An emerging role for comprehensive proteome analysis in human pregnancy research

Renu Shankar1, Neil Gude1,2, Fiona Cullinane1,2, Shaun Brennecke1,2, Anthony W Purcell3 and Eric K Moses4

1Department of Obstetrics and Gynaecology, The University of Melbourne, Melbourne, Australia, 2Pregnancy Research Centre, Department of Perinatal Medicine, The Royal Women’s Hospital, Melbourne, Australia, 3Department of Biochemistry and Molecular Biology, The University of Melbourne, Melbourne, Australia and 4Complex Disease Genetics Laboratory, Southwest Foundation for Biomedical Research, San Antonio, Texas, USA

Correspondence should be addressed to R Shankar, Pregnancy Research Centre, Department of Perinatal Medicine, The Royal Women’s Hospital, 132 Grattan Street, Carlton, Melbourne 3053, Australia; Email: r.shankar@pgrad.unimelb.edu.au

Abstract

Elucidation of underlying cellular and molecular mechanisms is pivotal to the comprehension of biological systems. The successful progression of processes such as pregnancy and parturition depends on the complex interactions between numerous biological molecules especially within the uterine microenvironment. The tissue- and stage-specific expression of these biomolecules is intricately linked to and modulated by several endogenous and exogenous factors. Malfunctions may manifest as pregnancy disorders such as preterm labour, pre-eclampsia and fetal growth restriction that are major contributors to maternal and perinatal morbidity and mortality. Despite the immense amount of information available, our understanding of several aspects of these physiological processes remains incomplete. This translates into significant difficulties in the timely diagnosis and effective treatment of pregnancy-related complications. However, the emergence of powerful mass spectrometry-based proteomic techniques capable of identifying and characterizing multiple proteins simultaneously has added a new dimension to the field of biomedical research. Application of these high throughput methodologies with more conventional techniques in pregnancy-related research has begun to provide a novel perspective on the biochemical blueprint of pregnancy and its related disorders. Further, by enabling the identification of proteins specific to a disease process, proteomics is likely to contribute, not only to the comprehension of the underlying pathophysologies, but also to the clinical diagnosis of multifactorial pregnancy disorders. Although the application of this technology to pregnancy research is in its infancy, characterization of the cellular proteome, unearthing of functional networks and the identification of disease biomarkers can be expected to significantly improve maternal healthcare in the future.


Introduction

A vital component of the highly controlled and synchronized biochemical changes that take place within the maternal uterus and its feto-placental compartment during the course of a pregnancy is the stage-specific expression of a wide array of cellular factors. The apparently seamless transition from the active cellular proliferation and differentiation phase of early pregnancy to the fully developed phase of late pregnancy and the switch from quiescence to active contractility during labour, involves an intricate interplay of intracellular and extracellular factors including hormones, adhesion molecules, growth factors and immunomodulators. Although the precise nature of the complex cross talk between these molecules is yet to be understood, there is no doubt that maternal pregnancy-associated physiological modifications, fetal development and parturition are significantly influenced by the subtle interactions between these bio-molecules.

Our current understanding of mammalian gestation and parturition has largely arisen through the application of a range of biochemical, endocrinological and physiological approaches involving animal models and cell lines (King et al. 2000, Lee & Demayo 2004). Although such data cannot directly be extrapolated to human pregnancy, they have contributed significantly to our understanding of human pregnancy and parturition. Until recently, such research has concentrated on identifying and localizing individual factors when, in fact, most function dynamically and in partnership with others within a system. Physiological processes involve complex molecular interactions and the coordinated functioning of a multitude
of interconnected intracellular pathways that are regulated by receptors recognizing specific paracrine/autocrine signals. Several such signalling pathways are critical during pregnancy and the precise progression of morphogenesis, embryogenesis and labour illustrates the synchronicity of these cellular events.

