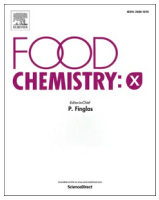
Food Chemistry: X 22 (2024) 101365

Contents lists available at ScienceDirect

Food Chemistry: X

journal homepage: www.sciencedirect.com/journal/food-chemistry-x



Analytical approaches for assessing protein structure in protein-rich food: A comprehensive review

Tian Lan [a,1], Yabo Dong [a,1], Lianzhou Jiang [a], Yan Zhang [b,*], Xiaonan Sui [a,*]

[a] College of Food Science, Northeast Agricultural University, Harbin 150030, China
[b] College of Horticulture and Landscape Architecture, Northeast Agricultural University, Harbin 150030, China

ARTICLE INFO

Keywords:
Food protein
Conformation
Structure
Detection methods

ABSTRACT

This review focuses on changes in nutrition and functional properties of protein-rich foods, primarily attributed to alterations in protein structures. We provide a comprehensive overview and comparison of commonly used laboratory methods for protein structure identification, aiming to offer readers a convenient understanding of these techniques. The review covers a range of detection technologies employed in food protein analysis and conducts an extensive comparison to identify the most suitable method for various proteins. While these techniques offer distinct advantages for protein structure determination, the inherent complexity of food matrices presents ongoing challenges. Further research is necessary to develop and enhance more robust detection methods to improve accuracy in protein conformation and structure analysis.

1. Introduction

Proteins are often considered to be nature's robots that are ubiquitous in living organisms. They are highly complex macromolecules responsible for essential biochemical processes in living organisms. As a central constituent of food due to its unique functionality and high nutritional value, protein is arguably the most crucial among the three macronutrients (protein, fat, and carbohydrate) (Fazzino, Amber, Juen, & Kevin, 2023). Food proteins have been widely studied for their positive contribution to human health. The most common food proteins encountered in daily life are animal and plant proteins. Animal proteins include gelatin, milk protein and egg protein, while plant proteins include soy protein, wheat protein and peanut protein. Microbial proteins are rapidly developing as a new sustainable alternative protein (Jach, Serefko, Ziaja, & Kieliszek, 2022). Proteins produced by microorganisms such as *Candida utilis* and Candida genus yeasts have been widely used in the food industry and feed production (Kieliszek et al., 2017; Kurcz, Błażejak, Kot, Bzducha-Wróbel, & Kieliszek, 2018). Animal proteins are superior to plant proteins in nutritional quality, absorption rate, essential amino acids, and minerals. However, their cost and disease-related risk factors (high calorie, high cholesterol, and high fat) are often deemed greater than those of plant proteins. As a cost-effective source of proteins, plant proteins possess functional characteristics such as emulsification, structural change, and color control. These proteins are broadly applied in food industries, however, their composition frequently includes anti-nutritional factors that negatively impact bioavailability (Samtiya, Rotimi, & Tejpal, 2020). Various functional characteristics of food proteins (gel formation and rheological behavior, emulsifying ability, foaming ability, etc.) cause changes in the appearance, taste, texture, and rheology of food products under different physical, chemical and enzymatic processing modifications. It is well established that the modification of food protein functionality is accompanied by conformational changes, and that small alterations in protein conformation have significant influences on its physicochemical and functional properties. Numerous studies have been conducted on the structure, characteristics, and relationship between the function of food proteins at the primary, secondary, tertiary, and quaternary structure levels. Our previously study found that addition of epigallocatechin-3-gallate (EGCG) causes a structure changes. At pH 4.5, with a greater proportion of EGCG, the α-helix and β-turn contents of protein were increased, with a corresponding reduction in β-sheet content (Lan et al., 2020). Guo et al. (2019) reported an increase in the β-sheets content and a decrease in α-helices content of preserved egg white during pickling. It was also observed that the protein vitro digestibility was negatively correlated with the content of β-sheets structure. Kar, Snigdha, Jeyamkondan, and Kaustav (2023) found that

* Corresponding author.
E-mail addresses: zhang_yan@neau.edu.cn (Y. Zhang), xiaonan.sui@neau.edu.cn (X. Sui).
[1] The first two authors contributed equally to this work.

https://doi.org/10.1016/j.fochx.2024.101365
Received 5 October 2023; Received in revised form 24 March 2024; Accepted 5 April 2024
Available online 8 April 2024

radiofrequency treatment does not cause changes in the secondary structure of proteins by examining the protein structure. Li et al. (2023) illustrated that ultrasound treatment could induce the exposure of protein hydrophobic groups and the formation of disulfide bonds, indicating changes in tertiary and quaternary structures, which could change structures of myofibrillar protein gel. Consequently, effective identification and analytical methods are essential to uncovering protein conformational changes and understanding the relationship between structural and functional properties during modification. In recent decades, novel rapid and nondestructive techniques have been developed and significantly improved for evaluating food protein structure and properties. Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE), size-exclusion chromatography (SEC) and other techniques have been developed to determine the relative molecular weight of proteins (Brunelle & Green, 2014; Striegel, 2023). Mass spectrometry plays an important role in protein primary structure analysis. Moreover, some spectroscopic techniques including Fourier transform infrared spectroscopy (FTIR) (Sazonova, Grube, Shvirksts, Galoburda, & Gramatina, 2019), Raman spectroscopy (Maiti, Apetri, Zagorski, Carey, & Anderson, 2004), circular dichroism spectroscopy (CD) (Phillips-Jones & Harding, 2019), and fluorescence spectroscopy (Wang, Sun, Pu, & Wei, 2017) are simple, rapid, and have drawn significant attention for their monitoring of secondary and tertiary structure changes of proteins.

The study of protein structure and functional properties has rapidly accelerated with the increasing demand for proteins by the food industry and consumers. Methods for identifying protein structures are diverse, and some have gained considerable popularity due to their unique advantages. We summarize a selection of examples commonly used in protein structure detection, as shown in Table 1. However, to the best of our knowledge, there is no extensive review of techniques for protein detection from isolation and primary structure identification to advanced structure detection. Due to the wide range of studies in this area, the current review intends to provide an overview of the methods including SDS-PAGE, two-dimensional SDS-PAGE, SEC, MS, Edman degradation, FTIR, CD, Raman spectroscopy, fluorescence spectroscopies, X-ray diffraction (XRD), nuclear magnetic resonance (NMR), and other techniques that are used for rapid evaluation of protein structure. This review both summarizes and compares these methods for determining protein structure in food systems. It discusses various application scenarios, along with the advantages and disadvantages of employing these techniques. The intention is to offer a valuable reference for the majority of food protein researchers.

2. Protein separation

2.1. Protein extraction

Protein extraction encompasses a range of techniques designed to isolate proteins from a variety of sources, including physical, chemical, or enzymatic strategies to separate and concentrate proteins from raw materials. The approach used is usually determined by the type of sample, the target protein's properties, and the intended application (Khursheed, Ashfaq, Yousuf, Anjum, & Younis, 2024). Chemical methods are categorized according to the different extraction solvents used (e.g., aqua, alkali, organic solvents, and acids). For example, alkali extraction, recognized as the most conventional method for extracting plant proteins, yields varying extraction rates ranging from 13% to 95%. Notably, performing alkali extraction at high pH levels can achieve protein yields exceeding 90% for oilseeds such as soybeans and rapeseed (Kumar et al., 2021). Physical methods mainly include techniques such as ultrasound, pulsed electric field, microwave, and high-pressure-assisted extraction, which utilize physical forces or forms of energy to disrupt the cellular structure, thereby facilitating the release and extraction of proteins. Each technique has distinct advantages and can be applied to different classes of protein extraction. For example, Kumar et al. (2021) proposed that ultrasonic-assisted extraction is suitable for protein extraction from oilseeds (sunflower, soybean, and peanut), whereas microwave-assisted extraction is more suited to milling industry byproducts. Enzyme-assisted extraction employs enzymes that degrade cell walls to enhance mass transfer and boost extraction rates, proving especially effective for isolating high-quality proteins (Naseri, Marinho, Holdt, Bartela, & Jacobsen, 2020). This approach excels at extracting high-quality proteins from plant sources and food processing by-products and can improve the solubility, emulsification, and foaminess of proteins, making it ideal for crafting functional foods. Furthermore, a variety of different extraction methods can be combined according to the nature of the product to further improve the efficiency and yield of protein extraction.

2.2. SDS-PAGE

SDS-PAGE is widely used in the separation of protein molecules, determination of the relative molecular weight of proteins by size, monitoring protein purification, as saying the purity of samples, and characterization of multimeric proteins. The schematic diagram is shown in Fig. 1A. This method typically utilizes charged yet gentle detergents, specifically SDS, to disrupt secondary and non-disulfide-linked

Table 1
Common methods for measuring protein structure.

Detection method	Proteins	Protein status	Reaction conditions	Measured structure	Reference
SDS-PAGE; DLS; CD fluorescence spectroscopy	Walnut protein isolate	Liquid	Ultrasound	Size; secondary and tertiary structure; molecular weight profile	(Shi et al., 2023)
2D SDS-PAGE, MS	*Cryptostegia grandiflora latex*	Liquid	Purification	Primary structure;	(Freitas et al., 2023)
SEC; FTIR; CD	Lysozyme; BSA; human insulin	Solid	Extrudate	Aggregation; secondary structure	(Dauer, Werner, Lindenblatt, & Wagner, 2023)
Chemical analysis; DLS; SDS-PAGE	Soy protein isolate	Liquid	Enzymatic hydrolysis; thermal	Hydrogen bonds; hydrophobic interactions; disulfide bonds; size	(Yuan, Zhou, Niu, Shen, & Zhao, 2023)
FTIR; fluorescence spectroscopy	Soy protein isolate	Solid; liquid	Enzyme crosslinking	Secondary and tertiary structure	(Dong et al., 2020)
SDS-PAGE; CD	Soy protein	Liquid	Heating	Secondary structure	(Ding et al., 2020)
Raman spectroscopy	Beef muscles	Solid	Freeze–thaw cycle	Secondary structure	(Zhu et al., 2023)
FTIR; CD; XRD	Tartary buckwheat protein	Liquid	Combined with polyphenol	Crystalline structure; secondary and tertiary structure	(Li et al., 2023)
Fluorescence spectroscopy; FTIR;	Soy protein isolate	Emulsion gels	Homogenization	Tertiary structure	(Zhang, Shen, Xue, Yang, & Lin, 2023)
NMR	Whey protein	Liquid	In aqueous solution	Secondary, tertiary and quaternary structure	(Edwards & Jameson, 2020)

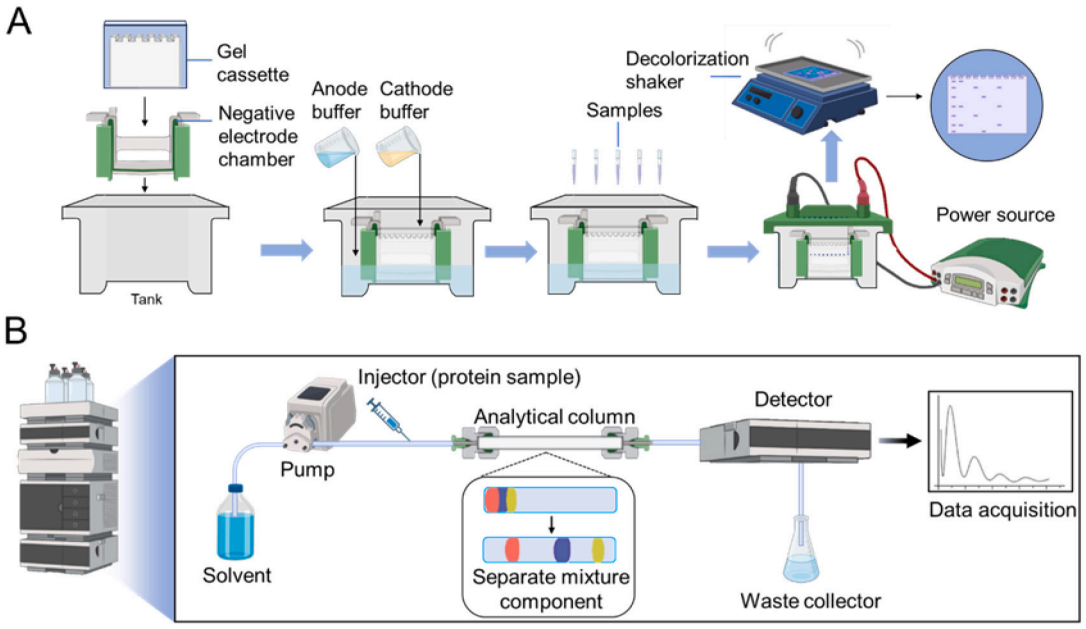


Fig. 1. The schematic diagram of (A) SDS-PAGE and (B) SEC.

tertiary protein structures. It induces a charge shift on the proteins, causing all protein molecules bound to the detergent to migrate in the same direction. This separation enables the estimation of molecular weights. (Brunelle & Green, 2014). The most commonly used SDS-PAGE gel system is the Laemmli system, which is a discontinuous gel containing different polyacrylamide concentrations and pH values. The upper stacking gel, with a pH of 6.8, has a lower percentage of polyacrylamide. While the lower separating gel, with a pH of 8.8, has a higher percentage of polyacrylamide. Based on the Laemmli system, SDS-PAGE is generally classified into two types. SDS-PAGE performed in the presence of a reducing agent like 2-mercaptoethanol (2-ME) or dithiothreitol (DTT), which further reduces disulfide linkages, unfolds proteins and breaks up quaternary protein structures (oligomeric subunits), is known as reducing SDS-PAGE (Roy & Kumar, 2014). On the contrary, non-reducing SDS-PAGE is free of reducing agents. With SDS-PAGE, proteins are analyzed under both reducing and non-reducing conditions, which can provide valuable information about disulfide-cross-linked subunits (Gallagher, 2012). Boye, Roufik, Pesta, and Barbana (2010) used reducing SDS-PAGE (in the presence of 2-ME), and observed missing bands, indicating that these peptides were cross-linked through disulfide bonds. Li et al. (2020) used reducing SDS-PAGE to analyze deformed soy proteins and revealed that the AB and A5B3 subunits are nearly entirely diminished, suggesting their significant involvement in the formation of aggregates through sulfhydryl/disulfide bond exchanges.

