine called a sequenator, using reagent to label each newly exposed amino terminal residue.

1.8 Lecture Journal Article #2

Notes:

This article is being used to illustrate different protein determination methods, especially those used in food analysis. Narrator's Note: This article is assigned reading for this lecture and the PDaF is posted in Canvas. The additional suggested reading in the thought questions, specifically question 2B, should address the abuse of protein determination mentioned in the Food journal article.

1.9 Knowledge Check Question 2 & 3

Notes:

Knowledge Questions 2 and 3

2. A) What are the five protein determination methods discussed in this paper? B) What is the most accurate method?

B) What is the most accurate method?

3. A) Which methods use nitrogen content to determine protein levels in food?

B) How was the use of nitrogen determination abused in 2007 and 2008?

Answers:

2. A) Kjeldahl Method, Modified Lowry Method, The Bradford Method, Amino Acid Analysis, Spectrometry Methods

B) Amino acid analysis

3. A) Kjeldahl Method, Dumas Method

B) Foods were adulterated with melamine, a non-protein substance rich in nitrogen. For further reading of food pulled by the U.S. FDA, see reference #24 in this assigned paper which is a good review of the problem and provides a list of the human foods that have been pulled by the FDA, although not much information is provided on the 2007 pet food contamination problem in the U.S., which was a very serious problem for pet owners.

1.10 Early Separation Techniques

Notes:

Early Separation Techniques

To study a protein in detail it must be separated from all other proteins and then techniques must be available to determine its properties. Crude protein extracts from tissues, cell organelles, culture or other sources are usually fractionated based on properties such as size or charge and makes use of differences in protein solubility, which is a complex function of pH, temperature, salt concentration and other factors. Ammonium sulfate ((NH4)2SO4) is often used due to its high solubility in water. Its addition, in the right amounts the method can effectively "salt out" or precipitate some proteins. Following protein precipitation, the ammonium sulfate is removed by dialysis. Dialysis separates proteins from solvents by taking advantage of the proteins larger size and utilizes a bag, tube, or cartridge made of a semipermeable membrane which is placed or suspended in a larger volume of buffered solution of appropriate ionic strength.

The most powerful methods for fractionating proteins make use of column chromatography. Column chromatography takes advantage of differences in protein charge, size, binding affinity and other properties of the protein(s) of interest. A porous solid material with the appropriate chemical properties is held in a column and called the stationary phase. A buffered solution, called the mobile phase, is percolated through the stationary phase. Then the protein containing solution is layered on the top of the column so it can percolate through the solid matrix. Individual proteins migrate faster or more slowly through the column depending on their properties.

1.11 Chromatography

Notes:

Chromatography

Illustrated here are the standard elements of a chromatographic column. Within the plastic column is the solid porous material or matrix – the stationary phase of the column. The solution or mobile phase flows through the matrix, passing out of the bottom of the column as the effluent. The protein solution to be separated is placed on top of the column to move through or percolate through the stationary phase. As proteins migrate through the column, they are retarded to different areas throughout the length of the column due to their different interactions with the matrix material of the column. Throughout the length of the column, individual types of proteins, represented as A, B, and C here, gradually separate from each other, forming bands. Note that separation improves (resolution of different proteins) as the length of the column increases. Also, in this illustration protein A is well separated from proteins B and C, while diffusional spreading decreases the separation or resolution of proteins B and C under these conditions.

1.12 Chromatography: Methods

Column chromatography: the stationary phase may consist of a particle of silica or polymer or it may be coated onto or bonded to these support particles. It may be packed into a tube or column or it may be coated onto the inner surface of the tube.

GC uses gas for the mobile phase and LC uses a liquid for the mobile phase.

HPLC uses high pressure pumps to increase resolution of solutes.

Notes:

Chromatography: Methods

Figure adapted from Tietz Fundamentals of Clinical Chemistry and Molecular Biology, 7th Ed. Chapter 12.

In planar chromatography: the stationary phase is coated on a sheet of paper or is bound to a solid surface (TLC).

Column chromatography: the stationary phase may consist of a particle of silica or polymer or it may be coated onto or bonded to these support particles. It may be packed into a tube or column or it may be coated onto the inner surface of the tube.