**Integrating proteomics with existing research techniques**

Contemporary global genomic strategies that identify novel genes and/or sets of genes simultaneously have been used to generate gene expression profiles in varying physiological conditions. Methodologies such as differential display and cDNA microarrays have been applied to investigate processes such as trophoblast differentiation and invasion (Aronow et al. 2001), chorionic villi gene expression (Chen et al. 2002), gene expression during labour (Marvin et al. 2002, Ogita et al. 2004) and expression of specific genes in gestational tissues (Tashima et al. 2002), and to compare the expression profiles of related genes in normal and pathological pregnancies (Pang & Xing 2004). However, the genome does not reflect the enormity and complexity of the human proteome, and genomic studies do not consider the vast number of co- and post-translational modifications (PTMs), nor the interactions and the sub-cellular localization of proteins, all of which collectively impact on cellular function.

The diversity created by such protein modifications is augmented by the fact that proteins can form multiple subunit combinations each with a specific function, as depicted by molecules such as the integrin receptors (Shimaoka et al. 1993). Furthermore, distinct proteomes specific to different cells, tissues or biological systems may exist, necessitating an in-depth examination of protein expression to understand the cellular distribution and function of proteins.

With the growing realization that more detailed information about proteins is necessary to elucidate biological function, more laboratories are integrating sophisticated mass spectrometry-based proteomic methodologies with the traditional, time-tested approaches such as Western blotting and immunohistochemical detection. Proteomic strategies are based on protein expression, function and structure (Lim & Elenitoba-Johnson 2004, Marko-Varga & Fehniger 2004). At the basic level, there is the identification and cataloguing of all proteins present in a system using expression-profiling approaches. Specific cellular populations can also be isolated and used to generate cell-specific expression profiles (Hoang et al. 2001) and monitor biological responses (Sawicki et al. 2003), while differential expression profiling can link dynamic changes in protein expression to various physiological stimuli or during disease processes (Hanash 2003, Marko-Varga & Fehniger 2004). Functional proteomic approaches offer a more focused analysis and place the proteins in their biological context, profiling PTMs (Mann & Jensen 2003), mapping interactions and pathways, detecting localization and investigating protein complexes that modulate protein expression and activity (Forler et al. 2003, Huh et al. 2003, Ranish et al. 2003). Determination of the three-dimensional structure of proteins allows for identification of drug targets (Marko-Varga & Fehniger 2004) in addition to being complementary to functional proteomics. A brief overview of contemporary proteomic techniques follows.

**Mass spectrometry and protein/peptide characterization**

Two ionization techniques have revolutionized the way proteins and peptides are characterized and sequenced by mass spectrometry (MS): electrospray ionization (ESI)- and matrix-assisted laser desorption/ionization (MALDI)-based MS overcome many of the inherent problems associated with peptide analysis using earlier ionization techniques and instrumentation (Gygi & Aebersold 2000), including their relatively poor sensitivity, the requirement of sample volatility and the sensitivity of ionization to sample preparation. Although both ionization techniques provide highly useful and often complementary information about peptide analytes, most studies have used MALDI-TOF (time-of-flight) MS for protein identification using the peptide mass fingerprinting technique (Gygi & Aebersold 2000) whilst ESI-tandem MS experiments are most frequently used to sequence individual polypeptides. With the advent of hybrid instruments and readily interchangeable ionization sources this distinction is now less apparent, with many examples of peptide sequencing using both MALDI-TOF and MALDI-TOF/TOF MS and the use of shotgun proteomics and tandem ESI-MS used in combination with multi-dimensional chromatography (the Mud-PIT approach) (Washburn et al. 2001, Wolters et al. 2001, Wu & MacCoss 2002) to identify proteins.

MALDI is a relatively gentle ionization technique, which involves the use of laser-induced ionization of a sample embedded in a crystalline matrix (Hillenkamp et al. 1991). The matrix absorbs at the laser wavelength and excitation of the molecules within the matrix leads to desorption (Spengler et al. 1993). The desorbed analyte ions are then accelerated into the mass analyser. The principal ion detected in MALDI is a singly charged parent ion (M + H)+; thus MALDI generates simple spectra for individual species and mixtures of polypeptides. At low analyte concentrations, MALDI is more tolerant than ESI towards organic and inorganic sample impurities, including detergents, denaturants and buffers commonly used in the isolation of polypeptides (Spengler et al. 1993). MALDI is a discontinuous ionization technique and analyte ions are only formed each time the laser is fired. Thus multiple laser shots are used to ionise the analyte and the data are collected on a cumulative basis.

