EnM
ENDOCRINOLOGY AND METABOLISM

Review Article

Endocrinol Metab 2017;32:18-22
https://doi.org/10.3803/EnM.2017.32.1.18
pISSN 2093-596X · eISSN 2093-5978

A Simple Outline of Methods for Protein Isolation and Purification

Chang-Hun Lee

School of Undergraduate Studies, College of Transdisciplinary Studies, Daegu Gyeongbuk Institute of Science and Technology (DGIST), Daegu, Korea

At the summer workshop of the Korean Endocrine Society held in 2016, some examples of protein experiments were discussed in the session entitled "All about working with proteins." In contrast to what the title suggested, it was unrealistic to comprehensively discuss all protein analytical methods. Therefore, the goal was to outline protein experimental techniques that are useful in research or in bench work. In conversations with clinicians, however, I have always felt that researchers who do not engage in bench science have different demands than those who do. Protein research tools that are useful in bench science may not be very useful or effective in the diagnostic field. In this paper, I provide a general summary of the protein analytical methods that are used in basic scientific research, and describe how they can be applied in the diagnostic field.

Keywords: Chromatography; Molecular imaging; Immunoenzyme techniques

INTRODUCTION

As a simple example, let us think about how to obtain a protein of interest. Proteins can be obtained from a wide variety of samples. For diagnostic purposes, they may be obtained from a patient's cells or tissues, whereas for experimental use on the bench, the proteins may originate from microorganisms or from cell lines derived from insects, vertebrate animals, or plants. In any case, endogenous proteins that we do not want are present in a much greater quantity than the proteins we do want. Regardless of the source, is not easy to isolate a specific protein for several reasons, some of which are presented below.

First, no generalized property can be applied in protein purification. In the case of nucleic acids (DNA or RNA), one can employ selective adsorption techniques using silica residues with positive charges, since nucleic acids have negative charges from their phosphodiester bonds. The development of specialized methods of nucleic acid purification, such as miniprep and midiprep, was made possible by the fact that all nucleic acid molecules are highly negatively charged. In contrast, the applicability of selective adsorption to proteins is limited for several reasons. Hydrophobicity or specific electrostatic properties can be found in a diverse range of molecules, including lipids, nucleic acids, and sugars. However, some proteins may have strong surface charges or specialized functional groups that can be useful for purification. For example, a purification method with phosphocellulose has been used for DNA-binding proteins that have many positively charged residues. Nonetheless, this does not guarantee high purity or loss-free purification. Unlike DNA, protein purification requires multiple steps, rather than a

Received: 31 December 2016, **Revised**: 16 January 2017,
Accepted: 8 February 2017
Corresponding author: Chang-Hun Lee
School of Undergraduate Studies, College of Transdisciplinary Studies, Daegu Gyeongbuk Institute of Science and Technology (DGIST), 333 Techno jungang-daero, Hyeonpung-myeon, Dalseong-gun, Daegu 42988, Korea
Tel: +82-53-785-6612, **Fax:** +82-53-785-6609, **E-mail:** leech@dgist.ac.kr

simple one-step procedure. As will be described later, fractionation during the isolation process and use of the proper column type during the purification process enable successful purification [1].

Second, the amount of the protein of interest tends to be quite small. Intracellular protein concentrations are known to be in the range of 300 mg/mL. However, this value refers to the total amount of proteins in a cell. Since the number of active proteins in cells is in the tens of thousands, most proteins are present in negligible quantities except for some housekeeping gene products. In some cases, the target proteins may be present in picomolar or femtomolar concentrations. Proteins such as keratins can be used as biomarkers for cancer diagnosis because they are overexpressed proteins. The typical analytical instruments used in biophysics are less sensitive than one might think. Spectroscopic or calorimetric methods usually require protein samples with concentrations ranging from 0.01 to 1 mM (or to tens of mg/mL). Therefore, most proteins directly derived from living organisms cannot be analyzed using those methods. In mass spectrometry, it is possible to analyze samples in femtogram units. However, mass spectrometry is not a panacea, because the ionization efficiency may be not high enough to obtain a useful signal, depending on the nature of the peptide fragments.

