Proteomics a Key Tool for a Better Understanding of Endometriosis: a Mini- Review

Bi bi Fatemeh Nobakht Motlagh Ghochani¹, Kambiz Gilany²,*

¹ Faculty of Paramedical Sciences, Shahid Beheshti University of medical sciences, Tehran, Iran.
² Reproductive Biotechnology Research Center, Avicenna Research Institute, ACECR, Tehran, Iran.

*Corresponding author: e-mail address: k.gilany@avicenna.ac.ir. (K.Gilany)

ABSTRACT

Endometriosis is a painful reproductive disease afflicting about up to 20% of women. It is one of the most frequent benign gynaecological diseases, however, little is known about the pathological of endometriosis. Over the past decade, high-throughput proteomics technologies have evolved considerably and have become increasingly more commonly applied to the investigation of female reproductive disease, including endometriosis. In this mini-review the authors look at the application of proteomics technologies in order to find biomarker associated with endometriosis.

Keyword: Endometriosis; Proteomics; Diagnosis

INTRODUCTION

Endometriosis is a gynecological pathogenesis manifested by the growth of endometrial glands and stroma outside the uterine cavity. Endometriosis is mostly observed in the fertility years and it is only rarely observed in adolescent and postmenopausal women. The common cause of endometriosis is pain symptoms (dysmenorrhea, dyspareunia, chronic pelvic pain) and infertility [1]. Women with endometriosis is estimated to be 4-19%, however, endometriosis is more frequently observed in women with pain symptoms and infertility [2, 3].

Pathogenesis of endometriosis

The etiology of endometriosis remains a mystery. Several theories about pathogenesis of endometriosis have been suggested. The most accepted one is Sampson's theory. This theory believes that the disease originated from retrograde menstruation of viable endometrial tissue through the fallopian tubes into the peritoneal cavity, where it implants on the peritoneal surface of pelvic organs. However, retrograde menstruation occurs almost in all women and it does not explain presence of endometriosis in the remote areas such as the lungs, skin, and breast. Another theory suggest that endometriosis is due to the occurrence of celomic metaplasia in the pelvic peritoneum [4]. Most recently, adult stem cell has been suggested that is the cause of endometriosis. It is suggested that adult stem cells reside in the basalis. Since the endometrium comprises glands, surface epithelium and supportive stroma, it is hypothesized that both epithelial and stromal stem/progenitor cells are responsible for the regenerative capacity of endometrium [5].

Proteome definition

The proteome has been defined as the protein complement of the genome. However, the definition of proteome has changed since its first defined by Wilkins et al. in 1995 [6]. Today, the term of proteome has developed to be: "The proteome of an individual is defined by the sum and the time dynamics of all protein species occurring during the life-time of this individual ". This definition of the proteome includes the protein expression of the individual protein, the isoforms of a protein and post-translational modifications of a protein [7].

Proteomics technologies

The first proteome analysis goes back to 1975, although the proteome was not defined at that moment. Protein separation and comparison by
two-dimensional gel electrophoresis (2-DE) is the classical method for quantitative analysis of the proteome [8]. Klose [9] and O’Farrell [10] first described the technique in 1975. The major advantage of 2-DE lies in its potential to simultaneously resolve thousands of proteins, at the same time revealing their MW, pI, and reflecting changes in protein expression and isoforms [11]. Apart from these advantages, 2-DE has also some drawbacks. One of its limitations includes the difficulty identification of low (<15 kDa) and high (>150 kDa) molecular weight of proteins and separation is generally limited to proteins that are neither too acidic/basic, nor too hydrophobic. A further limitation of traditional 2-DE is time-consuming and labor-intensive. In addition, the lack of reproducibility between gels leads to significant system variability making it difficult to distinguish between system variations and induced biological change, which means that real differences between protein abundance attributed.