Commonly, a 15% polyacrylamide gel is utilized as the separation gel for SDS-PAGE. This concentration facilitates the separation of proteins with relative molecular masses ranging from 10 to 100 kDa (Walker, 2002). The choice of polyacrylamide gel percentage depends on the size of the protein being separated. Gallagher (2012) provided an approximate guide for denaturing discontinuous electrophoresis: 5% acrylamide for 25–200 kDa, 10% acrylamide for 14–200 kDa, 15% acrylamide for 14–66 kDa and 5%–20% acrylamide for 10–300 kDa. However, traditional SDS-PAGE encounters challenges when it comes to separating small peptides and proteins below 10 to 15 kDa due to their low affinity with SDS, which reduces resolution. This can be resolved with the use of gradient gels, or by using different electrophoresis conditions, such as tricine-sodium dodecyl sulphate-polyacrylamide gel electrophoresis (Tricine-SDS-PAGE). The Tricine-SDS-PAGE system provides a simple and efficient alternative for separating low-molecular-mass proteins/peptides in the range 1–100 kDa with high resolution, particularly for proteins smaller than 30 kDa (Schägger, 2006). Compared with SDS-PAGE, Tricine-SDS-PAGE consists of three gels (stacking, spacer, and separating gel), and includes urea in the separating gel. In addition, tricine-SDS-PAGE requires the use of two running buffers: a cathode buffer and an anode buffer consisting of different pH and different concentrations of Tris, tricine, and SDS (Yim, Ahn, Kim, & Yun, 2002). Uniform acrylamide Tricine-SDS-PAGE can cover smaller molecular weight – for example, 10% gels and 16% gels cover the ranges 1–100 kDa and 1–70 kDa, respectively (Schägger, 2006). Claeys, De Smet, Balcaen, Raes, and Demeyer (2004) undertook a quantitative analysis of fresh meat peptides using tricine-SDS-PAGE and found that with time, the peptide concentration increased significantly in the molecular weight (MW) range 3–17 kDa. Ying et al. (2015) performed a tricine-SDS-PAGE analysis of soybean oil bodies with SDS/protein mass ratio of 1.52/1 to avoid protein loss caused by defatting.

Currently, different electrophoresis systems are used for detection of protein of different mass ranges. For proteins with a molecular weight of $<$30 kDa, Tricine-SDS-PAGE is the preferred method, while for proteins exceeding 30 kDa, SDS-PAGE is utilized. When dealing with monomeric structures, reduced SDS-PAGE will usually be selected to assess protein purity and molecular weight. For determination of the molecular weight of protein subunit structure, reducing SDS-PAGE is always used, as protein dimers and multimers are degraded into monomers and may even degrade into lower subunit structures (such as antibody heavy chain and light chain).

2.3. 2D-SDS-PAGE

While gel electrophoresis is widely used for protein separation, purification and identification, SDS-PAGE is often inadequate for effectively providing accurate information for complex protein mixtures (Walker, 2002). The emergence of two-dimensional gel electrophoresis (2D-DE) presents a possible solution to this problem. The 2D-DE is a well-known tool for the study of protein mixtures, and can classify similar molecular weight based on two inherent properties of proteins (McDonald & Sugrue, 2007). Generally, the 2D-DE couples two independent separation systems (isoelectric focusing and SDS electrophoresis) to separate proteins based on their different charges and sizes (Issaq & Veenstra, 2008)—this is known as two-dimensional sodium

dodecyl sulphate polyacrylamide gel electrophoresis (2D SDS-PAGE). In the first dimension, proteins are separated by isoelectric focusing (IEF) according to their isoelectric point. When the protein reaches their isoelectric point, the protein is highly concentrated, as the electric field no longer affects the protein (Friedman, Hoving, & Westermeier, 2009). The IEF gels are incubated in SDS buffer and in the second dimension, proteins are separated by SDS-PAGE according to the molecular weight of proteins. Generally, in the first dimension, the IEF gel would be cast into a glass tube, and then extruded and loaded onto a second-dimensional slab gel (Rabilloud, Chevallet, Luche, & Lelong, 2010). However, this is a rather difficult process, as a large amount of non-linear deformation of the tubular gel may occur, with separation of the sample taking up to 16 h. At present, one common improvement that is often utilized is the inclusion of an immobilized pH gradient (IPG) in the first dimension (Issaq & Veenstra, 2008). The IPGs enable the formation of stable and reproducible pH gradients, capable of focusing acidic and basic proteins on a single gel. IPG strips can be divided into three ranges: wide range (e.g., pH 3.0–10.0), medium ranges (e.g., pH 4.0–7.0, pH 7.0–11.0) and narrow ranges (e.g., pH 5.0–6.0, pH 5.5–6.5), with length between 7.0 cm and 22.0 cm (Friedman et al., 2009). This spread allows proteins with similar isoelectric point values to be separated with higher resolution. Due to its resolution and sensitivity, this method facilitates the simultaneous identification of multiple protein complexes, as well as the ability to analyze and study the abundance of micron-type proteins in a small number of samples (Lilley, Razzaq, & Dupree, 2002). Panchal, Hati, & Sakure (2020) detected 39 protein spots, roughly 10 kDa to 51 kDa, in goat milk proteins by isolation with 2D SDS-PAGE. Raikos, Hansen, Campbell, and Euston (2006) used 2D SDS-PAGE as a method for visualizing and identifying glyco-isoforms of egg proteins (egg white and egg yolk proteins). McDonald and Sugrue (2007) reported the 2D SDS-PAGE method (immobilize strip gels 7 cm, pH 3.0–10.0) for characterizing the maturity and heterogeneity of respiratory syncytial virus fusion protein. 2D SDS-PAGE also helps in the confirmation of protein-protein interactions, as well as highlighting almost undetectable complexes in the crude extract by purifying multiple protein complexes (Bernarde et al., 2010; Lasserre et al., 2010).

Compared with one-dimensional SDS-PAGE, the 2D gel has the advantage of separating a large number of complex protein complexes, coupled with higher resolution. The 2D technology has relatively complicated operation steps, where the sample running time is longer than SDS-PAGE, and sample preparation requirements are stricter. In addition, proteins of extreme MW (<10 KDa or > 200 KDa) and pI (< 3 or > 11), very hydrophobic proteins (e.g., integral membrane proteins with multiple trans-membrane domains), and low-abundance proteins are typically difficult to resolve or detect on 2D gels (Friedman et al., 2009).

2.4. Size-exclusion chromatography

Size exclusion chromatography (SEC) is a chromatographic technique that separates substances according to their molecular weight and hydrodynamic volume. The schematic diagram is shown in Fig. 1B. In SEC, analytes are separated using a column packed with porous polycarbonate. Larger molecules do not enter the pores as readily as smaller molecules and therefore elute from the column faster, resulting in different separation elution bands (Striegel, 2023). The molecular weight and size distribution of the sample are determined by comparing the separation elution bands against a calibration curve. This curve is generated by calibrating the chromatography column with standards of known molecular weight, thereby establishing a relationship between elution volume and molecular weight. The same is true for the identification of aggregates, which are usually larger than the monomeric form of the molecule and will therefore elute earlier from the column. The presence of dimers, oligomers or higher-level polymers can be recognized by comparing the elution profile of the sample with the elution profile of known monomers (Some, Amartely, Tsadok, & Lebendiker, 2019). The separation elution band can be determined based on the resolution, which is usually optimized by column parameters such as particle size, flow rate, column length, and number of columns. In general, the pore size of 150–200 Å is suitable for proteins with molecular weights in the range of 15–80 kDa, while the 200–300 Å pore size column is usually used for proteins of about 150 kDa. For very large plant proteins (MW > 200 kDa), columns with pore sizes of 500–1000 Å presents as the best choice (Fekete, Beck, Veuthey, & Guillarme, 2014). Parameters such as the column length, internal diameters, pore size, and flow rate commonly used in SEC for some food proteins are summarized in Table 2.

Modern analytical SEC platforms are often combined with multiple detectors for more detailed and extensive characterization of simple proteins or protein complexes. Ultraviolet (UV) has always been the main method of detection and is often used in combination with SEC to assess protein aggregation. SEC coupled with “on-line” static laser light scattering, refractive index, and UV detection were used by Folta-Stogniew (2006) to determine oligomeric states of proteins. Multi-angle laser light scattering (SEC-MALLS) allows for detection of the molar mass and radius of gyration of macromolecules (Bean & Lookhart, 2001). The use of SEC-MALLS to investigate aggregation of soy protein isolate (SPI) induced by high pressure treatment was systematically examined by Tang and Ma (2009). This method can determine the molecular weights and distribution of the soluble aggregates formed by different pressure treatments. Moreover, SEC-MALS is capable of confirming and analyzing the high MW aggregates, and SEC-fluorescence allows detection of low levels of aggregates (Hong, Koza, & Bouvier, 2012).

Compared with traditional SDS-PAGE for protein analysis, the main advantage of this approach is the mild mobile phase conditions (e.g., phosphate buffered saline), which allow the characterization of proteins with minimal impact on conformational structure and local environment (Burgess, 2018). In addition, SEC can handle a very wide range of molecular weights (about 0.1–10,000 kDa) and almost any type of protein polarity. However, proteins can only be separated when their molecular weights differ by at least 50% (Möller, Albert, & Atze, 2022). When used in combination with multiple detectors, SEC can be used as the main technology for protein analysis due to its ease of use, good automation, and multiple functions.

3. Primary structure

3.1. Mass spectrometry

The primary structure of protein is a linear sequence of amino acids, which are connected through peptide bonds (Lehrman, 2017). The primary structure of a protein is directly responsible for determining its secondary, tertiary, and other higher-level structures, which in turn influence its bioactive functions and physicochemical characteristics. Therefore, it is crucial to examine and characterize the structure of proteins.

Mass spectrometry has become a pivotal analytical tool in biological research at the molecular level, experiencing a significant surge in both instrument development and applications, especially for analyzing large biomolecules like proteins and peptides (Jonsson, 2001). A MS consists of an ionization source, a mass analyzer and a detector. It works by ionizing the sample components, producing ions with different mass-to-charge ratios (m/z), which are then accelerated by an electric field into a beam that flows toward the mass analyzer. Here, an electric or magnetic field or filtration is used to separate the ions based on their m/z ratio. Finally, these ions are detected to produce a mass spectrum that relates mass to concentration or partial pressure. MS can be classified into different types based on various ionization and mass analysis modes, facilitating its broad application in analyzing complex samples (Meyer, Fröhlich, Nordhoff, & Kuhlmann, 2022). In protein characterization, electrospray ionization (ESI) and matrix-assisted laser desorption/ionization (MALDI) are the two typically used methods for ionizing proteins.

Table 2
Common parameters of proteins were analyzed using SEC.

Samples	Column				Eluent	Flow rate (mL/min)	Reference
	Length (mm)	Internal Diameter (mm)	Particle size (μm)	Porosity (Å)			
Whey proteins	300	7.8	3	–	50 mM sodium phosphate buffer; pH 7.0	0.5	(Diamante, 2023)
Corn; soybean protein hydrolysates	300	7.8	1.7; 1.8; 5	100; 125; 130	50 mM sodium phosphate buffer; pH 7.0	0.1	(Li et al., 2023)
Lentils protein	300	10	8.6	–	0.1 M sodium phosphate buffer; pH 7.0	0.5	(Shrestha, Hag, Haritos, & Dhital, 2023)
Pea protein isolates	300	4.6	5	100	0.1 M sodium phosphate buffer; pH 7.0	–	(Sajib, Forghani, Vate, & Abdollahi, 2023)
Pea protein	300	10	8.6	–	0.01 M sodium phosphate buffer; pH 6.8	0.5	(Assad-Bustillos et al., 2023)
Ovalbumin	300	7.8	3.5	150	0.01 M sodium phosphate buffer; pH 7.0	1.0	(Huang et al., 2023)
Pea seed protein	60	26	50	–	50 mM phosphate buffer containing 0.15 M NaCl; pH 8.0	0.3	(Olías, Rayner, Clemente, & Domoney, 2023)
Whey protein concentrates	60	7.5	10	–	0.5 M Tris-HCl buffer; pH 7.0	0.6	(Gantumur et al., 2023)
Amaranth protein	300	10	13	–	Water/acetonitrile/trifluoroacetic acid (TFA): 69.9/30/0.1 (*v*/*v*) solvent	0.5	(Beaubier et al., 2023)

In MALDI, analytes are ionized through co-crystallization with a matrix and laser irradiation, leading to proton transfer reactions, while in ESI, ionization occurs by dispersing the sample into an aerosol and desolating it through coulombic fissions, resulting in the emergence of desolated analyte ions (Kebarle & Peschke, 2000; Moon, Yoon, Bae, & Kim, 2015). After ionization, specific ions (typically peptides) undergo further fragmentation into a set of ions that reflect sections of the peptide in methods like tandem mass spectrometry (MS/MS). The resulting mass spectrum from MS/MS reveals a sequence of peaks, each corresponding to the mass of the fragmented ions. These peaks signify fragments of the original peptide, and the mass differences between adjacent peaks can reveal the presence of distinct amino acids, given their unique masses. The inferred peptide sequences or mass fingerprints are matched against theoretical mass patterns derived from protein sequence databases, facilitating the identification of their amino acid sequences (Hunt, Yates 3rd, Shabanowitz, Winston, & Hauer, 1986).

