Proceedings of Anticancer Research, 2026, Volume 10, Issue 3
http://ojs.bbwpublisher.com/index.php/PAR
Online ISSN: 2208-3553
Print ISSN: 2208-3545

Abnormal Protein Structure Characterization Based on Proteomics and Its Application in Disease Diagnosis

Zekun Duan*

University of Queensland, Sir Fred Schonell Drive, St Lucia, Brisbane, QLD 4072, Australia

Author to whom correspondence should be addressed.

Abstract: As the key agents of biological processes, proteins' functionality is closely tied to the integrity of their spatial structure. Structural abnormalities can lead to protein dysfunction, which, in consequence, may induce a variety of diseases. Proteomics focuses on the entire proteome. The combination of high-throughput detection techniques and bioinformatics analysis has opened up new possibilities for characterising abnormal protein structures. It also provides a crucial molecular basis for the early diagnosis, classification and prognosis assessment of diseases. This review summarises the core proteomics techniques used to characterise abnormal protein structures. It focuses on the distinguishing features of such structures in fields such as neurodegenerative diseases, cancer and metabolic disorders, as well as their diagnostic applications, and provides a reference for proteomics-based research into abnormal proteins and the precise diagnosis of diseases.

Keywords: Proteomics; Abnormal proteins; Structural characterization; Disease diagnosis; Molecular biomarkers

---

1. Introduction

Proteins are the end products of gene expression; their function is determined not only by their amino acid sequence, but also by their secondary, tertiary and quaternary structures. In physiological conditions, proteins rely on molecular chaperones to fold correctly. Multiple regulatory mechanisms, including synthesis, modification, transport and degradation, maintain the structural stability of proteins [1]. This results in a system of protein stability within the cellular environment, ensuring the normal functioning of all vital processes. When this stability is disrupted, proteins may become misfolded, aggregate abnormally or be mislocalized, resulting in the loss of their original physiological functions and even causing toxic damage, which can subsequently induce pathological changes. Such conditions are also known as protein conformation disorders and are common diseases that threaten human health [2]. Previous research has mainly focused on individual proteins, making it difficult to capture dynamic changes in intracellular proteins or to meet the requirements of combined clinical testing. Proteomics focuses on all proteins in the body; by utilising high-throughput technologies such as mass spectrometry and two-dimensional gel electrophoresis, combined with bioinformatics methods, it can systematically analyse protein expression, structure and interactions [3]. With the continuous advancement of relevant technologies, the accuracy of detecting abnormal proteins has improved significantly, and proteomics is now widely used in the early screening, precise diagnosis and assessment of treatment efficacy for clinical diseases.

2. Methods of characterising abnormal protein structures based on proteomics

2.1. Laboratory analysis techniques

The core of the characterization of abnormal protein structures is the provision of direct evidence through experimental methods that directly detect the spatial conformation, modification status and aggregation properties of proteins. Recently, analytical techniques based on mass spectrometry (MS), nuclear magnetic resonance (NMR) and cryo-electron microscopy (Cryo-EM) have become mainstream, significantly improving the resolution and efficiency of structural determination.

2.1.1. Mass spectrometry

Mass spectrometry (MS) has become a core technique in proteomics due to its high sensitivity, high resolution and high throughput. MS enables the sequencing of abnormal proteins, the detection of post-translational modifications, the analysis of conformational changes and the identification of aggregation states [4]. The main types of MS currently in use include the following:

Cross-Linking Mass Spectrometry (XL-MS) involves adding a cross-linking agent to a protein sample to form covalent cross-links between amino acid residues. The cross-linked peptides are then detected by mass spectrometry and analysed using bioinformatics to determine the spatial distances between amino acid residues. This enables the construction of three-dimensional structural models, allowing for the detection of conformational changes and abnormal interactions within proteins [5]. For example, in Alzheimer's disease, this technique can detect cross-linking sites in β-amyloid (Aβ) and characterise the conformational changes that occur during its aggregation process.

Hydrogen-deuterium exchange mass spectrometry (HDX-MS) utilises the exchange properties of amide hydrogen with deuterium water; by detecting changes in exchange rates via mass spectrometry, it analyzes the secondary structure and conformational dynamics of proteins—regions that are misfolded or unfolded exhibit significantly higher exchange rates than normal regions, enabling rapid differentiation of conformational differences [6]. This technique requires only a small sample size and offers rapid detection. It is widely used for the characterization of abnormal proteins associated with tumours, such as the detection of conformational abnormalities in the p53 protein.

Matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry (MALDI-TOF-MS) involves mixing proteins with a matrix; laser excitation generates ions, and molecular weight is determined based on time-of-flight, enabling rapid detection of protein aggregation states. It is suitable for screening for amyloid aggregation, such as the detection of α-synuclein aggregation in Parkinson's disease [7]. MS/MS allows for further analysis of amino acid sequences, identifies sites of structural abnormalities, and provides a basis for precise identification.

2.1.2. Nuclear magnetic resonance (NMR) technology

Nuclear magnetic resonance (NMR) technology utilises the resonance of atomic nuclei in a magnetic field. By measuring the chemical shifts of nuclei such as hydrogen, carbon and nitrogen, NMR characterises the secondary and tertiary structures, and conformation dynamics, of small-molecule proteins (molecular weight < 50 kDa) [8]. This technology enables real-time monitoring of protein folding processes and captures intermediate states of misfolding. For example, in prion diseases, it can detect the transition of prion protein (PrP) from its normal conformation (PrP^C) to its pathogenic conformation (PrP^Sc), thus characterising the underlying molecular mechanisms.

Developments in High-Field NMR technology have broadened applications. This technique can resolve the structures of protein complexes, but it is subject to limitations such as high sample consumption, lengthy analysis times, and difficulties in resolving macromolecules. For this reason, it is most commonly employed in conjunction with MS for enhanced accuracy and comprehensiveness in characterization.

2.1.3. Cryo-electron microscopy

Cryo-EM quickly freezes protein samples to liquid nitrogen temperature to preserve their native structure. Two-dimensional images are then captured using an electron microscope, and these are combined with three-dimensional reconstruction techniques to construct a three-dimensional structural model. This method offers high resolution, no requirement for crystallisation, and minimal sample usage, and is especially suitable for the characterization of macromolecular proteins and aggregates [9,10]. This is a key method for characterising the structure of abnormal protein aggregates. In Alzheimer's disease, it enables the analysis of the structure of Aβ aggregates, clarifying their aggregation patterns and mechanisms of toxicity. Furthermore, in Parkinson's disease, it allows the structure of α-synuclein aggregate fibres to be resolved, providing direct evidence for research into the disease's pathogenic mechanisms. Last but not least, two-dimensional gel electrophoresis (2-DE), as a conventional technique, separates proteins of different molecular weights and isoelectric points by combining isoelectric focusing with SDS-PAGE electrophoresis. It enables preliminary screening by exploiting the differences in migration rates between abnormal and normal proteins, and when combined with mass spectrometry for identification, provides initial clues for characterization [11].

2.2. Bioinformatics prediction methods

Developments in bioinformatics have improved tools for protein structure prediction based on amino acid sequences. Presently, these tools can quickly predict secondary and tertiary structures, support the characterization of abnormal proteins, and compensate for limitations in laboratory techniques. Commonly used methods are as follows:

Protein structure prediction tools such as AlphaFold and RosettaFold analyse amino acid sequences using deep learning algorithms. These tools can predict three-dimensional structures with an accuracy of over 90%, enabling the rapid identification of the conformational features and abnormal sites of abnormal proteins [12]. For example, AlphaFold can be used to predict the three-dimensional structure of the mutant p53 protein, thus elucidating the mechanism of its loss of function; it can also predict the structure of Aβ, helping to analyse the structural changes that occur during the assembly process.

Misfolding prediction tools, like TANGO and AGGRESCAN, are designed to analyse amino acid sequences. They predict the likelihood of misfolding and identify aggregation domains, enabling the rapid screening of high-risk proteins. This technology is widely used in the prediction of proteins associated with neurodegenerative diseases. Examples include the prediction of abnormalities in the huntingtin protein in Huntington's disease and the SOD1 protein in amyotrophic lateral sclerosis [13].

NetPhos and NetGlyc are post-translational modification prediction tools that can forecast phosphorylation, glycosylation and other modification sites, and support the analysis of structural abnormalities resulting from modification defects [14]. Taking examples, when predicting phosphorylation sites on the Tau protein, the tool can be used to identify conformational abnormalities caused by hyperphosphorylation; it can also be employed to predict phosphorylation sites on the EGFR protein, to analyse the association between its mutations and the development of tumours.

Bioinformatics-based predictions have the advantages of being fast, efficient and cost-effective. It enables large-scale protein screening, but the results must be confirmed through experimental results to ensure accuracy.