The electrospray and ion-spray processes are also gentle ionization techniques in which the sample is pumped through a charged narrow capillary (Fenn et al. 1989).
As the liquid and analyte are charged, the mixture forms droplets and sprays from the orifice of the capillary. In ion-spray ionization, droplet formation is accelerated by using a nebulizing gas flow in an annular sheath surrounding the spraying needle. A counter-flow of dry gas assists sample evaporation and the droplets decrease in size until they become unstable and undergo a coulombic explosion to form even finer droplets. Ultimately, electrostatic repulsion is sufficient to cause desorption of the analyte ions, which are directed into the mass spectrometer. Unlike MALDI, ions generated by ESI usually bear multiple charges \((M+nH)^{n+}\), with several charge states apparent for some peptides. Thus ESI spectra are frequently more complex than those generated by MALDI. Being a continuous flow-based technique, ESI is amenable to on-line detection for liquid chromatography (LC) in LC-MS experiments.

Whilst ESI typically uses scanning analysers such as quadrupoles, MALDI requires a discontinuous method of analysis. For this reason time-of-flight (TOF) analysers are most commonly used with this form of ionization. The TOF analyser is a simple device in which ions are accelerated from the ion source into a long field-free tube and subsequently travel at constant velocity onto a detector. The time taken for the ions to pass down the tube is measured from the laser pulse to meeting the detector at the other end of the tube. The \(m/z\) ratio of each ion in the sample is then calculated based on the time of flight (which is proportional to \(m/z^2\)). Further enhancement of resolution is achieved in instruments with an electrostatic mirror or reflector located at the end of the linear flight tube that refocuses ions onto another trajectory so as to impact with an alternative detector. Instruments fitted with a reflector enable linear and reflector modes of analyses and are suitable for peptide sequencing.

More recently, MALDI-TOF/TOF MS technology (Medzihradzsky et al. 2000, Rejtar et al. 2002, Suckau et al. 2003) has been introduced, allowing very accurate ion selection using one TOF analyser, fragmentation in an adjoining collision cell and high resolution analysis of ion fragments in the second TOF analyser yielding high quality sequence information (Rejtar et al. 2004). The MALDI-TOF/TOF MS is now arguably the technique of choice for high-throughput protein identification (Pan et al. 2005). Another emerging technology incorporates Fourier transform (FT) or ion cyclotron resonance MS (Bergquist 2003, Heeren et al. 2004, Hopfgartner et al. 2004, Page et al. 2004, Ramstrom & Bergquist 2004, Schrader & Klein 2004). These instruments are capable of extremely accurate mass determination and have been used for top down sequencing (fragmentation analysis of intact proteins) (Lin et al. 2003) as well as high accuracy MS and tandem MS (MS/MS) experiments. As these instruments become more popular, and with hybrid instruments such as the triple quadrupole linear ion trap-FT-MS instruments emerging as proteomic tools (Hopfgartner et al. 2004), these instruments may form the vanguard of proteomics analysis allowing combined high resolution peptide mass fingerprinting, accurate PTM analysis and de novo sequencing of peptides and small proteins (Yergey et al. 2002). They also have great application in functional proteomic studies (Foster et al. 2003, Leung et al. 2002, Rappsilber et al. 2003, Olsen & Mann 2004).

**Ion fragmentation and peptide sequencing**

Modulating the conditions of ion formation and subsequent manipulation of these ions can induce fragmentation during both MALDI and ESI, which, in the case of peptides, leads to nested sets of ion fragments from which amino acid sequence information can be derived (Purcell & Gorman 2001). For ESI-based techniques, tandem mass spectrometry or MS/MS instrumentation can be used to generate and analyse fragments. These instruments consist of two mass analysers arranged in tandem and separated by a collision cell. Thus, in an MS/MS instrument the first mass analyser can be used to select a particular ionic species of defined mass. The selected ion is subsequently channelled into the collision cell that usually contains an inert gas. Collision with the gas molecules induces fragmentation and ion fragments are analysed in the second mass analyser. The most common instruments of this type of collision-induced dissociation (CID) MS/MS experiment are the triple quadrupole instruments. Other types of mass analysers may be used in tandem such as those used in hybrid instruments such as electrospray ionization-quadrupole-quadrupole-time-of-flight (ESI-QqTOF) MS instruments.