Workflows for identifying amino acid sequences typically fall into two categories: "top-down" and "bottom-up" approaches. In traditional bottom-up protein analysis methods, proteins are first chemically or enzymatically digested into peptides (1–3 Kda) with predictable termini and the resulting peptides are analyzed by MS after separation by liquid chromatography (LC) (Van Schaick, Haselberg, Somsen, Wuhrer, & Domínguez-Vega, 2022). Bottom-up approaches typically attain 5% to 70% sequence coverage, often employing trypsin for digestion. To enhance sequence coverage, a strategy involving the use of multiple proteases, each with distinct cleavage patterns, can be implemented. Additionally, the introduction of reagents designed to aid in protein unfolding may further improve coverage (Leurs, Mistarz, & Rand, 2015). MALDI-MS is often the preferred method for primary structural analyses in bottom-up MS due to its ease of use, low susceptibility to contamination, and relatively high tolerance to additives (Maus, Mignon, & Basile, 2018). Meanwhile, the coupling of chromatographic systems with ESI-MS is widely used for advanced peptide mapping experiments, as ESI-MS can be easily combined with LC systems for separation and automated processing, thus increasing sequence coverage and enabling higher throughput analysis, especially for complex mixtures (Takahashi, Isobe, Desiderio, & Nibbering, 2007). However, since protein identification is typically inferred from a limited number of characteristic peptides, the limited sequence information derived from small peptides is not always sufficient to identify specific protein forms. The identification of protein isoforms and (post-translational modifications) PTMs without prior knowledge is extremely limited (Dupree et al., 2020). On the other hand, the top-down approach bypasses the protein hydrolysis step and couples' intact proteins with MS/MS, allowing in-depth analysis of protein modifications and ab initio sequencing of proteins. The top-down MS is better suited for protein isoform studies and does not suffer from losses and/or degradation associated with enzyme-digested, preserving the destabilizing structural features that are destroyed in bottom-up MS (Greer et al., 2018). In MS/MS, after the ionization process, fragmentation of the protein backbone can lead to the production of terminal fragment ions through various methods. Collision-based fragmentation methods such as collision-induced dissociation (CID), infrared multiphoton dissociation (IRMPD), and post-source decay (PSD) predominantly cleave the CO-NH backbone bond, efficiently generating *b* and *y* fragment ions. On the other hand, nonergodic methods like Electron Capture Dissociation (ECD) and Electron Transfer Dissociation (ETD) facilitate localized, rapid cleavages ($<10^{-12}$ s) of the main chain covalent bond (NH-CHR), primarily producing *c* and *z* fragment ions (Armirotti & Damonte, 2010). Recently, ultraviolet Photodissociation (UVPD) has emerged as an effective method for fragmenting large proteins and protein complexes. UVPD typically involves the cleavage of the alkyl-carbonyl bond, leading to the generation of *a* and *x* fragment ions (Zenaidee et al., 2021). The amino acid sequence of a protein can be ascertained by analyzing various fragment ions, such as *a, b, c* (N-terminal), and *x, y, z* (C-terminal) ions, since each type of ion offers information on the precise place of a protein chain break. In particular, the fragments generated by nonergodic electronic matrix dissociation (e.g., ECD and ETD) are capable of identifying the precise location within the protein sequence where these chemical alterations take place. For example, Zhang and Ge (2011) utilized top-down MS/MS with ECD to identify two potential monophosphorylation sites on the positional isoform of monophosphorylated human cardiac troponin I (cTnI): the well-documented Protein Kinase A (PKA) site at Serine 22 (Ser22) and a newly discovered site at either Serine 76 (Ser76) or Threonine 77 (Thr77). While some technical challenges remain, developments in top-down protein analysis over the past few years have improved its ability to unambiguously identify, characterize and quantify thousands of protein forms at high throughput (Révész et al., 2023).

In addition, top-down and bottom-up mass spectrometry methods need to be used in a complementary manner for optimal protein identification and characterization. Bottom-up MS is excellent at identifying proteins and their constituent peptides in complex mixtures, while top-down MS offers precise information about the complete protein, including the PTM and its distribution along the protein backbone. The combination of these two approaches provides a more comprehensive understanding of the proteome, including the identification and characterization of proteins, as well as functional state determination. Moreover, sophisticated mass spectrometry techniques can be used to probe higher order structures of proteins, such as secondary, tertiary and

quaternary structures, by using appropriate sample preparation, ionization conditions and instrumentation. This is often achieved by techniques such as ion mobility spectrometry (IMS) and cross-linking mass spectrometry (XL-MS), which provide insights into the spatial arrangement of atoms within a molecule as well as the interaction sites between protein subunits (Chavez, Wippel, Tang, Keller, & Bruce, 2021). MS can also be used to investigate protein-ligand interactions. For example, affinity mass spectrometry and hydrogen/deuterium exchange MS (HDX-MS) can identify binding sites, binding affinities, and conformational changes caused by ligand binding (Chen et al., 2019).

3.2. Edman degradation

In addition to mass spectrometry, Edman degradation is a widely used method for primary structure identification of proteins. The technology for determining amino acid sequences of proteins is historically rooted in degradative sequencing approaches, and the most critical recent development is the emergence of automated procedures for degradation by the phenyl isothiocyanate (PITC) method. The reagent PITC couples with the terminal α-amino group of peptides or proteins to form phenylthiocarbamyl derivatives. In the presence of a strong acid, the initial amino acid can be detached from the protein, resulting in a protein that has a new amino terminus and is shorter by one amino acid. By repeating this chemical reaction and utilizing chromatographic analysis, the primary sequence of the protein could be obtained. Edman degradation represents a cost-effective technique for sequencing amino acids, making it ideal for obtaining long degradations on proteins or peptides containing between 60 and 150 residues. The protein sample requirements are comparatively minimal, enabling the analysis of minimal quantities of protein. However, for longer protein sequences, this approach grows more time-intensive and complicated. Moreover, with an increase in cycle repetitions, the accumulation of errors makes accurately determining long sequences challenging (Floyd & Marcotte, 2022; Niall, 1973a, 1973b).

The evolution of the Edman degradation method began in the early 1950s and was initially used for the sequencing of insulin. It can be performed without breaking peptide bonds between other amino acid residues. However, this method only identifies amino acid sequences in peptides. In the late 1990s, researchers combined MS and HPLC techniques as an alternative to Edman degradation to efficiently identify and analyze protein molecule (Suhaib Al Huq, Raja, & Oviya, 2024). Although the Edman degradation method is gradually being replaced due to the development of mass spectrometry. Combination of MS and Edman sequencing with database search tools remains important for identifying unknown proteins. The Edman degradation method is now widely used for product development in the food, chemical and pharmaceutical industries (König, Obermann, & Eble, 2022). Maky and Zendo (2023) used the Edman degradation method as a means of identifying novel bioactive peptides in frozen chicken breast hydrolysates and analyzing their potential for application as bio preservatives. Sangiorgio et al. (2022) employed three hydrolysis approaches to degrade soy isolate protein. After purification, a novel peptide was identified from the hydrolysis products through HPLC and Edman degradation, which possesses the potential to be used as a functional food ingredient for the prevention of chronic diseases. In addition to technology association, Borgo and Havranek (2015) has proposed a Edmanase substitution method for the harsh Edman reagents, which reduces the inactivation of fluorophores on peptides by the harsh Edman reagents. Although sequencing of full-length proteins is still not achieved, this method currently allows for efficient quantification of diagnostic peptides in samples of reduced complexity.

4. Secondary structure

4.1. Fourier transform infrared spectroscopy

FTIR spectroscopy is a method that combines the mathematical processing of Fourier transform, computer technology, and infrared spectroscopy to characterize protein structure and chemical properties. By absorbing the energy of chemical bonds, a spectrum is generated in the range of 400–4000 cm^{-1}, which provides information on the relevant molecular level (functional groups, bond types and molecular conformations) and protein secondary structure content (Kuan, Bhat, Patras, & Karim, 2013). The schematic diagram is shown in Fig. 2A. The infrared spectrum is obtained by processing the degree of absorption of different frequency beams by the sample in a short time. The interpretation of the spectrum is mainly based on the vibration of the repeating unit of the structure. Peptide and protein repeat units generate nine characteristic infrared absorption bands, of which the amide I and II bands are the two most prominent vibrational bands in the protein backbone (Kong & Yu, 2007). However, the most sensitive spectral region is the amide I band with a vibration absorption region of 1700–1600 cm^{-1} (Fig. 2A), which is related to the existence of α-helix, anti-parallel and parallel β-sheet, and random-coil structures. Furthermore, it is usually thought to be caused by the C═O stretching vibration of the peptide bond (Movasaghi, Rehman, & Ur Rehman, 2008). In a study by Jiang et al. (2019), the absorbance of SPI-anthocyanins at amide I band was decreased, indicating changes in the secondary structure of protein. Dong et al. (2020) reported that the analysis of the amide I band in FTIR showed an increase in the α-helix content along with a decrease in the random coils content for protein coated $CaCO_3$ microparticles without T-Gase treatment. Wen, Zhao, Jiang, and Sui (2024) discovered that increasing the SPI content in zein-SPI composite nanoparticles caused a large red shift in the amide I band peak, implying that the secondary structure of zein had changed.

The types and quantities of protein secondary structures are usually obtained through enhanced resolution and multivariate analysis techniques. However, factors such as the size of the molecule and the Doppler effect can cause different absorption bands to overlap (Al-Mbaideen & Benaissa, 2011). In order to more accurately obtain the content of protein secondary structure, resolution is improved and

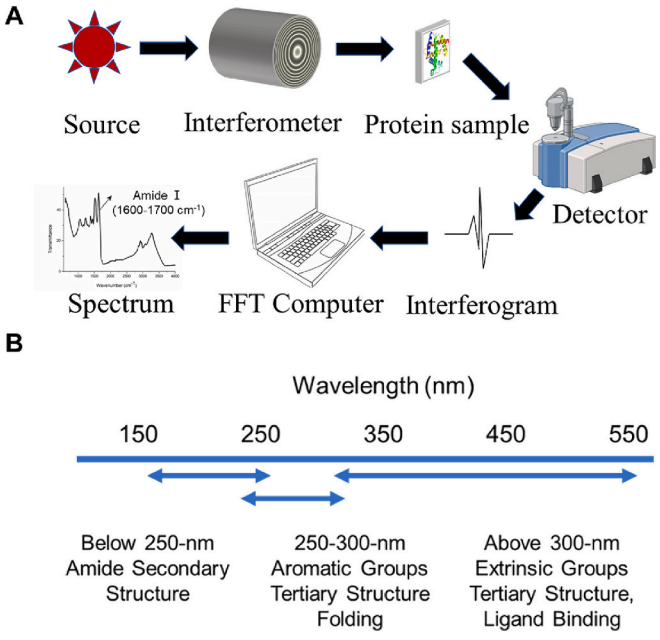


Fig. 2. (A) The schematic diagram of FTIR and (B) CD spectral regions and contributing chromophores in proteins.

multivariate analysis is often used to analyze the spectrum. Common methods to improve the resolution are frequency inversion convolution (FSD) and differential methods. FSD mainly involves determining the number, position, and relative amplitude of peaks, while differential methods can effectively eliminate spectral background and baseline drift. Multivariate analysis provides information by building an empirical model based on the obtained spectrum. Conventional techniques include partial least squares and principal component regression.

From a practical perspective, FTIR is the most widely used analytical technique. Almost all compounds have specific absorption characteristics in the infrared region, and the spectrum of the sample can be recorded in a very short time (1 s) (Bacsik, Mink, & Keresztury, 2004). FTIR has the advantages of having simple sample preparation procedures, a high spectral quality, low cost, and fast speed. However, the infrared spectral characteristic band of proteins is susceptible to the conformational changes of proteins. Therefore, FTIR is often used to detect the conformation of proteins in a dry state, in aqueous conditions and deuterium oxide solutions, and to derive the structure and dynamic characteristics of proteins. In addition, FTIR is often recommended for use with CD spectroscopy to improve the accuracy of protein secondary structure content (Schwaighofer, Alcaráz, Araman, Goicoechea, & Lendl, 2016).

