Review

Nanoproteomics: a new sprout from emerging links between nanotechnology and proteomics

Lee Jia[1], Yusheng Lu[1], Jingwei Shao[1], Xing-Jie Liang[2], and Yan Xu[3]

[1] Cancer Metastasis Alert and Prevention Institute, College of Chemistry and Chemical Engineering, Fuzhou University, Fuzhou 350002, China
[2] Chinese Academy of Sciences Key Laboratory for Biomedical Effects of Nanomaterials and Nanosafety, National Centre for Nanoscience and Technology of China, Beijing 100190, China
[3] University of California Los Angeles, Olympic Analytical Laboratory, Los Angeles, CA 90025, USA

The complexity of proteomics challenges current methods to provide all peptide mass fingerprints in an *ensemble* measurement of various proteins at differing concentrations. To detect those low-abundance proteins, nanotechnology provides a technical platform to improve biocompatibility, specificity, reproducibility, and robustness of the current proteomic methods. Here, we comprehensively analyze the weaknesses of traditional proteomic methods and evaluate the importance of nanomaterials in significantly improving the quality of proteomic methods by manipulating individual proteins. We also illustrate how the large surface-to-volume ratio of nanomaterials can facilitate mass transfer, enhance the efficiency of separation and high-throughput capability, and reduce assay time and sample consumption. The marriage of the two subjects and the resulting new nanoproteomics will revolutionize proteomics research.

The need for nanotechnology in proteomics

With completion of the human genome project in 2003, proteomics has become the point of interest for a deeper insight into the cellular processes that have not been depicted by genomics. The concept of proteomics involves a comprehensive study on the structures, localizations, post-translational modifications, functions, and interactions of all proteins expressed by an organism at a certain time and under certain conditions. Characterization of proteins will in turn lead to identifying the functional genome. Current estimates are that the proteome of eukaryotic cells contains thousands of proteins with up to 20 000 expressed at any time [1]. Such complexity poses a serious challenge to current proteomic techniques, including both the top-down 2D slab gel electrophoresis introduced 35 years ago [2,3] and the relatively new bottom-up approach using liquid chromatography/tandem mass spectrometry (LC/MS/MS) for peptide sequencing and mass fingerprinting (Box 1) [4,5]. In the 1990s, proteomics was not as strong as genomics. For example, proteomics missed tools like the polymerase chain reaction (PCR) in genomics, which can amplify a single nucleic acid molecule in a complex biological sample up to measurable levels that can be picked up by modern detectors [6]. Consequently, proteomics suffers from limit of detection (LOD): it is still impossible to detect all protein molecules existing in biological samples. Another problem in proteomics is the large dynamic concentration range of different proteins that co-exist in a single biosample – the so-called dynamic concentration barrier. Typical blood samples can contain more than 10 000 different protein species, with concentrations varying over nine orders of magnitude. Proteins at concentrations of 10^{-12}–10^{-15} M usually co-exist with those at concentrations of 10^{-3}–10^{-4} M, which makes detection of low-abundance proteins extremely difficult. Development of improved analytical techniques that can recognize a single protein molecule in the presence of others of high abundance is a top priority for proteomic methodology.

To face these proteomic challenges, protein microarray technologies have undergone rapid developments over the last decade, with cell-free expression systems now being employed for microarray generation [7]. Moreover, application of nanotechnologies to proteomics over the last few years has established a novel technical platform termed 'nanoproteomics' to study the dynamic concentration range of various proteins in complex biological samples. Advantages of nanotechnologies over conventional proteomic analysis include real-time multiplexed assays, low sample and reagent consumption, high sensitivity and specificity, short runtime, and miniaturized robust analytical systems composed of nanoscale separation media and channels. Unique nanomaterials such as carbon nanotubes (CNTs), nanowires, quantum dots (QDs), and gold nanoparticles (GNPs) are increasingly being used for proteomic applications. In this article, we briefly review conventional proteomic techniques and their weaknesses. We evaluate the benefits and risks of the merger between proteomics and nanotechnology and further analyze the scientific novelty, strength/weakness, and approaches, and discuss the perspectives of the technical marriage.

Corresponding authors: Jia, L. (pharmlink@gmail.com); Xu, Y. (yanxu@mednet.ucla.edu)
Keywords: nanoproteomics; nanomaterials; carbon nanotubes; gold nanoparticles; nanostructured surface; nanofluidics.