GC uses gas for the mobile phase and LC uses a liquid for the mobile phase.

A modern refinement of liquid chromatography is HPLC or high performance liquid chromatography, using high pressure pumps that speed the movement of the protein molecules in solution down the column and the method uses higher quality chromatographic materials that can withstand the crushing force of the pressurized flow. Reducing the transit time on the column limits the diffusional spreading and improves resolution of the solutes.

Many adaptations to these methods have been made over the past decades for the purification of biological macromolecules. For example chromatographic columns are part of most purchased nucleic acid purification kits.

1.13 Gas chromatography: Instrumentation

- 1. A supply of carrier gas and flow controller.
- 2. An injector or injection port to introduce samples into the system.
- 3. A chromatographic column to separate the analytes.
- 4. A column oven to heat the column.
- 5. A detector that responds to solute/analyte concentration.
- 6. A computer to control the system and to process the data.

Notes:

Gas Chromatography: Instrumentation

This is a brief overview of the components of a gas chromatograph instrument. For gas chromatography (GC) the mobile phase is a gas. Gas chromatography when combined or connected to a mass spectrophotometer is termed GCMS.

- 1.A supply of carrier gas and flow controller.
- 2.An injector or injection port to introduce samples into the system.
- 3.A chromatographic column to separate the analytes.
- 4.A column oven to heat the column.
- 5.A detector that responds to solute/analyte concentration.
- 6.A computer to control the system and to process the data.

1.14 Protein Structure

Notes:

This is a review slide to consider protein structure:

A protein's primary structure consists of a sequence of amino acids linked together by peptide bonds and includes any disulfide bonds. The resulting polypeptide can be coiled together into units of secondary structure, such as an α helix. The helix, or other secondary structure, is part of the tertiary structure of the folded polypeptide. In this example, the protein hemoglobin, four subunits or polypeptide chains make up the quaternary structure, that is two α chains and two β chains and their arrangement in space is depicted here.

Of course, globular (spherical) structure of hemoglobin is just one example…proteins come in all shapes and sizes, dependent upon the function of the protein, for example: ligaments are made of extended, linear proteins.

1.15 Three types of liquid column chromatographic methods - 1

Notes:

Three types of liquid chromatographic methods are used in protein purification using columns. These are ion exchange chromatography, size-exclusion chromatography, and affinity chromatography. These three methods use a liquid as the mobile phase or eluent. For ion-exchange chromatography, shown here, separation is based on exchange of ions between a matrix surface and eluents.

1.16 Three types of liquid column chromatographic methods - 2

Notes:

Shown here is size-exclusion chromatography. Also known as gel filtration, gel permeation, steric exclusion molecular exclusion or molecular sieve chromatography. Solutes are separated based on their size in solution. For example, a preparative step is used to separate large molecules such as proteins or nucleic acids from small molecules such as salts or oligonucleotides and may be performed in small spun columns where elution is driven by spinning in a centrifuge.

1.17 Three types of liquid column chromatographic methods - 3

Notes:

The third type of column chromatography used in protein purification is affinity chromatography. Also called Adsorption chromatography, with separation due to a series of adsorption/desorption steps.

1.18 Electrophoresis: Protein Separation

Notes:

Electrophoresis: Protein Separation

Electrophoresis, is another important technique for the separation of proteins and is based on the migration of charged proteins in an electric field. This procedure is not generally used to purify proteins in a large amount as other methods are available and electrophoretic methods generally affect the structure and thus the function of a protein. It is a very useful analytical method, as it permits visualization, permitting the estimation of the number of different proteins in a mixture or the purity of a protein of interest. Plus, importantly, this method permits determination of two properties of a protein: isoelectric point and approximate molecular weight.

Electrophoresis is usually carried out in gels made up of the cross-linked polymer polyacrylamide which acts as a molecular sieve, slowing the migration of proteins approximately in proportion to their charge–to-mass ratio. The shape of the protein may affect migration also. The migration of a protein in a gel during electrophoresis is a function of the size and shape of the protein. The force moving the protein is the electrical potential, E. The electrophoretic mobility of the molecule, μ, is the ratio of the particle, V, to the electrical potential, E. Electrophoretic mobility, μ, is also equal to the net charge of the molecule, Z, divided by the frictional coefficient ,f , which reflects in part a protein's shape.