4.2. Circular dichroism spectroscopy

Circular dichroism spectroscopy (CD) includes electronic transition CD, vibration CD, fluorescence detection CD and ultraviolet CD. It is mainly used to study chiral molecules, including the recognition of secondary structures, and the study of the conformational changes in tertiary structures of biological monomers and polymer molecules, such as proteins, nucleic acids, and carbohydrates (Phillips-Jones & Harding, 2019).

In the case of CD, planar polarized light can be regarded as being composed of two equal-sized circularly polarized light components. The interaction of light with asymmetric molecules generates elliptically polarized light due to differences in absorption and refractive index. This phenomenon results in CD signals (Greenfield, 2006). The CD signal appears only where it can absorb radiation, such that the spectral band is reflected on the different structural characteristics of the molecule. For instance, α-helical structures exhibit negative bands at 222 nm and 208 nm and positive bands at 193 nm. β-sheet structures display a negative band at 216 nm and a positive band ranging between 195 and 200 nm. Disordered proteins show negative bands around 195 nm and shallow positive bands above 210 nm (Wei, Thyparambil, Latour, & Proteomics, 2014). The spectrum of a protein is the cumulative result of its conformational element spectrum, enabling CD to estimate secondary structure content. CD spectral regions and contributing chromophores in proteins are shown in Fig. 2B. In a study by Zhang, Chen, Qi, Sui, and Jiang (2018), CD was employed to examine the structural changes of SPI during its interaction with anthocyanin-rich black rice extracts, leading to the determination of changes in the secondary structure of SPI. In addition, CD can be used to investigate different proteins, macromolecular complexes, and the structure, dynamics, and similarity of protein families (Wallace, 2009).

As a core method for studying protein structure, CD has strict sample requirements. Sample preparation requires close attention to be paid to the choice of the cuvette, buffer, and protein concentration. In general, high-transparency quartz cuvettes are used, of which round cuvettes are used for room temperature measurement (Pelton & McLean, 2000). The CD spectral buffer must be transparent and not contain any optically active materials. Numerous buffer compounds and salts exhibit significant absorption in the far-UV region, potentially impacting the outcomes of CD analyses. It is advisable to either lower their concentrations or refrain from employing them. The wavelength ranges appropriate for CD measurements across various buffer compounds and salts are detailed in Table 3 (Micsonai, Bulyáki, & Kardos, 2021). For high-quality data, the optical absorbance of the sample should be 0.87 (where the signal-to-noise ratio is the largest). However, lower concentrations can reduce turbidity and concentration-dependent sample aggregation. Thus, the content of protein in the sample used for CD spectroscopy must be higher than 95%, and the concentration of the sample used to measure the secondary structure is usually between 0.005 and 5 mg/mL (Greenfield, 2006).

Table 3
Absorption of various buffer compounds and salts in the far-UV region (Micsonai et al., 2021).

Compound	180 nm	190 nm	200 nm	210 nm	No absorption above
$NaClO_4$	0	0	0	0	170 nm
NaCl	>0.5	>0.5	0.02	0	205 nm
Na_2HPO_4	>0.5	0.3	0.05	0	210 nm
NaH_2PO_4	0.15	0.01	0	0	195 nm
NaOH	>2	>2	>2	>0.5	230 nm
DMSO (0.1%)[b]		>3	>3	1.8	233 nm
Urea (1 M)[b]		>3	>3	0.29	227 nm
Citric acid[b]	>2.5	0.45	0.22	0.21	240 nm
Boric acid, NaOH	0.3	0.09	0	0	200 nm
TRIS	>0.5	0.24	0.13	0.02	220 nm

In recent years, a notable trend is the use of Synchrotron radiation CD (SRCD), which has more advantages than the traditional CD spectrum. SRCD can measure fewer samples due to its higher signal-to-noise ratio as compared to traditional measuring instruments. It is also possible to obtain the characteristics of the secondary structure in the protein mixture more accurately, especially to distinguish the proline helix from the disordered protein structure (Yoneda, Miles, Araujo, & Wallace, 2017). Deller, Carter, Zampetakis, Scarpa, and Perriman (2018) studied the structural properties of antifreeze protein III by chemical cations in extreme environments through far-UV SRCD and other techniques. SRCD can also be used to detect the formation of protein ligands. Hussain, Longo, and Siligardi (2018) used high-photon flux diamond B_{23} beamlines to perform UV denaturation experiments on SRCD and presented a new method for screening protein-ligand binding.

4.3. Raman spectroscopy

Raman spectroscopy is a relatively mature molecular structure spectroscopy technology based on the Raman scattering effect that mainly uses the interaction between light and matter. Raman spectroscopy is capable of providing information about the molecular vibrational energy level (vibration lattice energy level) and rotational energy level structure in the spectrum, as well as providing insights into other low-frequency mode information (Nemecek, Stepanek, & Thomas Jr, 2013). The schematic diagram is shown in Fig. 3A. The typical Raman spectrometer usually consists of a light source, sample equipment, filter, interferometer, and detector. A charge-coupled device detector captures the filtered Raman signal to produce a spectrum (Vankeirsbilck et al., 2002). The signal-to-noise ratio, resolution, and stability of the instrument will affect the accuracy of the data. At present, Raman spectroscopy has overcome the problems of fluorescence, reduced sensitivity, and poor reproducibility. Additionally, deep ultraviolet resonance Raman spectroscopy (UVRRS) has become recognized as a valuable tool for protein structure characterization (Shashilov, Sikirzhytski, Popova, & Lednev, 2010).

Raman spectroscopy is often used to assess the secondary structure of proteins due to amide vibrational bands being sensitive to structure. The method uses a 206.5 nm continuous wave laser to directly excite the UVRRS in the $\pi \rightarrow \pi^*$ transition of the peptide bond (Moon et al., 2015). The spectral results are determined by the vibrations of the amide group, and the vibrational frequency, Raman cross-section and bandwidth of the amide are in turn closely related to the secondary structure of the protein. Therefore, Raman spectroscopy obtains information about the

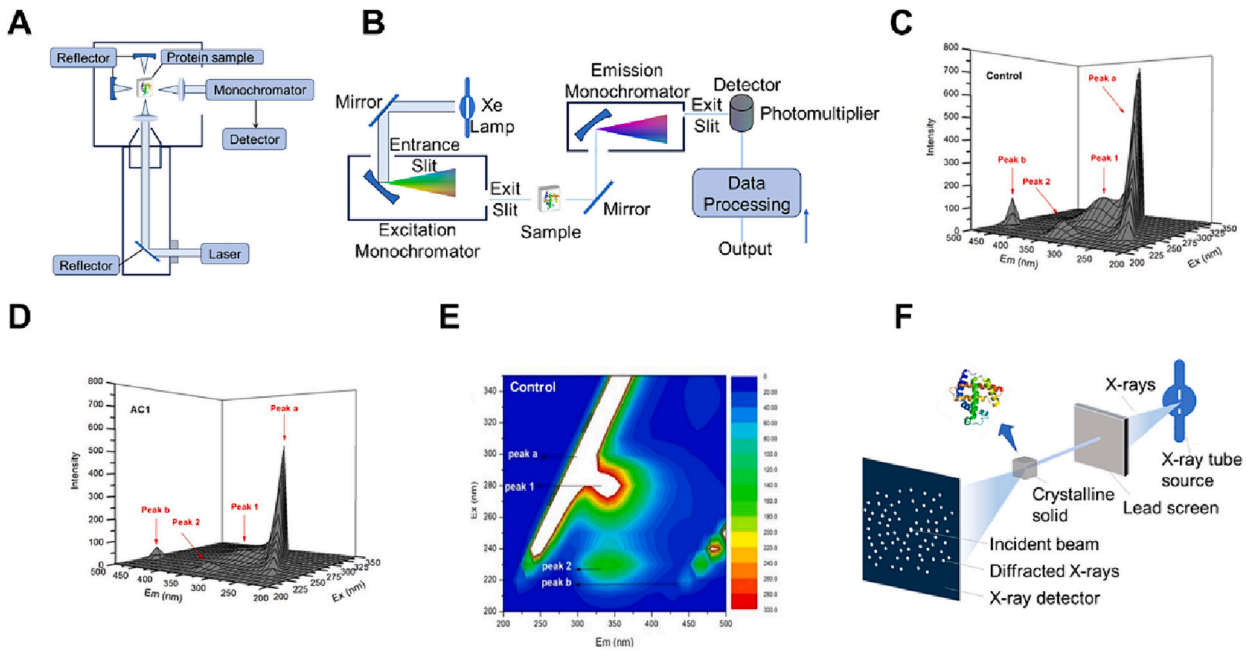


Fig. 3. The schematic diagram of (A) Raman spectroscopy, (B) Fluorescence spectroscopy, (C) Three-dimensional fluorescence spectra of SPI and (D) SPI-anthocyanin conjugates, (E) Three-dimensional fluorescence contour map (Jiang et al., 2019; Sui et al., 2018), and (F) XRD.

secondary structure mainly using analyzed peptide amide bands (amide bands I, II, and III). Yin, Jin, and Zhang (2018) utilizing Raman spectroscopy to examine soy globulin, findings indicated a negative correlation between the surface hydrophobicity of soybean 11S globulin and the presence of α-helix and β-sheet structures. In addition, Raman can be used to identify protein local conformational information. The side chains of particular specific amino acid residues produce distinctive Raman scattering signals, allowing the inference of local conformational changes in proteins. By using Raman spectroscopy, Yu et al. (2021) detected that Trp and Tyr residues in SPI after electrolysis tended to be exposed from their original embedded states. And based on the characteristic vibrational bands of S—S in the range of 500–550 cm^{-1}, it was determined that the side chain conformations changed from gauche–gauche–gauche mode to gauche–gauche–trans mode and trans–gauche–trans mode. The measurement of Raman spectroscopy does not require special sample preparation and can be collected directly in the aqueous solution using a glass capillary. In addition, it can detect proteins in different states: fibers precipitated from dilute aqueous solutions, amorphous aggregates, solids and crystals (Maiti et al., 2004).

In the analysis of the secondary structure method, Raman spectroscopy, infrared spectroscopy, and CD complement each other. FTIR provides information about the structure of molecules by measuring their vibrational frequencies. It is especially sensitive to the vibrations of polar bonds, making it ideal for identifying specific functional groups of proteins. However, it requires high sample purity. CD is primarily used to measure the optical activity of chiral centers in proteins and is very sensitive to the overall folded state of the protein structure, providing direct information about the ratio of the secondary structure of proteins. CD requires sample concentration and transparency and can be affected by solvents and other absorbing substances (interferences may occur when assessing proteins having a large number of aromatic amino acid residues). Raman spectroscopy relies on changes in the vibrational energy levels of molecules and is particularly sensitive to nonpolar bonds, so it can complement the limitations of FTIR in identifying nonpolar bonds in protein (Pelton & McLean, 2000).

5. Higher order structure

5.1. Fluorescence spectroscopy

Fluorescence spectroscopy has long been used as a very important and powerful method for extensively studying protein structure transitions, dynamics, protein conformations and interactions of protein (Wang et al., 2017). Unlike FTIR, CD, and other spectroscopic techniques, fluorescence spectroscopy is a probing method that relies on changes in the fluorophore microenvironment to characterize the exposure of amino acid hydrophobic groups, protein folding and unfolding, rotational diffusion rate of proteins, and inter-protein distances (Raghuraman, Chatterjee, & Das, 2019). The schematic diagram is shown in Fig. 3B. The fluorophore is a fluorescent molecule or substructure that emits light after absorbing ultraviolet or visible light. It absorbs energy in the form of light at a specific wavelength and releases energy in the form of light at a higher wavelength. Proteins contain three types of fluorophores: intrinsic, co-enzymic, and extrinsic. The intrinsic fluorophore is a naturally fluorescent amino acid residue of protein such as tyrosine (Tyr), tryptophan (Trp) and phenylalanine (Phe). Among the three fluorescent amino acid residues, Trp is the most highly fluorescent and widely used, followed by Tyr. Phe is the least fluorescent and is rarely used for protein research (Ladokhin, 2000). The application of Tyr and Phe is limited to Trp-free proteins, and the fluorescent release of Tyr is usually stronger (Chattopadhyay & Raghuraman, 2004). When three fluorescent amino acid residues are present in a protein, the fluorescence spectrum of Tyr and Phe residues appears at wavelengths below 290 and 270 nm, while pure Trp emission can be observed at wavelengths above 295 nm (Eftink, 1991).

The influence of the molecular environment on the fluorophore is important, and the polarity and non-polarity of the solution will change the maximum emission of Trp fluorescence spectrum. The unfolding of the protein changes the Trp molecular environment from a hydrophobic to aqueous environment with increased polarity, thereby shifting Trp fluorescence emission maximum toward a longer wavelength (red shift). The Trp environment moving from an aqueous to hydrophobic environment causes blue shift (shift to a shorter wavelength) of maximum emission intensity. Due to the high sensitivity of Trp to its local environment, the intrinsic tryptophan fluorescence of the protein provides

information about its structure and kinetics (Hellmann & Schneider, 2019). Zhang et al. (2023) used Synchronous fluorescence spectra to obtain information about the Trp microenvironmental change from SPI-hemin solutions, in order to obtain information about SPI tertiary structure.