To overcome some of the shortcomings of the above approaches non-gel-based quantitative proteomics methods have been developed significantly in recent years. The development of gel-free proteomic techniques such as Multidimensional Protein Identification (MudPIT) has provided powerful tools for studying large scale protein expression and characterization in complex biological systems [12, 13]. This approach was first described by Link et al. [14]. The technique consists of a 2-D LC peptide separation. A protein mixture is digested using proteolytic enzyme, such as trypsin. The resulting peptides are then separated by multidimensional LC coupled to mass spectrometry. First, the peptide mixtures are separated by strong cation exchange (SCX, first dimension). It is then followed by reverse-phase (RP, second dimension) separation coupled to electrospray mass spectrometer (ESI-MS/MS) [14]. MudPIT not only overcomes the shortcomings of 2-DE but also provides the following advantages: elimination of the time-consuming step for protein separation; high sensitivity and requirement of small sample size; and versatile mechanisms for peptide separation. A major disadvantage of MudPIT is an inability to readily provide information on protein isoforms or posttranslational modifications [15].

Another gel-free proteomic technique based on the principle of diagonal electrophoresis/chromatography is combined fractional diagonal chromatography (COFRADIC). Briefly, whole proteome digests are first separated by RP-HPLC into distinct fractions. Each primary fraction or a combination thereof is then treated with an enzyme or a chemical compound modifying the structure of a selected class of peptides. This modification reaction is chosen such that peptides holding such modified structures are differently retained by chromatographic columns. Thus, when such modified primary fractions are separated a second time under identical chromatographic conditions as during the primary separation, they separate from non-modified peptides and are isolated for LC-MS/MS analysis [16]. COFRADIC method has been used to isolate N-terminal, methionyl peptides etc. from whole proteome digests [17, 18]. The main disadvantage of the COFRADIC technology compared with the 2-DE approach is the lack of information regarding the pI of the proteins identified, which would miss relevant information about potential PTMs. However, both methods complement each other, and thus COFRADIC allowed the identification of several membrane-spanning proteins not previously identified by 2-DE, demonstrating that gel-free methods are able to overcome some of the limitations of the 2-DE approach [19].

One way to determine quantification of proteins is difference gel electrophoresis (DIGE). Due to the limitations of 2-DE that mentioned above Unlu et al. [20] first described a method, 2-DE DIGE, that enabled more than one sample to be separated in a single 2-DE. The development of 2-DE DIGE gives more accurate and reliable quantification information of protein abundance because the samples to be compared are run together on the same gel, eliminating potential gel-to-gel variation. However, spots on a given 2-DE often contain more than one protein, making quantification ambiguous since it is not immediately apparent which protein in the spot has changed. In addition, any 2-DE approach is subject to the restrictions imposed by the gel method [21].

A number of stable isotope labeling approaches have been developed for “shot-gun” quantitative proteomic analysis. These are divided into two
categories: in vivo (metabolic labeling) and in vitro (enzymatic or chemical labeling) [22]. The stable isotope labeling methods have provided valuable flexibility while using quantitative proteomic techniques to study protein changes in complex samples. However, most labeling-based quantification approaches have potential limitations. These include increased time and complexity of sample preparation, requirement for higher sample concentration, high cost of the reagents, incomplete labeling, and the requirement for specific quantification software [23].

Metabolic or in vivo labeling involves the incorporation of stable isotopes during protein biosynthesis [24]. Initially described for total labeling of bacteria using 15N-enriched cell culture medium it has gained wider popularity in the form of the stable isotope labeling by amino acids in cell culture (SILAC) approach introduced by Mann and co-workers in 2002 [25, 26]. The SILAC technique relies on the incorporation of isotopically labeled amino acids into proteins formed by the growing organism. Isotopically labeled amino acids are usually added to the growth medium or the labeled amino acids can be generated by the organism through the addition of isotopically labeled salts to the growth medium [25-27].

The advantages of SILAC are that it has higher fidelity than ICAT (incorporating nearly 100% efficiency) and does not require multiple chemical processing and purification steps, thus ensuring that the samples to be compared have been subjected to similar conditions throughout the experiment. However, this approach requires viable active cell lines to allow for the incorporation of the respective heavy/ light amino acids into the protein samples and may not always be available for all experimental samples [28].