 $\mu = V = Z$

E f

1.19 Protein Separation

Notes:

Protein Separation

For SDS-PAGE (sodium dodecyl sulfate polyacrylamide gel electrophoresis) uses the detergent (SDS) to bind to the proteins (hydrophobic interaction) in amounts roughly proportional to the molecular weight (MW) of the protein, and contributes a large net negative charge, making each protein in a mixture being analyzed exhibit a similar charge to mass ratio. The SDS also denatures the protein, altering the native shape or conformation of all proteins in the mixture, all of the proteins in the mixture have a similar shape.

In this illustration: different samples are loaded into the wells at the top of the polyacrylamide gel. The proteins move into the gel when an electrical field is applied. Following electrophoresis, a stain is applied (Examples: Coomassie blue, Amido black) which binds to the proteins only and not the gel. Each band on the stained gel represents a different protein (or protein subunit, for example with hemoglobin). Smaller proteins move through the gel more rapidly than the larger proteins and are located nearer the bottom of the gel. This gel represents sequentially purified extracts of the bacterium E. coli, from a crude cellular extract on the left to preparations following each additional purification step. The last lane represents the final purified protein (the enzyme RNA polymerase, with four subunits).

1.20 SDS PAGE: Visualization

Notes:

SDS Page: Visualization of Separation and Estimating MW.

Determination of unknown protein weight uses known molecular weight (MW) markers, which are run in adjacent well. Standard proteins of known MW are electrophoresed (Lane 1) and used to estimate the MW of an unknown protein (Lane 2).

A plot of log Mr of the marker proteins versus relative migration during electrophoresis is linear, allowing the MW of the unknown protein to be read from the graph.

1.21 Protein separation: Isoelectric focusing

Notes:

Protein Separation: Isoelectric Focusing

Separates proteins according to their isoelectric points.

A protein mixture is applied to a pH gradient gel.

After applying an electric field to the gel, proteins enter the gel and migrate until each reaches a pH equivalent to its pI.

Remember that when pH = pI, the net charge of a protein is zero.

1.22 Mass Spectrometry (MS)

Mass Spectrometry (MS)

- Useful for determining elemental composition and structure of inorganic and organic compounds
- A powerful qualitative and quantitative analytical technique
- Coupling with GC or LC expands analytical capabilities and clinical applications
- Ability to identify and quantify makes it a key analytical tool in proteomics especially in sequencing unknown proteins

Notes:

Mass Spectrometry (MS)

Useful for determining elemental composition and structure of inorganic and organic compounds

A powerful qualitative and quantitative analytical technique

Coupling with GC or LC expands analytical capabilities and clinical applications

Ability to identify and quantify makes it a key analytical tool in proteomics especially in sequencing unknown proteins

Mass spectrometry was used for many, many years in chemistry before applied to study macromolecules such as proteins and nucleic acids. Early on, measurements were made on molecules in the gas phase and macromolecules would rapidly decompose. In the mid to late 1980s two techniques were developed to overcome this obstacle. One was the use of a short pulse of laser light to ionize the proteins followed by desorbing the mixture from the matrix into a vacuum system. This process is termed matrix-assisted laser desorption/ionization mass spectrophotometry or MALDI MS. For the second and equally successful adaptive method, macromolecules in solution are forced directly from the liquid to the gas phase.

In proteomics, mass spectrometry is used to sequence short stretches of polypeptides, making the instrument invaluable for quickly identifying unknown proteins.

1.23 Electrospray Mass Spectrometry

Notes:

Electrospray Mass Spectrometry

This figure shows a protein solution being dispersed into highly charged droplets by passage through a needle under the influence of a high voltage field. The droplets evaporate and the ions, with added protons shown here, enter the mass spectrometer for m/z measurement.

1.24 Electrospray Mass Spectrometry (continued)

Notes:

Electrospray Mass Spectrometry (continued)

The spectrum shown here is a family of peaks with each successive peak from right to left corresponding to a charged species that is increased by 1 in both mass and charge.