Fluorescence spectroscopy is also extensively used to investigate protein-ligand interactions. Numerous studies on the interaction between polyphenols and proteins have employed fluorescence spectroscopy, revealing that the addition of polyphenols quenches the fluorescence of proteins. The Stern–Volmer equation is often utilized to analyze the fluorescence quenching of proteins induced by polyphenols (Kanakis et al., 2011; Zhang et al., 2018). Fluorescence quenching is the decrease in the fluorescence intensity from a fluorophore induced by intra- or intermolecular interactions with a quencher molecule, and can be divided into dynamic quenching and static quenching. In short, dynamic quenching occurs when the quencher collides with the fluorophore, deactivating its excited state and reducing fluorescence intensity. On the other hand, static quenching results from the formation of a non-fluorescent ground state complex between the fluorophore and quenching molecule (Karoui & Blecker, 2011). Combining fluorescence quenching and changes in Trp environment allows for further understanding of the protein tertiary structure. Kanakis et al. (2011) indicated that the quenching of β-Lactoglobulin fluorescence in catechin, epicatechin and epicatechin gallate-protein complexes is due to protein structure folding and the reduced polarity of Trp microenvironment.

The fluorescent coenzyme is a natural biomolecule that is present in a small fraction of proteins, such as dihydronicotinamide adenine dinucleotide, flavin adenine dinucleotide (FAD) and flavin mononucleotide (FMN) (Su, Hung, Huang, & Lin, 2011). Another class of chromophore is extrinsic fluorophore that can covalently and non-covalently bind with protein unique active sites, with the most prominent being 1-anilino-8-naphthalene sulfonate (ANS). The extrinsic probe is combined at a specific site on the protein molecule, with the polarity of the protein site being determined by the characteristics of the probe, thereby revealing the structural characteristics of the protein (Hawe, Sutter, & Jiskoot, 2008). ANS as a hydrophobic probe combined with fluorescence spectroscopy is commonly used in the determination of protein surface hydrophobicity and the characterization of protein structure (Deshpande & Sathe, 2018; Jia, Gao, Hao, & Tang, 2017; Zhu, Lin, Ramaswamy, Yu, & Zhang, 2017).

5.2. Three-dimensional fluorescence spectra

The three-dimensional (3D) fluorescence spectra, also known as excitation-emission matrices or total fluorescence, is a powerful technique that can provide more detailed information about the configuration changes of proteins. Compared with 2D fluorescence, where the fluorescence intensity changes with only one wavelength (excitation wavelength or emission wavelength), the fluorescence intensity of 3D fluorescence changes with the change in both excitation and emission wavelengths (Bortolotti et al., 2016). The 3D fluorescence spectra and contour plots (visual representations of the spectra) provide spectral information, as shown in Fig. 3C - E. Two peaks are common in all spectra: peak (a) represents the Rayleigh scattering peak (excitation wavelength = emission wavelength, $\lambda_{ex} = \lambda_{em}$) and is due to the scattering effect of excitation light in the sample solutions; peak (b) is the second-order scattering peak ($\lambda_{em} = 2\ \lambda_{ex}$) (Zaroog & Tayyab, 2012). Zhang et al. (2009) showed that the increase in the fluorescence intensity of peak (a) in malachite green-bovine serum albumin complex was due to the increased diameter of the macromolecule, which in turn enhanced the scattering effect. Besides peaks (a) and (b), there are two other typical fluorescence peaks (1 and 2), depending on the protein type. Peak (1) mainly reveals the spectral characteristic of the intrinsic fluorophore, such as the polarity of the microenvironment of Trp and Tyr residues when the protein sample is excited at 280 nm. As for peak (2), it provides information about the polypeptide backbone as a result of $n \rightarrow \pi^*$ transition with secondary structural changes (Ma, Yan, Xu, Guo, & Li, 2016). Sui et al. (2018) showed a decrease in the fluorescence intensity of peaks 1 and 2 of soybean protein-anthocyanidin, indicating the strong interaction between protein and anthocyanin, the unfolding of the peptide backbone, and changes in the environment of lysine and Trp residues.

Consequently, the 3D fluorescence spectrum provides insights into protein conformational changes, encompassing both tertiary and polypeptide backbone structures. It can also provide a rough indication of the connection between the secondary structure and the three-dimensional structure in conformational changes.

5.3. X-ray diffraction

XRD and NMR spectroscopy are potent techniques used to study the 3D structure of proteins with near atomic resolution. In XRD, X-rays with energies ranging from 5 to 20 keV are utilized. The schematic diagram is shown in Fig. 3F. When X-rays with wavelengths comparable to interatomic distances interacts with the scattering objects, they are elastically scattered by different atoms, leading to interference and resulting in XRD patterns in specific directions (Kim, Lee, & Kafle, 2013). In an ordered protein crystal, high energy X-rays are scattered by the atoms, forming a 3D interference pattern of diffraction spots. This pattern can be recorded and transformed into the electron density, which is closely related to the scattering molecules in the crystal. It can be interpreted as an atomic structure model, allowing for the analysis of protein structures (Parker, 2003).

XRD has high resolution, and is considered to be one of the best methods for protein structure analysis (Wang et al., 2017). In general, XRD is used to analyze crystallizable proteins; the crystals must be in a well-organized arrangement, as amorphous materials cannot provide sufficient diffraction information (Drenth, 2001). Samples of crystallizable proteins are generally single crystals containing only one molecular species with good diffraction effects. As such, in XRD spectroscopy, contaminants generally have no effect on protein diffraction. XRD spectroscopy can provide the most accurate atomic details (Parker, 2003). In a study by Chen and Subirade (2009), XRD spectroscopy was applied to analyze the structure of pure protein powder and composite protein microspheres prepared by soy protein and zein. The structural changes between protein microspheres and pure protein powder were obtained, as was the structural difference of protein microspheres prepared in different proportions. However, XRD spectroscopy does not excel in obtaining information related to the dynamics of macromolecular motion (Ma et al., 2015).

5.4. Nuclear magnetic resonance

In the case of NMR, the nucleons within an atom nucleus “spin” under the action of an applied magnetic field, and the rate of spin is closely related to the surrounding electronic environment and the nuclear structure. Angular momentum from spin generates a nuclear magnetic moment that can be projected on the applied magnetic field; the projection can take up unique, discrete ‘allowed’ value related to the interaction with the magnetic field (Levitt, 2013). These discrete energy levels in different types of nuclei can be probed by transient bursts of energy in the radio-frequency range. Radio frequency pulse arrays are used to induce transitions between discrete energy levels (“excitations”) in specific nuclei, which will produce detectable magnetization, and transfer the magnetization to other nuclei. Information on the local chemical structure and the global spatial arrangement of the atoms can thus be obtained.

In the NMR experiment, the measurement process includes many important motions of biological macromolecules, and the measurement time can be controlled on the time scale of nanoseconds to seconds. NMR can even directly measure the motions of molecular groups or entire domains (Delgado, Tironi, & Añón, 2011). However, NMR has strict

sample preparation requirements for protein samples. NMR excels in studying macromolecules in solution – as such the purity should be higher than 95% to provide a measurable signal, and in general, molecular weight of protein molecules to be analyzed should be below 50 kDa; above which lower resolution of protein structures is obtained instead (Joachimiak, 2009). Currently, isotope labeling is a common technique that aids in the NMR analysis of proteins. In a study by Kainosho et al. (2006), isotope labelling assisted NMR was used to analyze protein structure through the method of stereo-array isotope labelling, overcoming problems that may occur in NMR spectra, such as low signal ratio.

In general, the structure of the protein sample after crystallization can be analyzed by XRD spectroscopy. NMR spectroscopy is commonly used to analyze protein structures in solutions. XRD and NMR can deliver complementary information, and are among the most powerful and prominent techniques for studying the 3D structure of protein. Fig. 4 is an illustration of the collection of technologies covered herein.

5.5. Laser light scattering and diffusing wave spectroscopy

Laser light scattering (LLS) is a powerful technique for characterizing the structure of protein molecules in solution. The combination of dynamic light scattering (DLS) and static light scattering (SLS) is usually used to characterize the colloidal behavior of proteins such as protein aggregation, conformational change and complex formation. SLS depends on the diffraction pattern of the scattered photons; measurements of the intensity of scattered light are determined as a function of scattering angle in suspensions of macromolecules or colloids. It is commonly used to measure larger droplets that have sizes ranging from 100 nm to 1000 μm (Jin et al., 2016). In the measurement of DLS, the particles in the solution are constantly in Brownian motion, leading to shifts in the diffraction pattern and a correspondingly fluctuating intensity of scattered light reaching the detector. Swiftly moving smaller particles create rapid fluctuations, while larger particles induce slower ones (Zhu & Gao, 2019). The intensity of the scattered light is measured as a function of time, allowing for the size and dynamic information of the particles in the solution to be deduced (Alexander & Dalgleish, 2006). Currently, the suitable particle range characterized using commercial DLS instrument is about 3 nm to 5 μm. It can measure most of the particles with the advantages of accuracy, rapidity, good repeatability, but the measurement requires a highly transparent sample, which often requires a greater factor of dilution (Amin, Rega, & Jankevics, 2012). Xu et al. (2019) used LLS to analyze the structure of protein in soy whey wastewater, as well as to obtain a detailed grasp of the changes in protein structure after adding different ratios of epigallocatechin-3-gallate.

Diffusing wave spectroscopy (DWS), which is similar in principle to the DLS, has developed rapidly in recent years. Unlike LLS, DWS can be used to measure turbid or concentrated samples, making up for the main disadvantage of LLS – in that it can only measure highly transparent

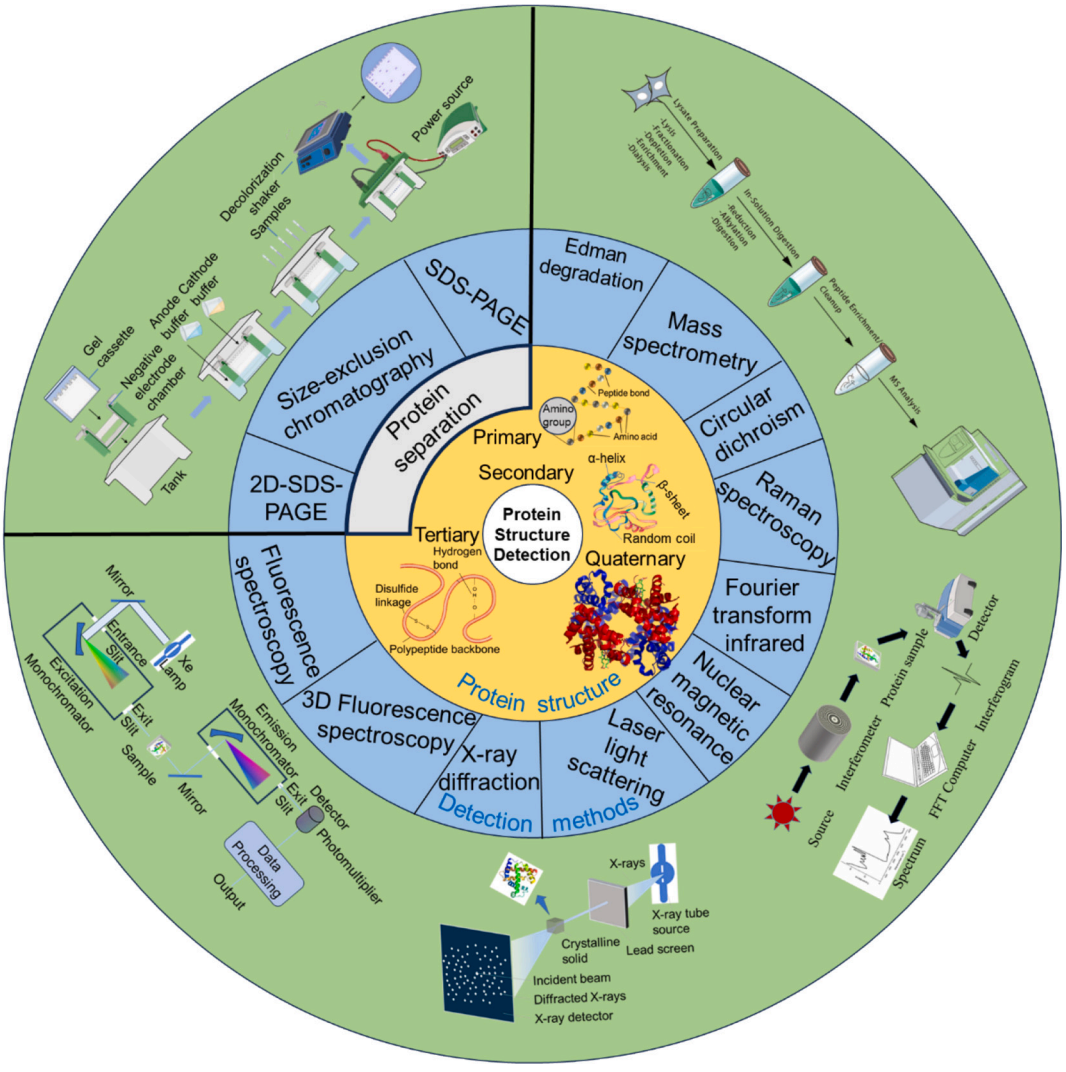


Fig. 4. Techniques for protein separation and detection.

samples (Harden & Viasnoff, 2001). In a study by Vasbinder, Van Mil, Bot, and De Kruif (2001), DWS was used to study the acid-induced aggregation of casein micelles from milk in the presence of whey protein preparations, allowing for the information on the interaction between particles in the aggregated sample to be obtained.