Another type of mass analyser that is useful for MS/MS experiments using ESI sources are the ion-trap instruments, which contain a single mass analyser (the ion trap) but can perform MS/MS and additional MS\(^n\) stages of fragmentation by virtue of their ability to trap ions, eject unwanted ions and perform the fragmentation post selection all within the one analyser (effectively separating the MS/MS experiment in time rather than by space as is the case for triple quadrupole instruments).

Fragmentation of ions in MALDI need not involve CID, since the ions generated by MALDI may be induced to undergo fragmentation in the ion source (in-source decay, ISD) or after leaving the ion source (post-source decay, PSD). Laser intensity is varied to induce both ISD and PSD, which are both mediated by collisions with the matrix cloud formed by MALDI. An inert gas may also be used in the source to induce CID fragmentation during MALDI.

**Strategies for protein identification**

In order to visualize the proteome of a cell, robust and high-resolution separation techniques are required to first separate the protein components to allow comparative studies. The most common front-end technology for proteomics studies is undoubtedly 2-dimensional gel electrophoresis (2DGE) in which, typically, proteins are first
subjected to isoelectric focusing (IEF) within a tube gel or on an immobilized pH gradient (IPG) strip. Following this focusing step, the proteins separated in this first dimension are subject to a second dimension typically of SDS-PAGE. Thus, in a standard IEF-SDS-PAGE analysis proteins are separated by isoelectric point and molecular size. Under appropriate conditions individual proteins can be visualized as spots on 2-D gels. These spots can subsequently be excised, an in-gel proteolysis step can be performed, typically using trypsin, and the tryptic fragments analysed by mass spectrometry. For proteins derived from an organism of known genome, the characteristic masses of tryptic fragments can be unique for the different theoretical gene products, allowing accurate determination of protein identity. This strategy, known as peptide mass fingerprinting (PMF), has been the staple tool of protein identification. However, more detailed and sensitive analysis frequently demands more stringent identification criteria. This typically combines PMF with sequencing of individual tryptic fragments to give more statistically reliable identification. This form of analysis can also help to identify sites of post-translational modification.

A limitation of 2-D gels is their resolution, with typically less than 2000 spots visible and a relatively poor dynamic range. Although strategies exist to remove abundant proteins from biological samples and pre-fractionation of cellular extracts is available to produce proteome subsets (e.g. nuclear, cytosolic or membrane protein fractionation), other separation techniques have been explored with great success. Thus, as an alternative or complementary separation technology, liquid chromatography is increasingly being used to resolve a complex proteome into fractions for analysis. This can vary from fractionation of solubilized extracts by various affinity or interactive modes of chromatography prior to further analysis by 2DGE or additional chromatographic steps, through to the high-throughput multi-dimensional chromatography protein identification technology (MuDPIT approach) pioneered by the Yates group (Washburn et al. 2001, Wolters et al. 2001, Wu & MacCoss 2002). In MuDPIT experiments the whole mixture of proteins are first digested into smaller fragments typically using trypsin. The highly complex mixture of proteolytic fragments then undergoes a form of shotgun sequencing that relies on the ability of at least two dimensions of chromatographic separation to resolve the mixture sufficiently to allow MS/MS-based sequencing of tryptic peptides by on-line LC-MS/MS analysis. The use of automated algorithms for peptide sequence assignment and collation of the huge datasets generated by this technique is obligatory.