Third, there is no way to amplify proteins. In the case of nucleic acids, amplification is possible, to the point that even a trace amount obtained in a biological sample can be used for further analysis. Polymerase chain reaction (PCR) amplifies DNA sequences, and RNA sequences can also be amplified by reverse transcription PCR. In contrast, if a protein is not overproduced at the beginning of the experiment, it may be lost continuously during the experiment, to the point that it is not detectable at the final analytical stage.

Fourth, contamination issues are inevitable. This is related to the second and third problems discussed above. More specifically, certain proteins make up an exceptionally large share of the proteins present in cells. Intermediate filament proteins and histones are proteins that appear as contaminants when analyzing trace proteins by mass spectrometry. Samples from human patients pose especially severe problems. These samples may be attached to blood vessels, or may be a mixture of lesions and healthy tissue. Lipid molecules, blood cells, and/or blood plasma may already be present in excess. Proteins that are classified as noise are present in much greater quantities than the target protein, which often makes data analysis difficult.

Fifth, proteins are unstable. It is well known that RNA is quickly degraded, but proteins are also degraded by enzymes both *in vivo* and *in vitro*. It is not easy to block all proteolytic activity by adding protease inhibitors to a test tube. If the temperature, pH, or salt concentration is not suitable, proteins are likely to be denatured. If the experimental conditions of the buffer are not optimal, the proteins may aggregate before or after the analysis. In proteins with cysteine residues, the structure may be destroyed by undesired oxidation-reduction reactions. Experts familiar with protein purification emphasize that researchers should not leave proteins in the refrigerator for a long time because proteins are unstable macromolecules.

PROTEIN ISOLATION

The simplest way to solve all the above problems for researchers is to use overexpression, or enrichment. The rule of thumb for protein experiments is to obtain as much protein as possible at the beginning of the experiment. As discussed above, proteins are inevitably lost over the course of the experiment because of their chemical and/or biological instability. Given these practical limitations, it is important to obtain as much protein as possible at the beginning of the process.

With this in mind, let us start our discussion of protein isolation by addressing methods for the high-efficiency isolation of proteins with little loss [2]. Choosing the best protein isolation method depends on the properties of the source sample (e.g., whether the sample is liquid or solid). Supposing that a solid sample contains a large number of cells, a process must be used for homogenizing the tissues and lysing the cells. In the case of tissue samples, mechanical homogenization methods are useful. Methods for lysing cells range from physical methods, such as heat treatment and sonication, to chemical methods, such as treatment with a detergent solution.

Detergents that increase the solubility of proteins can be used most effectively by taking into consideration the conditions of the experimental medium, especially the buffer. Using appropriate detergents, proteins that are difficult to extract (membrane proteins or nuclear proteins) can also be obtained in desirable amounts. In addition, chaotropic reagents, such as urea and guanidine hydrochloride, can be used to increase the efficiency of extraction because they break down the structure of the protein and dissolve well in water. However, treatment with chaotropic reagents usually requires high salt concentrations. Therefore, if not removed using membrane dialysis, the high salt concentration can cause problems in further steps of the experiment. The salt removal process itself can lead to loss of the protein. One should consider both the advantages and disadvantages of using

chaotropic reagents when designing an experimental procedure.

In the case of liquid samples, a decision should be made about whether to obtain the dissolved protein or to extract it from the cells it is contained in. In order to extract the protein from the cells where it is present, it is necessary to isolate the cells by centrifugation. In particular, centrifugation using media with different densities may be useful to isolate proteins expressed in specific cells. For example, to obtain only the immune cells from bodily fluids or to separate adipocytes or keratinocytes from skin tissue, centrifuging the liquid containing the cells in a high-density medium may precipitate the desired cells depending on the density of each constituent cell. Density-gradient ultracentrifugation is additionally applicable for eliminating undesired cellular impurities or obtaining certain cell organelles.