LLS is a label-free method that allows for characterization of protein. However, the development of DWS also makes up for the shortcoming of only testing diluted samples, thereby increasing the practicality of this technology in the future, and enabling a wider range of applications.

6. Assistive analyses

6.1. Chemical analysis

Chemical analysis can be seen as a method for analyzing protein structure using chemical reagents. Chemical reagents react with proteins molecules, and the characteristics of protein samples or chemical reagents may change (such as solubility, particle size, absorbance value, etc.), after which these changes are studied in order to characterize the changes in protein structures. The tertiary structure of the protein is strongly influenced by the secondary structure through the presence of various types of intramolecular interactions (disulfide bond, hydrophobic interaction, hydrogen bond, ionic bond, van der Waals force). This section reviews the use of different chemical reagents to analyze disulfide bonds and non-covalent interactions.

The disulfide bond is a covalent bond that stabilizes protein molecules by reducing the entropy of the denatured state. Disulfide bonds and sulfhydryl groups can be interconverted through redox reactions. For instance, the formation of a disulfide bond results from the oxidation of two cysteine sulfhydryl groups (Depuydt, Messens, & Collet, 2011). Several highly sensitive and accurate methods exist for quantitatively measuring sulfhydryl groups, such as electrochemical and fluorometric assays. However, these methods are time-consuming. Ellman's reagent is widely used to directly determine the content of sulfhydryl groups in proteins due to its simplicity and speed (Li & Zhao, 2006). This reagent utilizes the reaction between 5, 5′-dithiobis-2-nitrobenzoic acid (DTNB) and the sulfhydryl groups, resulting in the formation of 2-nitro-5-mercaptobenzoic acid (TNB). TNB2- exhibits strong absorption at 412 nm, while DTNB has no absorption at this wavelength, allowing the identification of free sulfhydryl groups in proteins through ultraviolet spectroscopy. At present, many studies have characterized the effect of different conditions on the changes of food protein disulfide bonds and sulfhydryl groups using Ellman's reagent. Samples with good light transmittance typically require no excessive treatment. Conversely, certain samples cannot be directly used with Ellman's reagent, for which special sample pre-treatment is required before spectroscopic measurements. For instance, soybean milk protein, with its low sulfhydryl group content and high turbidity, requires separation by acetone precipitation to create a low-turbidity protein solution. The content of sulfhydryl groups and disulfide bonds in this solution is then determined using Ellman's reagent (Ou, Kwok, Wang, & Bao, 2004). Although Ellman's reagent is a common reagent for the determination of protein sulfhydryl groups by spectrophotometry, there is a risk of incomplete measurement due to the presence of cystamine as a "reaction quencher" between the protein sulfhydryl groups and Ellman's reagent. Ellman's reagent is usually replaced by 4,4′-dithiodipyridine (DTDP), which has the advantages of hydrophobicity and requiring small sample volumes. DTDP reacts rapidly with poorly accessible protein sulfhydryl groups, thereby skipping the cysteine catalysis, shortening the reaction time (Hansen, Østergaard, Nørgaard, & Winther, 2007). However, this method is limited by the color of the sample, pH, and measuring instrument, and so on. Nevertheless, it still provides an alternative method of verification for the determination of protein sulfhydryl groups and disulfide bonds.

DTT, SDS and urea are also reagents commonly used to analyze proteins interactions. DTT is a strong reducing agent, which can reduce the disulfide bonds that maintain the protein tertiary and quaternary structure. Urea and SDS are both known to disrupt hydrophobic interaction and hydrogen bonds. SDS is more efficient than urea in disrupting hydrophobic interaction, whereas urea is more efficient in breaking hydrogen bonds (Chiang, Loveday, Hardacre, & Parker, 2019). These reagents, individually or in combination, can provide insights into the types of intramolecular interactions present in proteins. They affect various protein characteristics such as solubility, viscosity, and particle size. In short, these methods find extensive use in measuring chemical bonds in proteins due to their simplicity and lack of instrumental limitations.

7. Conclusions

There are several methods for the measurement of protein structure, each with its own unique characteristics and in most cases complementary. In the detection of protein separation, results obtained from the SDS-PAGE method are directly interpreted, while SEC is milder in terms of sample pre-treatment. Most techniques for determining the secondary structure have broad applicability and are suitable for various types of protein samples, whether in liquid or solid forms. However, it should be noted that secondary structure detection is often employed in conjunction with multiple methods due to the inherent characteristics of the sample and its environment. In the detection of protein tertiary structure, the 3D fluorescence spectrum presents more detailed conformational changes of the protein than 2D fluorescence, not only including changes inside the tertiary structure, but also changes in polypeptide backbone structures. XRD is mainly applied to detect crystallizable proteins, while NMR can provide effective and dynamic information of protein samples in solution. Although these techniques exhibit several advantages for measuring the structures of proteins, there are still some challenges to overcome before protein conformation detection is fully achieved due to the complexity of ingredients in actual foods. We believe that with the further development of technology, the limitations of some of these methods will be surpassed, in turn providing greater convenience, sensitivity, and accuracy for the measurement of protein structure.

CRediT authorship contribution statement

Tian Lan: Conceptualization, Writing – original draft. **Yabo Dong:** Writing – original draft. **Lianzhou Jiang:** Supervision. **Yan Zhang:** Supervision. **Xiaonan Sui:** Conceptualization, Supervision, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors have no conflicts of interest to be declared.

Data availability

Data will be made available on request.

Acknowledgement

We gratefully acknowledge the financial support received from the Distinguished Young Scientists Fund of NSFC (32325041) and the Key Program of NSFC (32230082).

References

Alexander, M., & Dalgleish, D. G. (2006). Dynamic light scattering techniques and their applications in food science. *Food Biophysics, 1*, 2–13.

Al-Mbaideen, A., & Benaissa, M. (2011). Frequency self deconvolution in the quantitative analysis of near infrared spectra. *Analytica Chimica Acta, 705*, 135–147.

Amin, S., Rega, C. A., & Jankevics, H. (2012). Detection of viscoelasticity in aggregating dilute protein solutions through dynamic light scattering-based optical microrheology. *Rheologica Acta, 51*, 329–342.

Armirotti, A., & Damonte, G. (2010). Achievements and perspectives of top-down proteomics. *Proteomics, 10*(20), 3566–3576.

Assad-Bustillos, M., Cázares-Godoy, A. C., de Lavergne, M. D., Schmitt, C., Hartmann, C., & Windhab, E. (2023). Assessment of the interactions between pea and salivary proteins in aqueous dispersions. *Innovative Food Science & Emerging Technologies, 84*, Article 103290.

Bacsik, Z., Mink, J., & Keresztury, G. (2004). FTIR spectroscopy of the atmosphere. I. Principles and methods. *Applied Spectroscopy Reviews, 39*, 295–363.

Bean, S., & Lookhart, G. (2001). Factors influencing the characterization of gluten proteins by size-exclusion chromatography and multiangle laser light scattering (SEC-MALLS). *Cereal Chemistry, 78*, 608–618.

Beaubier, S., Pineda-Vadillo, C., Mesieres, O., Framboisier, X., Galet, O., & Kapel, R. (2023). Improving the in vitro digestibility of rapeseed albumins resistant to gastrointestinal proteolysis while preserving the functional properties using enzymatic hydrolysis. *Food Chemistry, 407*, Article 135132.

Bernarde, C., Lehours, P., Lasserre, J.-P., Castroviejo, M., Bonneu, M., Mégraud, F., & Ménard, A. J. M. (2010). Complexomics study of two helicobacter pylori strains of two pathological origins: Potential targets for vaccine development and new insight in bacteria metabolism. *Molecular & Cellular Proteomics, 9*, 796–2826.

Borgo, B., & Havranek, J. J. (2015). Computer-aided design of a catalyst for Edman degradation utilizing substrate-assisted catalysis. *Protein Science, 24*(4), 571–579.

Bortolotti, A., Wong, Y. H., Korsholm, S. S., Bahring, N. H. B., Bobone, S., Tayyab, S., … Stella, L. (2016). On the purported "backbone fluorescence" in protein three-dimensional fluorescence spectra. *Royal Society of Chemistry, 6*, 12870–112876.

Boye, J. I., Roufik, S., Pesta, N., & Barbana, C. (2010). Angiotensin I-converting enzyme inhibitory properties and SDS-PAGE of red lentil protein hydrolysates. *LWT- Food Science and Technology, 43*, 987–991.

Brunelle, J. L., & Green, R. (2014). One-dimensional SDS-polyacrylamide gel electrophoresis (1D SDS-PAGE). *Methods in Enzymology, 541*, 151–159.

Burgess, R. (2018). A brief practical review of size exclusion chromatography: Rules of thumb, limitations, and troubleshooting. *Protein Expression and Purification, 150*, 81–85.

Chattopadhyay, A., & Raghuraman, H. (2004). Application of fluorescence spectroscopy to membrane protein structure and dynamics. *Current Science, 87*, 175–180.

Chavez, J. D., Wippel, H. H., Tang, X., Keller, A., & Bruce, J. E. (2021). In-cell labeling and mass spectrometry for systems-level structural biology. *Chemical Reviews, 122* (8), 7647–7689.

Chen, G., Fan, M., Liu, Y., Sun, B., Liu, M., Wu, J., Li, N., & Guo, M. (2019). Advances in MS based strategies for probing ligand-target interactions: Focus on soft ionization mass spectrometric techniques. *Frontiers in Chemistry, 7*, 703.

Chen, L., & Subirade, M. (2009). Elaboration and characterization of soy/zein protein microspheres for controlled nutraceutical delivery. *Biomacromolecules, 10*, 3327–3334.

Chiang, J. H., Loveday, S. M., Hardacre, A. K., & Parker, M. E. (2019). Effects of soy protein to wheat gluten ratio on the physicochemical properties of extruded meat analogues. *Food Structure, 19*, Article 100102.

Claeys, E., De Smet, S., Balcaen, A., Raes, K., & Demeyer, D. J. M. S. (2004). Quantification of fresh meat peptides by SDS–PAGE in relation to ageing time and taste intensity. *Meat Science, 67*, 281–288.

Dauer, K., Werner, C., Lindenblatt, D., & Wagner, K. G. (2023). Impact of process stress on protein stability in highly-loaded solid protein/PEG formulations from small-scale melt extrusion. *International Journal of Pharmaceutics: X, 5*, Article 100154.

Delgado, M. C. O., Tironi, V. A., & Añón, M. C. (2011). Antioxidant activity of amaranth protein or their hydrolysates under simulated gastrointestinal digestion. *LWT- Food Science and Technology, 44*, 1752–1760.

Deller, R., Carter, B., Zampetakis, I., Scarpa, F., & Perriman, A. (2018). The effect of surface charge on the thermal stability and ice recrystallization inhibition activity of antifreeze protein III (AFP III). *Biochemical and Biophysical Research Communications, 495*, 1055–1060.

Depuydt, M., Messens, J., & Collet, J.-F. (2011). How proteins form disulfide bonds. *Antioxidants & Redox Signaling, 15*, 49–66.

Deshpande, M., & Sathe, S. K. (2018). Interactions with 8-Anilinonaphthalene-1-sulfonic acid (ANS) and surface hydrophobicity of black gram (Vigna mungo) Phaseolin. *Journal of Food Science, 83*, 1847–1855.

Diamante, L. (2023). High performance liquid chromatography with size exclusion column (HPLC-SEC) method for identifying the major whey proteins of whey protein products. *Eurasian Journal of Biological and Chemical Sciences, 6*(1), 1–6.

Ding, J., Xu, Z., Qi, B., Liu, Z., Yu, L., Yan, Z., … Sui, X. (2020). Thermally treated soya bean oleosomes: The changes in their stability and associated proteins. *International Journal of Food Science and Technology, 55*, 229–238.

Dong, Y., Lan, T., Wang, X., Zhang, Y., Jiang, L., & Sui, X. (2020). Preparation and characterization of soy protein microspheres using amorphous calcium carbonate cores. *Food Hydrocolloids, 107*, Article 105953.

Drenth, J. (2001). *X-ray diffraction: Principles*. John Wiley & Sons.

Dupree, E., Jayathirtha, M., Yorkey, H., Mihasan, M., Petre, B., & Darie, C. (2020). A critical review of bottom-up proteomics: The good, the bad, and the future of this field. *Proteomes, 8*(3), 14.

Edwards, P. J., & Jameson, G. B. (2020). Structure and stability of whey proteins. In H. S. Boland, & A. K. Thompson (Eds.), *Milk proteins* (pp. 201–242). Elsevier.

Eftink, M. R. (1991). Fluorescence techniques for studying protein structure. *Methods of Biochemical Analysis, 35*, 127–205.

Fazzino, T. L., Amber, B. C., Juen, G., & Kevin, D. H. (2023). Ad libitum meal energy intake is positively influenced by energy density, eating rate and hyper-palatable food across four dietary patterns. *Nature Food, 4*, 144–147.

Fekete, S., Beck, A., Veuthey, J.-L., & Guillarme, D. (2014). Theory and practice of size exclusion chromatography for the analysis of protein aggregates. *Journal of Pharmaceutical and Biomedical Analysis, 101*, 161–173.