The inset figure is a computer generated transformation of this spectrum.

A mass spectrum is represented by the relative abundance of each ion plotted as a function of its mass to charge ratio or the m/z ratio. This measurement of the ionized species can be used to deduce the mass (M) of the analyte with very high precision. The most abundant ion is arbitrarily assigned 100%.

1.25 Tandem Mass Spectrometry

Notes:

Tandem Mass Spectrometry

Following proteolytic hydrolysis, a protein solution is injected into a mass spectrometer (MS-1). Next, the different peptides are sorted, so that only one type is selected for further analysis. The selected peptide is further fragmented in a chamber between the two mass spectrometers and m/z is measured in the second mass spectrometer (MS-2). Note that many of the ions generated during the second fragmentation result from breakage of the peptide bond, illustrated here. These are called b-type or y-type bonds, depending on whether the charge is retained on the amino- or carboxyl-terminal side, respectively.

Note that the fragment ions from the first MS-1 are sometimes known as product ions since they have been separated by the m/z value in the first stage of mass spectrometry (MS-1) and then selected for further analysis in MS-2.

1.26 Tandem Mass Spectrometry (continued)

Notes:

Tandem Mass Spectrometry (continued)

Shown here is a typical spectrum with peaks representing the peptide fragments generated from a sample of one small peptide of 10 residues. The labeled peaks are y-type ions. The large peak next to y5" is a doubly charge ion and is not part of the y set. The successive peaks differ by the mass of a particular amino acid in the original peptide. In this case the deduced sequence was Phe-Pro-Gly- Gln-(Ile/Leu)-Asn-Ala-Asp-(Ile/Leu)-Arg. Ile and Leu have the same molecular mass and thus the ambiguity. In this example the set of peaks derived from the y type ions predominates and the spectrum is greatly simplified. This is due to the Arg residue at the amino terminus of the peptide and most of the positive charges are retained on this residue.

1.27 Current Mass Spectrometers

Notes:

Current Mass Spectrometers

Since mass spectrometers measure m/z (Mass spectrum result) and not molecular mass it does have a fundamental impact on the physical operating principles of mass spectrometers and has influenced all aspects of instrument design since the mid 1980s and these instruments began being developed to be extended into the realm of biological molecular analysis.

A current mass spectrometer consists of five parts:

1)Ion source 2) Vacuum system 3) Mass analyzer 4) Detector and 5) Computer

Note that the ion source is required for the ionization step in which an ion is produced from a neutral atom or molecule. Many approaches are used and include: 1) Electrospray ionization (ESI) 2) atmospheric pressure chemical ionization (ACPI) 3) atmospheric pressure photoionization (APPI) 4) inductively coupled plasma (ICP) and matrixassisted laser desorption/ionization (MALDI).

When coupled with HPLC (or HPLC-MS) this step is usually of the ESI or APCI ionization method.

1.28 Mass analyzers and Ion detectors for

Mass analyzers and lon detectors for Mass spectrometers

Current classification of mass analyzers: 2 types Beam-type

1) quadrupole 2) magnetic sector 3) time-of-flight (TOF)

Trapping-type

1) quadrupole ion traps 2) linear ion traps 3) ion cyclotron resonance 4) orbitrap

Current classification of ion detectors: 3 types

1) discrete dynode multipliers 2) continuous dynode electron multipliers 3) microchannel plate (MCP) electron multipliers

Notes:

Mass Analyzers and Ion Detectors for Mass Spectrometers

Now the general classes of mass spectrometer instruments are:

1) Beam-type or 2) trapping

For beam type: the ions make one trip through the instrument and then strike the detector be destructively detected Beam-type, the entire process takes microseconds to milliseconds

For trapping type analyzers, the ions are held in a spatially defined region of space by a combination of magnetic, electrostatic, and/or RF fields. The trapping fields are manipulated in ways that allow m/z measurements to be performed. Trapping times vary from a small fraction of a second to minutes, with clinical applications being at the low end of this range.