3. Applications of abnormal protein structures in disease diagnosis

3.1. Neurodegenerative diseases

Core pathological features of neurodegenerative diseases include the misfolding and aggregation of specific proteins, resulting in insoluble amyloid deposits. This subsequently leads to neuronal damage and death [15]. Traditional diagnosis of this type of disease has been heavily reliant on clinical symptoms and imaging tests. This approach lacks specific molecular biomarkers, which makes early diagnosis particularly challenging. Proteomics technology, however, has opened up new paths for its early diagnosis.

Alzheimer disease (AD)'s core pathological features are the abnormal aggregation of Aβ and the hyperphosphorylation of tau protein [16]. Aβ is produced by the abnormal cleavage of the APP protein; in its normal state, it exists as a soluble monomer, but when abnormal, it misfolds and aggregates into amyloid fibrils, which deposit in the cerebral cortex and hippocampus. Overphosphorylation of the tau protein leads to the formation of neurofibrillary tangles, which disrupt the neuronal cytoskeleton. In the study, the misfolded intermediate state of Aβ was identified as a key factor in the early onset of AD and may serve as a potential diagnostic biomarker. Overexpression of Ser396 and Ser404 is associated with disease progression and can be used to assess disease status [17].

Parkinson's disease (PD) is characterised by the misfolding and aggregation of α-synuclein (α-syn) into Lewy bodies. α-syn aggregation occurs in three stages: monomers, oligomers and fibrils, with oligomers being the primary toxic form responsible for damage to dopaminergic neurons [18]. The levels of α-syn oligomers in the cerebrospinal fluid of Parkinson's disease patients are significantly higher than in healthy individuals and may serve as a biomarker for early diagnosis. The N-terminal and C-terminal regions are key sites for misfolding, and mutations in the relevant amino acids accelerate aggregation, making them useful for genetic risk assessment.

Amyotrophic lateral sclerosis (ALS) is characterised by the misfolding and aggregation of proteins such as TDP-43 and SOD1, which form inclusion bodies [19]. Mutations in the SOD1 gene lead to misfolding of the protein, resulting in oxidative stress and neuronal toxicity; abnormal modification of TDP-43 disrupts RNA metabolism. Research indicates that abnormal TDP-43 interactions can serve as a diagnostic biomarker, while the intermediate state content of misfolded SOD1 can be used to monitor treatment efficacy.

3.2. Tumour

Structural abnormalities in proteins are closely linked to the development and progression of cancers. The activation of proto-oncogenes and the inactivation of tumour suppressor genes are often caused by changes in protein conformation, abnormal post-translational modifications, or abnormal aggregation [20]. Through proteomic characterization of abnormal protein structures, tumour-specific molecular biomarkers can be identified. This provides a basis for early screening, precise classification and prognostic assessment, thereby advancing precision medicine.

The p53 protein is a major tumour suppressor, and its dysfunction is associated with more than 50% of human solid tumours. Normal wild-type p53 regulates the cell cycle and induces apoptosis [21]; mutations in the p53 gene can lead to misfolding of the protein, loss of its tumour-suppressing function, and even the acquisition of tumour-promoting activity. NMR and mass spectrometry analyses have shown that the core DNA-binding domain of p53 is a mutation hotspot. Following mutation, its ability to bind to DNA is reduced and its conformational stability is diminished, leading to the formation of insoluble aggregates. HDX-MS can rapidly distinguish between wild-type and mutant p53 based on differences in hydrogen-deuterium exchange rates. Results indicate that mutation sites such as R175H and R273H are associated with rapid disease progression and poor prognosis and can therefore be used for prognostic assessment.

The epidermal growth factor receptor (EGFR) is a critical therapeutic target in cancer treatment, and structural abnormalities in this receptor are associated with non-small cell lung cancer, breast cancer and other malignancies [22]. While the native EGFR exists as a monomer, ligand binding induces dimerisation, which activates downstream signalling pathways. Mutations such as the deletion of exon 19 and the L858R mutation in exon 21 can alter its conformation, leading to sustained activation in the absence of ligands and driving tumour proliferation. Cryo-EM analysis reveals that EGFR mutant dimers are more stable, with conformational changes in the kinase domain being key to sustained activation [23]; XL-MS detection has found that phosphorylation levels in the tyrosine kinase domain are significantly elevated, which can serve as a biomarker for diagnosis and the assessment of targeted therapy efficacy.