If a soluble protein is obtained from bodily fluids, it is treated similarly to a cell lysate from solid samples. Protein solutions are generally dilute when they undergo analysis. Thus, it is necessary to perform an enrichment process, such as concentration or precipitation. The traditional techniques of salting out and heat denaturation have the advantage of being very simple. In addition, the precipitated protein is very stable, meaning that this process may be used as a means of increasing the shelf life of a protein. Salting out is a method of lowering the solubility of proteins through competing solubility in water using salts that are more soluble in water, such as ammonium sulfate. Since proteins precipitate at a specific concentration of salt, this procedure also has the advantage that the desired proteins can be separated from other proteins and precipitated.

Instead of using salts, it may be desirable to use isoelectric precipitation by lowering the pH. Isoelectric precipitation can be performed when the pH reaches at isoelectric point (pI) of the target protein. Each protein has a different pI value, meaning that isoelectric precipitation can be used as a fractionation method as well. Generally, this method is also very simple since a mineral acid or trichloroacetic acid (TCA) is titrated until the target protein is obtained through precipitation. When changing the pH or salting out is not preferable, polymers such as polyethylene glycol or organic solvents such as methanol or acetone can be used to promote precipitation. If required, a cocktail of precipitating reagents (e.g., a mixture of acetone and TCA) can be developed.

However, the isolation technique should be chosen with the method of analysis in mind, as the isolation step may cause denaturation of the protein or require additional steps such as salt removal. The centrifugal membrane concentration method is a useful single-step method of concentrating a protein. However, it is important to keep in mind that this method may exhibit problems involving clogging when used with highly concentrated proteins, and such problems decrease the efficiency of the concentration procedure and increase the loss of the protein.

PURIFICATION AND ANALYSIS

After enrichment, purification is necessary. Proteins for analytic experiments must be pure enough to have a high signal-to-noise ratio. Methods for purifying target proteins from dirty mixtures vary widely, but preparation-grade purification is most commonly achieved using chromatography [1,2]. Since, as discussed above, proteins do not have a generalized purification method (unlike nucleic acids), the type of chromatography to be used depends entirely on the physical and chemical properties of the target protein. Proteins are usually purified by liquid chromatography (LC), and fast protein LC and high-performance LC can be chosen depending on whether the goal is preparation or quantitative analysis. For proteins, it is possible to use the following techniques either in a single step or sequentially: hydrophobic interaction column chromatography, size exclusion chromatography, ion exchange column chromatography, and affinity chromatography.

If it is not necessary to prepare a large quantity of the target protein, electrophoresis is a possibility. Electrophoresis separates proteins according to their molecular size. If the molecular weight of the target protein is known, the approximate band position on the gel can be cut and be used for further analysis like mass spectrometry. In general, sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) provides 1-dimensional (1D) results when analyzing proteins by size. Instead of directly loading the protein sample when performing SDS-PAGE, it is possible to perform 2-dimensional (2D) electrophoresis by performing isoelectric focusing and loading the resultant gel tube with proteins separated according to their pI values. Gels developed in 2D have much better resolution than gels developed in 1D, making this a favorite technique in proteomics.

If the target protein is too small for the band (1D) or spot (2D) to be easily stained after electrophoresis, a Western blot using antibodies can be performed. In order to perform Western blotting, it is necessary to have an antibody specific to the target protein. In recent years, Western blotting has become possible in most cases due to the commercial availability of a very wide range of antibodies.

It is also possible to separate and purify proteins using antibodies. Immunoprecipitation is a good enrichment method for

proteins. Since antibody binding is specific, most of the final products are obtained only from the target protein of the antibody. Either protein A-conjugated or protein G-conjugated agarose, which bind to most antibodies, can be used in most cases, meaning that virtually any kind of antibody can be applied to obtain a small amount of target protein through several rounds of centrifugation.

The enzyme-linked immunosorbent assay (ELISA) technique is a sensitive assay method using antibodies. ELISA is an enzyme-amplified reaction for antigens present in trace amounts in bodily fluids. Using ELISA, both the presence of the target protein and quantitative information about it can be obtained without purification. Of course, if enough antibodies are present, it is possible to identify the amount and location of proteins in tissues and cells *in situ* via immunohistochemistry or immunofluorescent labeling.