Box 1. Shotgun sequencing and peptide mass fingerprinting

Shotgun sequencing is an approach used for analyzing a complex protein sample by LC/MS/MS, where the sample is digested with a suitable protease (commonly trypsin), and the resulting peptides are separated by HPLC and then characterized by tandem MS. Proteins are identified by matching the MS fragmentation patterns with predicted information from genomics or proteomics databases. Although many studies have used 1D reverse phase HPLC before MS measurement, such an approach is inherently limited by the number of peptides that can successfully be loaded and resolved on the column and detected by the mass spectrometer.

Peptide mass fingerprinting is a method for protein identification in which the unknown protein of interest is first cleaved into smaller peptides, whose absolute masses can be accurately measured with a mass spectrometer such as MALDI-TOF or electrospray ionization time of flight (ESI-TOF). These masses are then compared to either a database containing known protein sequences or even the genome by using computer programs to theoretically cut the protein into peptides, and calculate the absolute masses of the peptides from each protein. Then, the masses of the peptides of the unknown protein are compared to the theoretical peptide masses of each protein encoded in the genome. The results are statistically analyzed to find the best match. Therefore, the protein sequence has to be present in the database. In practice, each protein in a mixture has to be separated and isolated by 2D gels before cleavage.

Traditional proteomic strategies

Analysis of protein complexes based on a proteomics approach provides insight into specific translational signaling pathways and cross-interactions. Proteomics research takes these forms: (i) global proteomics to identify and catalogue all possible proteins within an organism [8]; (ii) targeted proteomics to identify and elucidate mutations and changes of individual proteins in complex mixtures of disease foci in the course of disease development [9,10]; and (iii) pharmacoproteomics to reveal interactions between proteins and drugs, and identify site-specific modifications (phosphorylation, glycosylation, and S-nitrosylation [11,12]) on the proteins induced by the drugs [13].

The first and perhaps the most crucial step in any successful proteomics experiment is sample preparation. Protein solubilization, protein separation, protease digestion, peptide separation, and peptide selection, all involve steps and protocols that vary greatly among laboratories. Moreover, different commercially available tandem mass spectrometers have different mass accuracies and different rates of peptide selection for fragmentation [14]. Proper sample preparation enhances the reproducibility of results. There is no one unique way to prepare samples and hence each protocol developed differs depending on the type of sample and the type of experiment to be conducted. Table 1 summarizes typical proteomic analysis techniques that are used to identify and quantify functional proteins present in the proteome.

Application of nanomaterials to proteomic analysis

Rapid detection of low-abundance proteomic samples requires the development of ultrasensitive, robust, and high-throughput technology. The above-mentioned techniques, however, are basically ineffective for direct identification of membrane proteins, insensitive for detection of low-abundance proteins, and impossible for assessing biological changes in protein regulation or expression [14]. In order to overcome these technical limitations associated with sensitivity, dynamic range, detection time, and multiplexing, proteomics has begun using nanotechnology resources resulting in a novel analytical platform known as nanoproteomics.

The employment of nanomaterials smaller than 100 nm in at least one dimension with their unique size-related physical and chemical properties into analytical technologies was first suggested almost 30 years ago [15]. However, their uses in proteomic analysis still lags behind the rapid development of other nanotechnologies, despite the numerous progresses made already [16,17]. New separation media feature higher efficiency and better selectivity in proteomic analysis than the traditional methods [18].

Table 1. Biotechnical advantages and disadvantages of current proteomic approaches

Technical categories		Advantages	Disadvantages	Refs
Non-mass spectrometry-based	2D gel electrophoresis	Can detect 1000s of proteins at a single run based on their size and charge	Narrow dynamic range; difficult for automation	[3]
	Multi-dimensional protein identification technology (MudPIT)	Strong cation exchange and reverse phase HPLC combined enhance sensitivity and dynamic range	Insensitive for low-abundance proteins	[62]
	Protein arrays	Surface is arrayed with specific molecules to the surface of the chip; easy for automation; good dynamic range	Require cloning of 100s of proteins	[7]
	Two-hybrid systems	Can detect potential protein interactions; easy for automation; measure DNA	False negatives and positives	[63]
	Isotope-coded affinity tagging	Broader dynamic range; easily automated	Measure relative abundance	[27]
Mass spectrometry-based	MALDI-TOF	Accurate mass measurement; fast and low cost	Extensive sample process prior to analysis	[13]
	Atmospheric pressure-MALDI	Fast and low cost, allowing for high-throughput analysis; good sensitivity	Extensive sample process prior to analysis	
	LC/MS/MS	Rely on fragmentation; broader dynamic range; reproducible; generate rich y, b, and neutral loss ions	Extensive sample process prior to analysis	[64]
	2D gel/MS	Simple sample process	Inefficient extraction of proteins from excised gel	[13]
	NanoMate	Faster and automated; favorable ionization; good sensitivity and reliability	High cost	