Researchers from Fudan University analysed the plasma proteomes of 53,026 individuals and found that over 650 proteins are associated with at least 50 diseases. GDF15 emerged as a key biomarker for the diagnosis and prognosis of various tumours, whilst EDA2R, NTproBNP and others also demonstrated good diagnostic performance [24]. Furthermore, mutations in splicing factors within tumours lead to abnormal splicing of genes such as CD44 and VEGFR; the resulting abnormal protein isoforms can serve as diagnostic markers for tumour invasion, metastasis and drug resistance, and mass spectrometry can be used to detect them with high precision.

3.3. Metabolic disorders

Metabolic disorders (such as diabetes, obesity and fatty liver disease) are closely linked to metabolic disturbances caused by abnormal protein structures; proteomic characterization can elucidate the underlying molecular mechanisms, thus providing a basis for diagnosis and intervention.

The core pathological features of type 2 diabetes mellitus (T2DM) are insulin resistance and insufficient secretion, which are associated with structural abnormalities in the insulin protein and the insulin receptor (IR). Mass spectrometry and NMR analyses have shown that misfolding of the insulin protein reduces its activity, whereas abnormal phosphorylation of the tyrosine kinase domain of the IR leads to conformational changes, preventing the activation of downstream signalling pathways and resulting in insulin resistance [25]. HDX-MS analysis has revealed that the cleavage of disulfide bonds between the A and B chains of insulin leads to misfolding, forming inactive aggregates; the levels of these aggregates correlate with disease progression and may serve as diagnostic biomarkers. Furthermore, abnormal phosphorylation at Tyr1162 and Tyr1163 on the IR is associated with insulin resistance and can be utilised for disease assessment.

Abnormal protein structures are closely associated with obesity and adipocyte differentiation; peroxisome proliferator-activated receptor gamma (PPARγ) is a key regulatory factor, whose conformational changes influence ligand-binding capacity and regulate adipocyte differentiation and lipid metabolism. Cryo-EM analysis reveals that an abnormal conformation of the PPARγ ligand-binding domain leads to excessive activation, promoting excessive adipocyte differentiation; XL-MS detection has found that abnormal interactions with co-activators can serve as diagnostic markers and intervention targets for obesity.

4. Conclusion

Proteomics technologies provide a novel technical foundation for the systematic characterization of abnormal protein structures. Through the combined application of experimental detection and bioinformatics-based prediction, it is possible to comprehensively analyse the conformational changes, modification status and aggregation characteristics of abnormal proteins. Abnormal protein structures exhibit specific characteristics in a wide range of diseases, including neurodegenerative disorders, cancer and metabolic diseases. Characterising these structures not only aids in elucidating the pathogenic mechanisms of these diseases but also provides important molecular biomarkers for early diagnosis, disease classification and prognosis assessment. With continuous technological innovation and deepening research, their application in disease diagnosis will become increasingly widespread. This is expected to drive major breakthroughs in precision diagnosis and treatment, thereby providing stronger safeguards for human health.

Disclosure statement

The authors declare no conflict of interest.

References

[1] Chang C, Yan Q, Jiang Z, 2022, Research Progress in the Preparation of Bioactive Peptides by Enzymatic Hydrolysis of Legume Proteins. Food Research and Development, 43(11): 10–17.

[2] Zheng L, Feng Y, Rong Z, et al., 2026, Application of Nanomaterials in the Treatment of Ischemic Stroke. Chinese Journal of Cell Biology, 48(03): 823–842.

[3] Zhou T, Wang W, Zhang Z, et al., 2026, Application and Advances of Transcriptomics and Proteomics Technologies in the Study of Intervertebral Disc Degeneration. Chinese Journal of Tissue Engineering Research, 30(23): 5925–5933.

[4] Wang X, Zhang F, Wu X, et al., 2025, Advances in Nontargeted Analysis of Per- and Polyfluoroalkyl Substances Based on High-Resolution Mass Spectrometry. Journal of Analysis and Testing, 9(04): 457–481.

[5] Wang J, Gong Z, Dong X, et al., 2022, Fast Cross-Linking by DOPA2 Promotes the Capturing of a Stereospecific Protein Complex over Nonspecific Encounter Complexes. Biophysics Reports, 8(Z1): 239–252.

[6] Zhang Y, Wang P, Li T, et al., 2025, Application and Advances of HDX-NMR and HDX-MS in the Study of Protein Structural Dynamics. Journal of Magnetic Resonance, 42(04): 445–456.