Antibodies are usually obtained from rats or rabbits. If the antibody is obtained from camelids, a unique antibody protein known as a single-chain antibody or nanobody can be obtained [3,4]. A typical antibody usually consists of four polypeptide chains per molecule, whereas a camelid antibody has two chains per antibody molecule. Since the antigen-binding site of a camelid antibody is composed of a single chain, the cloning and the production process are simpler than those of regular antibodies. As antibody design studies evolve, the use of single-chain antibodies will become more common.

Instead of using antibodies, it may be desirable to use an aptamer for the target protein [5]. Aptamers are nucleic acid molecules that bind specifically to their target proteins. Aptamers can be obtained by screening RNA or DNA macromolecules for binding to target molecules (usually through a method known as SELEX [systematic evolution of ligands by exponential enrichment]). Aptamers are likely to replace antibodies because they are very easy to synthesize *in vitro*.

In recent years, techniques for identifying target proteins in which mass spectrometry and LC are combined have become very common. Mass spectrometry, of course, has the limitation of only allowing qualitative analysis. Metabolic labeling can be used for quantitative analyses, but only in controllable culture conditions in bench work. Nevertheless, it is becoming an accurate and reliable protein analysis method since sequence information can be obtained even from trace amounts of target proteins. Recently, mass spectrometry imaging (also known as imaging mass spectrometry) has been studied as a way of detecting and analyzing proteins directly in specific locations in cells or tissues by imaging mass profiles (or maps) on 2D samples [6,7]. This technique is very similar to finding a neurotransmitter in a specific section of the brain using magnetic resonance imaging. If this technique is generalized, it will be possible to detect the presence or absence of a target protein *in situ* just by looking at the tissue or cell. In other words, it may become possible to determine the presence or absence of a desired biomarker quickly and accurately even in a complex situation (e.g., when no appropriate antibody exists or the sample is a mixture of normal tissue and lesion tissue).

CONCLUSIONS

Much work must still be done in collecting and analyzing proteins, requiring considerable research on the part of bench scientists. As long as proteins continue to be used as the major biomarkers for diagnosing diseases, scientists will make every effort to develop easier methods for protein research. I am confident that communication both among bench scientists and between bench scientists and clinicians will be a major driving force in future protein research.

CONFLICTS OF INTEREST

No potential conflict of interest relevant to this article was reported.

ORCID

Chang-Hun Lee *http://orcid.org/0000-0002-3806-1170*

REFERENCES

1. Burgess RR, Deutscher MP. Methods in enzymology: guide to protein purification. 2nd ed. Vol. 436. San Diego: Academic Press; 2009.
2. Janson JC. Protein purification: principles, high resolution methods, and applications. Hoboken: John Wiley & Sons; 2011.
3. Wang Y, Fan Z, Shao L, Kong X, Hou X, Tian D, et al. Nanobody-derived nanobiotechnology tool kits for diverse biomedical and biotechnology applications. Int J Nanomedicine 2016;11:3287-303.
4. Konning D, Zielonka S, Grzeschik J, Empting M, Valldorf B, Krah S, et al. Camelid and shark single domain antibodies: structural features and therapeutic potential. Curr Opin Struct Biol 2016;45:10-6.

5. Ozalp VC, Kavruk M, Dilek O, Bayrac AT. Aptamers: molecular tools for medical diagnosis. Curr Top Med Chem 2015;15:1125-37.
6. Schwamborn K, Kriegsmann M, Weichert W. MALDI imaging mass spectrometry: from bench to bedside. Biochim Biophys Acta 2016 Oct 31 [Epub]. http://doi.org/10.1016/j.bbapap.2016.10.014.
7. Petras D, Jarmusch AK, Dorrestein PC. From single cells to our planet-recent advances in using mass spectrometry for spatially resolved metabolomics. Curr Opin Chem Biol 2017;36:24-31.