Floyd, B. M., & Marcotte, E. M. (2022). Protein sequencing, one molecule at a time. *Annual Review of Biophysics, 51*, 181–200.

Folta-Stogniew, E. (2006). Oligomeric states of proteins determined by size-exclusion chromatography coupled with light scattering, absorbance, and refractive index detectors. *New and Emerging Proteomic Techniques*, 97–112.

Freitas, C. D., Souza, D. P., Grangeiro, T. B., Sousa, J. S., Lima, I. V., Pedro Filho, N. S., … Ramos, M. V. (2023). Proteomic analysis of Cryptostegia grandiflora latex, purification, characterization, and biological activity of two osmotin isoforms. *International Journal of Biological Macromolecules, 252*, Article 126529.

Friedman, D. B., Hoving, S., & Westermeier, R. (2009). Isoelectric focusing and two-dimensional gel electrophoresis. *Methods in Enzymology, 463*, 515–540.

Gallagher, S. R. (2012). SDS-polyacrylamide gel electrophoresis (SDS-PAGE). *Current Protocols Essential Laboratory Techniques, 6*(1), 7-3.

Gantumur, M.-A., Hussain, M., Li, J., Hui, M., Bai, X., Sukhbaatar, N., Li, J., Chen, W., Hou, J., & Jiang, Z. (2023). Modification of fermented whey protein concentrates: Impact of sequential ultrasound and TGase cross-linking. *Food Research International, 163*, Article 112158.

Greenfield, N. J. (2006). Using circular dichroism spectra to estimate protein secondary structure. *Nature Protocols, 1*, 2876.

Greer, S. M., Sidoli, S., Coradin, M., Schack Jespersen, M., Schwämmle, V., Jensen, O. N., … Brodbelt, J. S. (2018). Extensive characterization of heavily modified histone tails by 193 nm ultraviolet photodissociation mass spectrometry via a middle–down strategy. *Analytical Chemistry, 90*(17), 10425–10433.

Guo, W., Zhao, Y., Yao, Y., Wu, N., Xu, M., Du, H., & Tu, Y. (2019). Relationship between protein structure changes and in vitro digestion of preserved egg white during pickling. *International Journal of Biological Macromolecules, 138*, 116–124.

Hansen, R. E., Østergaard, H., Nørgaard, P., & Winther, J. R. (2007). Quantification of protein thiols and dithiols in the picomolar range using sodium borohydride and 4, 4′-dithiodipyridine. *Analytical Biochemistry, 363*, 77–82.

Harden, J., & Viasnoff, V. (2001). Recent advances in DWS-based micro-rheology. *Current Opinion in Colloid & Interface Science, 6*, 438–445.

Hawe, A., Sutter, M., & Jiskoot, W. (2008). Extrinsic fluorescent dyes as tools for protein characterization. *Pharmaceutical Research, 25*, 1487–1499.

Hellmann, N., & Schneider, D. (2019). Hands on: using tryptophan fluorescence spectroscopy to study protein structure. *Protein Supersecondary Structures: Methods and Protocols*, 379–401.

Hong, P., Koza, S., & Bouvier, E. S. (2012). A review size-exclusion chromatography for the analysis of protein biotherapeutics and their aggregates. *Journal of Liquid Chromatography & Related Technologies, 35*, 2923–2950.

Huang, Z., Jiang, Y., Li, H., Li, Q., Gao, Z., Zhang, Y., Zhang, Q., & Fu, L. (2023). Effect of glycation derived from α-dicarbonyl compounds on the in vitro digestibility of ovalbumin: Tracing of advanced glycation end-products and immuno-active peptides. *Food Research International, 169*, Article 112842.

Hunt, D. F., Yates, J., 3rd, Shabanowitz, J., Winston, S., & Hauer, C. R. (1986). Protein sequencing by tandem mass spectrometry. *Proceedings of the National Academy of Sciences, 83*(17), 6233–6237.

Hussain, R., Longo, E., & Siligardi, G. (2018). UV-denaturation assay to assess protein photostability and ligand-binding interactions using the high photon flux of diamond B23 beamline for SRCD. *Molecules, 23*, 1906.

Issaq, H. J., & Veenstra, T. D. (2008). Two-dimensional polyacrylamide gel electrophoresis (2D-PAGE): Advances and perspectives. *Biotechniques, 44*, 697–700.

Jach, M. E., Serefko, A., Ziaja, M., & Kieliszek, M. (2022). Yeast protein as an easily accessible food source. *Metabolites, 12*(1), 63.

Jia, J., Gao, X., Hao, M., & Tang, L. (2017). Comparison of binding interaction between β-lactoglobulin and three common polyphenols using multi-spectroscopy and modeling methods. *Food Chemistry, 228*, 143–151.

Jiang, L., Liu, Y., Li, L., Qi, B., Ju, M., Xu, Y., Zhang, Y., & Sui, X. (2019). Covalent conjugates of anthocyanins to soy protein: Unravelling their structure features and in vitro gastrointestinal digestion fate. *Food Research International, 120*, 603–609.

Jin, W., Xu, W., Liang, H., Li, Y., Liu, S., & Li, B. (2016). Nanoemulsions for food: Properties, production, characterization, and applications. *Emulsions, 3*, 1–36.

Joachimiak, A. (2009). High-throughput crystallography for structural genomics. *Current Opinion in Structural Biology, 19*, 573–584.

Jonsson, A. (2001). Mass spectrometry for protein and peptide characterisation. *Cellular and Molecular Life Sciences, 58*, 868–884.

Kainosho, M., Torizawa, T., Iwashita, Y., Terauchi, T., Ono, A. M., & Güntert, P. (2006). Optimal isotope labelling for NMR protein structure determinations. *Nature, 440*, 52–57.

Kanakis, C., Hasni, I., Bourassa, P., Tarantilis, P., Polissiou, M., & Tajmir-Riahi, H.-A. (2011). Milk β-lactoglobulin complexes with tea polyphenols. *Food Chemistry, 127* (3), 1046–1055.

Kar, A., Snigdha, G., Jeyamkondan, S., & Kaustav, M. (2023). Effect of radiofrequency processing on the structural and bio-functional properties of egg white proteins. *Food Chemistry, 404*, Article 134533.

Karoui, R., & Blecker, C. (2011). Fluorescence spectroscopy measurement for quality assessment of food systems—A review. *Food and Bioprocess Technology, 4*, 364–386.

Kebarle, P., & Peschke, M. (2000). On the mechanisms by which the charged droplets produced by electrospray lead to gas phase ions. *Analytica Chimica Acta, 406*(1), 11–35.

Khursheed, N., Ashfaq, A., Yousuf, O., Anjum, Z., & Younis, K. (2024). An overview of extraction of animal-based protein via various extraction techniques and its functional properties. *Green Chemistry in Food Analysis*, 45–61.