Post-labeling of proteins and peptides is performed by chemical or enzymatic derivatization in vitro. The first in vitro labeling the ICAT (Isotope Coded Affinity Tags) approach developed by Reudi Aebersold and et al. [29]. Because cysteine is a rare amino acid, ICAT and related methods significantly reduce the complexity of the peptide mixture which can be advantageous when highly complex samples are analyzed. In addition, as it is a solution-phase labeling, those proteins not amenable to 2-D PAGE, such as very acidic/basic, too large/small proteins can now be analyzed using an LC-MS workflow. However, ICAT is obviously not suitable for quantifying the significant number of proteins that do not contain any (or a few) cysteine residues and is of limited use for analysis of post-translational modifications and splice isoforms [30]. One limitation of ICAT is that there are only two labels available. This could result in multiple experiments if more than two treatments need to be compared, and would increase the cost accordingly. The need for comparisons of larger numbers of treatments led to the development of the 4- or 8-plex iTRAQ, which can compare up to four or eight samples in a single analysis, respectively. The iTRAQ technique was first described by Ross et al. in 2004 [31]. The iTRAQ label is an isobaric tagging compound consisting of a reporter group (variable mass of 4-plex: 114–117 Da or 8-plex: 113–121 Da), a balance group and an amino-reactive group that introduces a highly basic group at lysine side chains and at peptide N-terminal (Fig. 1). During the initial MS scan, labeled peptides appear as a single peak due to the isobaric masses. The isobaric nature of iTRAQ-labeled peptides allows the signal from all peptides to be summed in both MS and MS/MS modes thus enhancing the sensitivity of detection. During MS/MS, the label releases the reporter group as a singly charged ion of masses at m/z 114–117 (4-plex) or m/z 113–121 (8-plex) [22]. The advantages of iTRAQ are to identify and quantify low-abundance proteins in complex samples – coupled with the ability to multiplex up to eight samples in parallel suggests that iTRAQ and similar mass balanced labels holds the most promise for quantitative biomarker discovery [32].

A primary computerized search was performed in PubMed of publications for quantitative proteomics. Figure 2 show the number of publications per year using various types of quantitative proteomics up to 2010. Although it was impossible to select ‘key words’ to include every relevant publication, it was clear that all of the methods described above are still in use. Moreover, with the possible exception of ICAT, the use of most of these methods is increasing each year.
In fact, the numbers of publications using DIGE, iTRAQ methods are all increasing at approximately equal rate. Surface-enhanced laser desorption/ionization (SELDI-MS) uses chemically defined or antibody-coated protein biochip arrays for rapid protein detection. An advantage of SELDI-TOF-MS is its relatively high tolerance for salts and other impurities. The sample requirement is low (1–10 μg total protein per spot) and sample volume can be freely chosen from 0.5 μl up to around 400 μl. However, it is incompatible with TOF MS/MS and does not allow reliable protein identification [33-35].
Application of proteomics technologies for the study of endometriosis

Over the past decade, high-throughput proteomics technologies have been developed rapidly and have become increasingly more commonly applied to the investigation of endometriosis. Proteomics technology enables comparison of hundreds or thousands of proteins to identify disease-specific biomarkers. The discovery of novel candidate biomarkers is one of crucial problems for the early diagnosis of endometriosis. In the area of biomarker discovery, proteomic techniques are proving particularly powerful tools to identify protein in blood or tissues that may be markers of disease. New classes of biomarkers derived from mass spectroscopy analysis of the low molecular weight proteome have shown improved abilities in the early detection of disease and hence in patient risk stratification and outcome. This review mainly presents current advances in the problems and prospects of candidate biomarker for the early diagnosis of endometriosis, discovered by technologies of proteomics. To our knowledge, only several studies have been published in which 2DE-PAGE was applied to the analysis of serum, endometrium, peritoneal fluid and endometrial fluid from women with endometriosis, with some promising preliminary findings of differential protein expression between diseased and nondiseased subjects (is explained below).