Beam-type instruments include 1) quadrupole 2) magnetic sector and 3) time-of-flight (TOF). Since TOF are now being used in labs, some details are below:

Modern TOFs produce single digit parts per million (ppm) mass accuracy, allowing confirmation of a compound molecular formula. It is a pulsed technique and couples readily to pulsed ionization methods with MALDI being the most common example. MALDI-TOF is used in protein and peptide and bacterial identification.

Likewise ion detectors have improved and current classification includes 3 types

1) discrete dynode multipliers 2) continuous dynode electron multipliers 3) microchannel plate (MCP) electron multipliers

1.29 Mass Analyzers

Notes:

Mass Analyzers

MALDI – Ion Source

Ion Source – production of ions , gas phase techniques, desorption/ionization techniques – process of an atom or molecule acquiring a charge cation (+ charge), anion (- charge).

1.30 Mass Analyzers

Notes:

Mass Analyzers TOF

Separates ions based on their mass-to-charge ration.

Magnetic and/or electric field(s) exert a force on an ion, and this force accelerates ions to different amounts, depending on the mass of the ion.

1.31 Mass Analyzers

Notes:

Mass Analyzers – MCP - Detector

Turns ions into signals the computer can read and produce a spectrum with.

1.32 MS Applications

MS Applications

\cdot GC-MS

Application: drug testing for clinical or forensic purposes, also for detecting xenobiotic compounds, anabolic steroids, pesticides, pollutants and inborn errors of metabolism

• HPLC-MS/MS

Application: Screening and confirmation of genetic disorders and inborn errors of metabolism

• MALDI-TOF

Application: expanding field of proteomics, identification of bacteria and fungi

Notes:

MS Applications

GC-MS

Application: drug testing for clinical or forensic purposes, also for detecting xenobiotic compounds, anabolic steroids, pesticides, pollutants and inborn errors of metabolism

HPLC-MS/MS

Application: Screening and confirmation of genetic disorders and inborn errors of metabolism

MALDI-TOF

Application: expanding field of proteomics, identification of bacteria and fungi

Mass spectrometers coupled with either gas or liquid chromatographs (GC-MS and LC-MS) are versatile analytical instruments combining the separation power of chromatographs and the specificity and low detection limits of a mass spectrophotometer.

MALDI-TOF: has been used with ICP ionization techniques which have enhanced the analytical capabilities. An important use is a primary analytical tool for discovery in the rapidly developing and expanding field of proteomics. Also important is its use in the identification of bacteria and fungi. For identification purposes, maintaining an up to date reference library file is a necessity, especially for identifying emerging infection organisms.

1.33 Knowledge Check Question 4 & 5

Notes:

Question 4: What are the three main components of a mass spectrophotometer (MS)?

Answer: Ion source, Mass analyzer, and Detector

Question 5: Which type of MS instrument is being applied to identify bacteria in clinical and PHLs?

Answer: MALDI-TOF

1.34 Conclusions

Conclusions

- Proteins as diagnostic tools have been important since the beginning of the 20th century beginning with the study of immunoglobulins and their role in health and their resultant critical role in diagnostics.
- By using newer, refined protein analytical methods to study proteins and thus, more accurately define their role in health, we should begin seeing advanced diagnostics using these newly characterized biomarkers of disease.

Notes:

Conclusions

Proteins as diagnostic tools have been important since the beginning of the 20th century beginning with the study of immunoglobulins and their role in health and their resultant critical role in diagnostics.

By using newer, refined protein analytical methods to study proteins and thus, more accurately define their role in health, we should begin seeing advanced diagnostics using these newly characterized biomarkers of disease.

1.35 Review

Review

- For this session we learned about protein sequencing's history and some of the techniques and laboratory methodologies that have been used to study and characterize proteins.
- Plus, we learned of the important adaptations in techniques for MS, making the instruments and methods critical for the renewed use of protein sequencing to define peptides and proteins as possible diagnostic biomarkers of infectious disease.

Notes:

Review

For this session we learned about protein sequencing's history and some of the techniques and laboratory methodologies that have been used to study and characterize proteins. Plus, we learned of the important adaptations in techniques for MS, making the instruments and methods critical for the renewed use of protein sequencing to define peptides and proteins as possible diagnostic biomarkers of infectious disease.