Kieliszek, M., Kot, A. M., Bzducha-Wróbel, A., Błażejak, S., Gientka, I., & Kurcz, A. (2017). Biotechnological use of Candida yeasts in the food industry: *A review*. *Fungal Biology Reviews, 31*(4), 185–198.
Kim, S. H., Lee, C. M., & Kafle, K. (2013). Characterization of crystalline cellulose in biomass: Basic principles, applications, and limitations of XRD, NMR, IR, Raman, and SFG. *Korean Journal of Chemical Engineering, 30*, 2127–2141.
Kong, J., & Yu, S. (2007). Fourier transform infrared spectroscopic analysis of protein secondary structures. *Acta Biochimica et Biophysica Sinica, 39*, 549–559.
König, S., Obermann, W. M., & Eble, J. A. (2022). The current state-of-the-art identification of unknown proteins using mass spectrometry exemplified on de novo sequencing of a venom protease from Bothrops moojeni. *Molecules, 27*(15), 4976.
Kuan, Y.-H., Bhat, R., Patras, A., & Karim, A. (2013). Radiation processing of food proteins–A review on the recent developments. *Trends in Food Science & Technology, 30*(2), 105–120.
Kumar, M., Tomar, M., Potkule, J., Verma, R., Punia, S., Mahapatra, A., … Berwal, M. K. (2021). Advances in the plant protein extraction: Mechanism and recommendations. *Food Hydrocolloids, 115*, Article 106595.
Kurcz, A., Błażejak, S., Kot, A. M., Bzducha-Wróbel, A., & Kieliszek, M. (2018). Application of industrial wastes for the production of microbial single-cell protein by fodder yeast Candida utilis. *Waste and Biomass Valorization, 9*, 57–64.
Ladokhin, A. S. (2000). Fluorescence spectroscopy in peptide and protein analysis. *Encyclopedia of Analytical Chemistry*, 5762–5779.
Lan, T., Dong, Y., Zheng, M., Jiang, L., Zhang, Y., & Sui, X. (2020). Complexation between soy peptides and epigallocatechin-3-gallate (EGCG): Formation mechanism and morphological characterization. *LWT- Food Science and Technology, 134*, Article 109990.
Lasserre, J.-P., Nicaud, J.-M., Pagot, Y., Joubert-Caron, R., Caron, M., & Hardouin, J. (2010). First complexomic study of alkane-binding protein complexes in the yeast Yarrowia lipolytica. *Talanta, 80*(4), 1576–1585.
Lehrman, S. R. (2017). Protein structure. *Fundamentals of Protein Biotechnology*, 9–38.
Leurs, U., Mistarz, U. H., & Rand, K. D. (2015). Getting to the core of protein pharmaceuticals–comprehensive structure analysis by mass spectrometry. *European Journal of Pharmaceutics and Biopharmaceutics, 93*, 95–109.
Levitt, M. H. (2013). *Spin dynamics: Basics of nuclear magnetic resonance*. John Wiley & Sons.
Li, D., Zhu, L., Wu, Q., Chen, Y., Wu, G., & Zhang, H. (2023). Identification of binding sites for Tartary buckwheat protein-phenols covalent complex and alterations in protein structure and antioxidant properties. *International Journal of Biological Macromolecules, 233*, Article 123436.
Li, J., Dai, Z., Chen, Z., Hao, Y., Wang, S., & Ma, X. (2023). Improved gelling and emulsifying properties of myofibrillar protein from frozen shrimp (Litopenaeus vannamei) by high-intensity ultrasound. *Food Hydrocolloids, 135*, Article 108188.
Li, W., Huang, J., Zheng, L., Liu, W., Fan, L., Sun, B., Su, G., Xu, J., & Zhao, M. (2023). A fast stop-flow two-dimensional liquid chromatography tandem mass spectrometry and its application in food-derived protein hydrolysates. *Food Chemistry, 406*, Article 135000.
Li, X., Chen, L., Hua, Y., Chen, Y., Kong, X., & Zhang, C. (2020). Effect of preheating-induced denaturation during protein production on the structure and gelling properties of soybean proteins. *Food Hydrocolloids, 105*, Article 105846.
Li, Y., & Zhao, M. (2006). Simple methods for rapid determination of sulfite in food products. *Food Control, 17*, 975–980.
Lilley, K. S., Razzaq, A., & Dupree, P. (2002). Two-dimensional gel electrophoresis: Recent advances in sample preparation, detection and quantitation. *Current Opinion in Chemical Biology, 6*, 46–50.
Ma, P., Xue, Y., Coquelle, N., Haller, J. D., Yuwen, T., Ayala, I., … Skrynnikov, N. R. (2015). Observing the overall rocking motion of a protein in a crystal. *Nature Communications, 6*, 8361.
Ma, X., Yan, J., Xu, K., Guo, L., & Li, H. (2016). Binding mechanism of trans-N-caffeoyltyramine and human serum albumin: Investigation by multi-spectroscopy and docking simulation. *Bioorganic Chemistry, 66*, 102–110.
Maiti, N. C., Apetri, M. M., Zagorski, M. G., Carey, P. R., & Anderson, V. E. (2004). Raman spectroscopic characterization of secondary structure in natively unfolded proteins: α-synuclein. *Journal of the American Chemical Society, 126*, 2399–2408.
Maky, M. A., & Zendo, T. (2023). Identification of a novel bioactive peptide derived from frozen chicken breast hydrolysate and the utilization of hydrolysates as biopreservatives. *Biology, 12*(9), 1218.
Maus, A., Mignon, R., & Basile, F. (2018). Enhanced protein identification using graphite-modified MALDI plates for offline LC-MALDI-MS/MS bottom-up proteomics. *Analytical Biochemistry, 545*, 31–37.
McDonald, T. P., & Sugrue, R. J. (2007). The use of two-dimensional SDS-PAGE to analyze the glycan heterogeneity of the respiratory syncytial virus fusion protein. In R. J. Sugrue (Ed.), *Glycovirology protocols* (pp. 97–108).
Meyer, H. E., Fröhlich, T., Nordhoff, E., & Kuhlmann, K. (2022). Massenspektrometrie. *Bioanalytik*, 359–414.
Micsonai, A., Bulyáki, É., & Kardos, J. (2021). BeStSel: From secondary structure analysis to protein fold prediction by circular dichroism spectroscopy. *Structural Genomics: General Applications*, 175–189.
Möller, A., Albert, V., & Atze, J. (2022). Influence of the fractionation method on the protein composition and functional properties. *Innovative Food Science & Emerging Technologies, 81*, Article 103144.
Moon, J. H., Yoon, S., Bae, Y. J., & Kim, M. S. (2015). Formation of gas-phase peptide ions and their dissociation in MALDI: Insights from kinetic and ion yield studies. *Mass Spectrometry Reviews, 34*(2), 94–115.
Movasaghi, Z., Rehman, S., & Ur Rehman, D. I. (2008). Fourier transform infrared (FTIR) spectroscopy of biological tissues. *Applied Spectroscopy Reviews, 43*, 134–179.
Naseri, A., Marinho, G. S., Holdt, S. L., Bartela, J. M., & Jacobsen, C. (2020). Enzyme-assisted extraction and characterization of protein from red seaweed Palmaria palmata. *Algal Research, 47*, Article 101849.
Nemecek, D., Stepanek, J., & Thomas, G. J., Jr. (2013). Raman spectroscopy of proteins and nucleoproteins. *Current Protocols in Protein Science, 71*(1), 17–18.
Niall, H. D. (1973a). [36] Automated edman degradation: The protein sequenator. In , *Vol. 27*. *Methods in enzymology* (pp. 942–1010). Elsevier.
Niall, H. D. (1973b). [36] Automated edman degradation: The protein sequenator. In , *27*. *Methods in enzymology* (pp. 942–1010).
Olías, R., Rayner, T., Clemente, A., & Domoney, C. (2023). Combination of three null mutations affecting seed protein accumulation in pea (Pisum sativum L.) impacts positively on digestibility. *Food Research International, 169*, Article 112825.
Ou, S., Kwok, K., Wang, Y., & Bao, H. (2004). An improved method to determine SH and–S–S–group content in soymilk protein. *Food Chemistry, 88*, 317–320.
Panchal, G., Hati, S., & Sakure, A. (2020). Characterization and production of novel antioxidative peptides derived from fermented goat milk by L. fermentum. *LWT, 119*, 108887.
Parker, M. (2003). Protein structure from X-ray diffraction. *Journal of Biological Physics, 29*, 341–362.
Pelton, J. T., & McLean, L. R. (2000). Spectroscopic methods for analysis of protein secondary structure. *Analytical Biochemistry, 277*, 167–176.
Phillips-Jones, M. K., & Harding, S. E. (2019). Tapping into synchrotron and benchtop circular dichroism spectroscopy for expanding studies of complex polysaccharides and their interactions in anoxic archaeological wood. *Heritage, 2*(1), 121–134.
Rabilloud, T., Chevallet, M., Luche, S., & Lelong, C. (2010). Two-dimensional gel electrophoresis in proteomics: Past, present and future. *Journal of Proteomics, 73*(11), 2064–2077.
Raghuraman, H., Chatterjee, S., & Das, A. (2019). Site-directed fluorescence approaches for dynamic structural biology of membrane peptides and proteins. *Frontiers in Molecular Biosciences, 6*, 96.
Raikos, V., Hansen, R., Campbell, L., & Euston, S. R. (2006). Separation and identification of hen egg protein isoforms using SDS–PAGE and 2D gel electrophoresis with MALDI-TOF mass spectrometry. *Food Chemistry, 99*(4), 702–710.
Révész, Á., Hevér, H., Steckel, A., Schlosser, G., Szabó, D., Vékey, K., & Drahos, L. (2023). Collision energies: Optimization strategies for bottom-up proteomics. *Mass Spectrometry Reviews, 42*(4), 1261–1299.
Roy, S., & Kumar, V. (2014). A practical approach on SDS PAGE for separation of protein. *International Journal of Science and Research, 3*, 955–960.
Sajib, M., Forghani, B., Vate, N. K., & Abdollahi, M. (2023). Combined effects of isolation temperature and pH on functionality and beany flavor of pea protein isolates for meat analogue applications. *Food Chemistry, 412*, Article 135585.
Samtiya, M., Rotimi, E., & Tejpal, D. (2020). Plant food anti-nutritional factors and their reduction strategies: An overview. *Food Production, Processing and Nutrition, 2*, 1–14.
Sangiorgio, S., Vidović, N., Boschin, G., Aiello, G., Arcidiaco, P., Arnoldi, A., … Scarabattoli, L. (2022). Preparation, characterization and *in vitro* stability of a novel ACE-inhibitory peptide from soybean protein. *Foods, 11*(17), 2667.
Sazonova, S., Grube, M., Shvirksts, K., Galoburda, R., & Gramatina, I. (2019). FTIR spectroscopy studies of high pressure-induced changes in pork macromolecular structure. *Journal of Molecular Structure, 1186*(15), 377–383.
Schägger, H. (2006). Tricine–sds-page. *Nature Protocols, 1*, 16–22.
Schwaighofer, A., Alcaráz, M. R., Araman, C., Goicoechea, H., & Lendl, B. (2016). External cavity-quantum cascade laser infrared spectroscopy for secondary structure analysis of proteins at low concentrations. *Scientific Reports, 6*(1), 33556.
Shashilov, V. A., Sikirzhytski, V., Popova, L. A., & Lednev, I. K. (2010). Quantitative methods for structural characterization of proteins based on deep UV resonance Raman spectroscopy. *Methods, 52*, 23–37.
Shi, L., Yang, X., Gong, T., Hu, C., Shen, Y., & Meng, Y. (2023). Ultrasonic treatment improves physical and oxidative stabilities of walnut protein isolate-based emulsion by changing protein structure. *LWT- Food Science and Technology, 173*, Article 114269.
Shrestha, S., Hag, L., Haritos, V., & Dhital, S. (2023). Comparative study on molecular and higher-order structures of legume seed protein isolates: Lentil, mungbean and yellow pea. *Food Chemistry, 411*, Article 135464.
Some, D., Amartely, H., Tsadok, A., & Lebendiker, M. (2019). Characterization of proteins by size-exclusion chromatography coupled to multi-angle light scattering (SEC-MALS). *Journal of Visualized Experiments, 148*, Article e59615.
Striegel, A. M. (2023). Size-exclusion chromatography. In *Liquid chromatography* (pp. 509–537).
Su, Y.-Z., Hung, M.-W., Huang, K.-C., & Lin, L.-W. (2011). Resolution of skin auto-fluorescence for non-invasive application. In *IEEE international instrumentation and measurement technology conference* (pp. 1–4).
Suhaib Al Huq, M., Raja, K., & Oviya, I. R. (2024). Prokaryotic cell membrane-based protein technologies. *World Academy of Sciences Journal, 6*(2), 1–8.
Sui, X., Sun, H., Qi, B., Zhang, M., Li, Y., & Jiang, L. (2018). Functional and conformational changes to soy proteins accompanying anthocyanins: Focus on covalent and non-covalent interactions. *Food Chemistry, 245*, 871–878.
Takahashi, N., Isobe, T., Desiderio, D. M., & Nibbering, N. M. (2007). *Proteomic biology using LC/MS: Large scale analysis of cellular dynamics and function*. John Wiley & Sons.
Tang, C.-H., & Ma, C.-Y. (2009). Effect of high pressure treatment on aggregation and structural properties of soy protein isolate. *LWT- Food Science and Technology, 42*, 606–611.
Van Schaick, G., Haselberg, R., Somsen, G. W., Wuhrer, M., & Domínguez-Vega, E. (2022). Studying protein structure and function by native separation–mass spectrometry. *Nature Reviews Chemistry, 6*(3), 215–231.

Vankeirsbilck, T., Vercauteren, A., Baeyens, W., Van der Weken, G., Verpoort, F., Vergote, G., & Remon, J. P. (2002). Applications of Raman spectroscopy in pharmaceutical analysis. *TrAC Trends in Analytical Chemistry, 21*, 869–877.

Vasbinder, A. J., Van Mil, P. J., Bot, A., & De Kruif, K. G. (2001). Acid-induced gelation of heat-treated milk studied by diffusing wave spectroscopy. *Colloids and Surfaces B: Biointerfaces, 21*, 245–250.

Walker, J. M. (2002). SDS polyacrylamide gel electrophoresis of proteins. In *The protein protocols handbook* (pp. 61–67).

Wallace, B. A. (2009). Protein characterisation by synchrotron radiation circular dichroism spectroscopy. *Quarterly reviews of biophysics, 42*(4), 317–370.

Wang, K., Sun, D.-W., Pu, H., & Wei, Q. (2017). Principles and applications of spectroscopic techniques for evaluating food protein conformational changes: A review. *Trends in Food Science & Technology, 67*, 207–219.

Wei, Y., Thyparambil, A. A., Latour, R. A., & Proteomics. (2014). Protein helical structure determination using CD spectroscopy for solutions with strong background absorbance from 190 to 230 nm. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics, 1844*, 2331–2337.

Wen, J., Zhao, J., Jiang, L., & Sui, X. (2024). Oil-water interfacial behavior of zein-soy protein composite nanoparticles in high internal phase Pickering emulsion. *Food Hydrocolloids, 149*, Article 109659.

Xu, Z., Hao, N., Li, L., Zhang, Y., Yu, L., Jiang, L., & Sui, X. (2019). Valorization of soy whey wastewater: How Epigallocatechin-3-gallate regulates protein precipitation. *ACS Sustainable Chemistry & Engineering, 7*, 15504–15513.

Yim, S.-K., Ahn, T., Kim, J.-S., & Yun, C.-H. J. (2002). Polyacrylamide gel electrophoresis without a stacking gel: Application for separation of peptides. *Analytical Biochemistry, 305*, 277–279.

Yin, H., Jin, H., & Zhang, H. (2018). Study on isolation and Raman spectroscopy of glycinin in soybean protein. *Grain & Oil Science and Technology, 1*(2), 72–76.

Ying, Y., Zhao, L., Kong, L., Kong, X., Hua, Y., & Chen, Y. J. (2015). Solubilization of proteins in extracted oil bodies by SDS: A simple and efficient protein sample preparation method for Tricine–SDS–PAGE. *Food Chemistry, 181*, 179–185.

Yoneda, J. S., Miles, A. J., Araujo, A. P. U., & Wallace, B. A. (2017). Differential dehydration effects on globular proteins and intrinsically disordered proteins during film formation. *Protein Science, 26*, 718–726.

Yu, D., Zhang, X., Zou, W., Tang, H., Yang, F., Wang, L., & Elfalleh, W. (2021). Raman spectroscopy analysis of the effect of electrolysis treatment on the structure of soy protein isolate. *Journal of Food Measurement and Characterization, 15*, 1294–1300.

Yuan, D., Zhou, F., Niu, Z., Shen, P., & Zhao, M. (2023). Formation of mucus-permeable nanoparticles from soy protein isolate by partial enzymatic hydrolysis coupled with thermal and pH-shifting treatment. *Food Chemistry, 398*, Article 133851.

Zaroog, M. S., & Tayyab, S. J. (2012). Formation of molten globule-like state during acid denaturation of aspergillus Niger glucoamylase. *Process Biochemistry, 47*, 775–784.

Zenaidee, M. A., Wei, B., Lantz, C., Wu, H. T., Lambeth, T. R., Diedrich, J. K., … Loo, J. A. (2021). Internal fragments generated from different top-down mass spectrometry fragmentation methods extend protein sequence coverage. *Journal of the American Society for Mass Spectrometry, 32*(7), 1752–1758.

Zhang, F., Shen, R., Xue, J., Yang, X., & Lin, D. (2023). Characterization of bacterial cellulose nanofibers/soy protein isolate complex particles for Pickering emulsion gels: The effect of protein structure changes induced by pH. *International Journal of Biological Macromolecules, 226*, 254–266.

Zhang, H., & Ge, Y. (2011). Comprehensive analysis of protein modifications by top-down mass spectrometry. *Circulation. Cardiovascular Genetics, 4*(6), 711.

Zhang, J., Jia, Y., Wu, W., Zhang, Y., Chen, P., Li, X., Wei, X., Li, C., & Li, K. (2023). Influence of hemin on structure and emulsifying properties of soybean protein isolate. *Food Chemistry, 421*, Article 136183.

Zhang, Y., Chen, S., Qi, B., Sui, X., & Jiang, L. (2018). Complexation of thermally-denatured soybean protein isolate with anthocyanins and its effect on the protein structure and in vitro digestibility. *Food Research International, 106*, 619–625.

Zhang, Y.-Z., Zhou, B., Zhang, X.-P., Huang, P., Li, C.-H., & Liu, Y. (2009). Interaction of malachite green with bovine serum albumin: Determination of the binding mechanism and binding site by spectroscopic methods. *Journal of Hazardous Materials, 163*, 1345–1352.

Zhu, J., Li, S., Yang, L., Zhao, Z., Xia, J., Zhu, Y., & Li, C. (2023). Effect of multiple freeze–thaw cycles on water migration, protein conformation and quality attributes of beef longissimus dorsi muscle by real-time low field nuclear magnetic resonance and Raman spectroscopy. *Food Research International, 166*, Article 112644.

Zhu, S., Lin, S., Ramaswamy, H., Yu, Y., & Zhang, Q. (2017). Enhancement of functional properties of rice bran proteins by high pressure treatment and their correlation with surface hydrophobicity. *Food and Bioprocess Technology, 10*(2), 317–327.

Zhu, X., & Gao, T. (2019). Spectrometry. In *Nano-inspired biosensors for protein assay with clinical applications* (pp. 237–264). Elsevier.