[7] Zhou H, Zhao H, Qi Y, 2026, Application of MALDI-TOF Mass Spectrometry in the Study of Endogenous Peptides in Dairy Products. Food Science, 1–16.

[8] Chen S, Zhang L, Li S, et al., 2024, Detecting Biomarkers by Dynamic Nuclear Polarization Enhanced Magnetic Resonance. National Science Review, 11(09): 71–87.

[9] Liu N, Wang H, 2019, Research Status and Future Prospects of Single-Particle Cryo-Electron Microscopy Technology. China Basic Science, 21(05): 1–7.

[10] Yang Y, Gong Y, Gu T, et al., 2023, Advances in Cryo-Electron Microscopy Technology and Its Applications in Environmental Research. Acta Chimica Sinica, 81(08): 990–1001.

[11] Zhao Y, Yang L, Yu Y, et al., 2024, Proteomics Technology and Its Application in the Study of Fluorosis. Heilongjiang Animal Science and Veterinary Medicine, (10): 27–31 + 39.

[12] Che Y, Wu R, Zheng J, 2026, Application of AI Protein Structure Prediction in the Biomedical Field. Chemistry Education (Chinese and English), 47(04): 1–8.

[13] Shi J, Teng J, Li J, et al., 2025, Research Progress on the Mechanism of Action of the Gut-Brain Axis in Amyotrophic Lateral Sclerosis and the Intervention of Traditional Chinese Medicine. Journal of Neurology and Neurorehabilitation, 21(05): 376–385.

[14] Xie X, Yang X, Shen H, et al., 2026, Artificial Intelligence Aids Mass Spectrometry Analysis of Protein Post-Translational Modifications and Their Clinical Applications. Journal of Mass Spectrometry, 1–21.

[15] Fan X, Xie P, Hou X, 2026, Research Progress on the Role of Ataxin-3-Mediated Protein Homeostasis Imbalance in Neurodegenerative Diseases. Journal of Nanjing Medical University (Natural Science Edition), 46(04): 512–519.

[16] Zheng Q, Wang X, 2025, Alzheimer’s Disease: Insights into Pathology, Molecular Mechanisms, and Therapy. Protein & Cell, 16(02): 83–120.

[17] Liu B, Li J, Wang M, et al., 2023, The Effects of Flavonoids on Reducing Aβ Deposition and Tau Hyperphosphorylation in Alzheimer’s Disease. Chemistry of Life, 43(04): 570–578.

[18] Zhang S, Peng Y, Zhang Z, et al., 2026, Research Progress on the α-Synuclein E46K Mutation Exacerbating the Pathological Progression of Parkinson’s Disease. Chinese Pharmacological Bulletin, (03): 411–415.

[19] Jing W, Saereetaveekun N, Voratunyakit N, et al., 2024, Neurodegenerative Diseases, Misfolded and Abnormally Aggregated Proteins, and Sleep Disorders. Journal of Neurology and Neurorehabilitation, 20(02): 58–66.

[20] Qin X, Yu Q, Li X, et al., 2025, Activation and Chemical Targeting Intervention of the Neddylation Pathway of Proteins in Lung Cancer. Science China Chemistry, 55(04): 914–936.

[21] Hou M, Mao J, 2024, Research on the Role of MYBL2 and P53 in the Occurrence and Development of Endometrial Cancer. Inner Mongolia Medical Journal, 56(03): 312–315 + 318.

[22] Jiao J, Li S, Liu L, 2025, Mechanism of the Hippo-YAP Pathway in Promoting the Development of Advanced Non-Small Cell Lung Cancer and Its Clinical Translation Research Progress. Chinese Journal of Clinical Oncology and Rehabilitation, 32(08): 517–526.

[23] Ni R, Hu H, Jiang T, 2026, Research Progress on EGFR Mutations and Targeted Therapy in Glioma. Journal of Neurological Diseases and Mental Health, 26(03): 153–159.

[24] Deng TY, You J, He Y, et al., 2024, Atlas of the Plasma Proteome in Health and Disease in 53,026 Adults. Cell, 188(1): 253–271.e7.

[25] Liu K, Li R, Rui F, et al., 2026, Exploring the Pathogenesis and Treatment of Insulin Resistance Based on the Correlation Between "Spleen Qi Dispersing Essence" and Glucose Metabolism Reprogramming. Journal of Traditional Chinese Medicine, 41(04): 695–701.