Zhang et al. [36] compared the protein expression maps of eutopic endometria and sera of women with or without endometriosis using two-dimensional gel electrophoresis. After the comparative proteomic study, they have identified 13 and 11 differentially expressed proteins in sera and eutopic endometrium between the two study groups, respectively. These proteins were characterized by searching a computerized database using molecular weight and isoelectric points, but some of sera proteins remain elusive. Some of the matched proteins with different expression may be cytoskeletons, and others may be the regulatory proteins of cell cycle, signal transduction, or immunological function. Those proteins include the G antigen family B1 protein, actin-related protein 6, actinlike-7- anhydrase I, Dentin matrix acidic phosphoprotein I, CD166 antigen, cyclin A1, and 14-3-3 protein sigma et al. Another group used 2D-PAGE for the Comparison of protein expression during both the secretory and proliferative phase eutopic endometrium from women with and without endometriosis [37]. They identified dysregulated proteins in women with endometriosis which included: molecular chaperones including heat shock protein 90 and annexin A2, proteins involved in cellular redox state, such as peroxiredoxin 2, proteins involved in protein and DNA formation/breakdown, including ribonucleoside-diphosphate reductase, prohibitin and prolyl 4-hydroxylase, and secreted proteins, such as apolipoprotein A1.

Have and colleagues [38] studied eutopic endometrium in women with endometriosis, using 2DE-PAGE. A total of 820 protein spots were matched on 2-D gels, with 119 proteins regulated differentially between the two study groups. Of the 50 highest fold change proteins 21 proteins were found only in the endometriosis affected sample group. The authors observed that several molecules were up- and down-regulated in several areas including apoptosis, immune reaction, glycolytic pathway, cell structure, and transcription factors.

In one study applied proteomic techniques to the analysis of peritoneal fluid (PF) proteome in fertile and infertile women with endometriosis, the investigators used 2-DE to analyze the PF of patients. A total of 114 protein spots were presented in PF of fertile and infertile women with endometriosis. Nine protein spots had significantly higher expression in PF of infertile women when compared with fertile controls with endometriosis. These proteins were identified as 1 isoform of α-1-antitrypsin, 2 isoforms of serotransferrin, 1 isoform of complement C3, 1 isoform of serum amyloid P-component and 1 isoform of clusterin. Three protein spots remain unidentified. Abnormalities in the immune response have been hypothesized to be involved in the pathophysiology of endometriosis associated infertility [39].

Ametzazurra et al. were described the use of endometrial fluid aspirate from the uterine cavity of women as a biological sample for the discovery of biomarkers associated with endometriosis. Samples were collected during the post-ovulatory
secretory phase of the menstrual cycle. They have used 2-DE to examine endometrial fluid. Endometrial fluid exhibits a complex and rich proteome composition with more than 800 protein spots detected. The changes in protein expression observed in this comparison have predominantly identified proteins involved in cell signalling, cell death and cell movement. Among the differentially expressed proteins Ametzazurra et al. found a high representation of cytoskeletal proteins, such as moesin, beta-actin, tubulin beta chain, F-actin capping protein subunit beta, WD repeat protein 1, heat-shock protein beta-1 and septin-11. Also, Levels of 14-3-3 protein sigma and gamma (signal transduction) were significantly higher in the endometrial fluid aspirate of women suffering from endometriosis [40]. Other studies utilized SELDI-TOF-MS technology as a proteomic tool in discovering proteins that are differentially expressed in women with and without endometriosis [41-49]. But, as mentioned above SELDI does not allow reliable protein identification. So this technique is not suitable for diagnosis of endometriosis.

Future perspective
Clinically engineered mass spectroscopy systems are essential for the further development and validation of multiplexed biomarkers that have shown tremendous promise for the early detection of disease. Future endometriosis research in the area of noninvasive diagnosis needs to be performed using endometriotic tissue, endometrium and peritoneal fluid in the quest to identify new biomarkers for this complex disease and increase understanding of the pathophysiology. Newer technologies such as gel free proteomic techniques like for example iTRAQ and iTRAQ combined MudPIT and proteomic pattern profiling should be explored in the area of development of a possible screening test for the disease.

REFERENCES