**Differential display and quantitative analysis**

Although a robust front-end technology, subtle changes in protein expression are often difficult to observe by gel-to-gel comparisons. In order to visualize subtle changes in protein composition in different samples, the use of differential image gel electrophoresis (DIGE) technology has been developed. This analysis incorporates a combination of three Cy dyes including Cy2, Cy3, and Cy5 with different emission wavelengths (blue, green and red in colour respectively), for protein differential displays. Three dyes allow the simultaneous comparison of three samples in a single 2-D gel, minimizing errors introduced by gel-to-gel comparison. Typically, one colour is used as an internal standard to control for protein loading. The inclusion of the internal standard (aliquots of each sample pooled into one) allows each resolved protein to be compared with itself within the internal standard to generate a ratio of relative expression. The same internal standard is run on each gel within an experimental series allowing the ratio of relative expression of the same protein across several gels to be compared directly, distinguishing gel-to-gel variation from biological variation. Accordingly, quite small differences in expression levels can be determined by comparing the ratio obtained from one fluorescent-labelled sample directly with another. As a result, it is possible to see <10% differences in protein expression between samples, with >95% statistical confidence (Patton 2002).

Likewise, two samples used for 2-dimensional liquid chromatography (2DLC) experiments can be labelled and run simultaneously to allow quantitation of individual species in a sample. This can be accomplished using isotopically coded affinity tags (ICAT)-based technology, which involves labelling a subset of tryptic peptides from a protein via cysteine residues with ICAT. The tags incorporate a ‘heavy’ and a ‘light’ form, added to respective samples for comparison during an alkylation step that labels cysteine residues found within certain tryptic fragments. ICAT reagents contain a biotin moiety allowing affinity purification of labelled peptides post-tryptic digestion. New generation ICAT are acid-cleavable (allowing removal of the biotin group) and incorporate $^{13}$C rather than deuterium into the tags, thereby improving co-elution of ‘heavy’ and ‘light’ tagged species during RP-HPLC. Other approaches that involve stable isotope labelling of other amino acids or similar isotopically labelled reagents of different coupling chemistries are becoming common (Ong et al. 2002).

**Proteomic studies on early pregnancy**

The establishment of pregnancy starts with the adhesion of an implantation-competent blastocyst to the deciduaized endometrium of the uterus. Contact between the outer epithelial trophoderm of the differentiated blastocyst and the uterine luminal epithelium, initiates invasion and differentiation of the uterine stromal cells into decidual cells, leading to implantation (Dey et al. 2004). The receptivity of the uterine endometrium to the blastocyst and the process of implantation are modulated by oestrogen and progesterone (Rider 2002, Ma et al. 2003). Dysfunction
during the preimplantation period can affect not only the development of the fetus, but may also result in faulty placentation and lead to associated pregnancy complications. Recent genomic and proteomic studies have focused on signalling pathways associated with blastocyst formation and the uterine receptivity to blastocyst implantation (Daikoku et al. 2005, Natale et al. 2004). A host of proteins derived from the blastocyst, decidua, placenta and the extracellular matrix (ECM) have been implicated in the cell-to-cell and cell-to-ECM interactions essential for a successful implantation to be carried to term (Lala & Hamilton 1996, Aplin 1997, Kimber 2000, Lindhard et al. 2002, Dey et al. 2005, Natale et al. 2004). Previous studies had speculated that factors regulated by Hoxa10, a transcription factor associated with preimplantation embryonic viability (Satokata et al. 1995) and with progesterone-mediated uterine stromal cell proliferation (Lim et al. 1999), might be associated with the impaired decidualization and implantation (Satokata et al. 1995).

Daikoku and colleagues (2005) applied an initial quantitative proteomic differential display method, DIGE, in conjunction with MALDI-TOF/TOF to identify proteins linked to the Hoxa10 signalling pathway and then used conventional techniques to delineate their localization and function. Comparison of the protein expression profiles generated from purified stromal cells obtained from pseudopregnant uteri of wild-type and Hoxa10-deficient mutant mice (Hoxa10−/−) revealed significant differences in expression among twenty-nine proteins including actin-binding proteins associated with decidualization (Shaw et al. 1998). FBKP52 (FK506 binding protein 4), an Hsp90 binding immunophilin and a co-chaperone of steroid hormone receptors (Barent et al. 1998), was observed to be downregulated in Hoxa10−/− mice and was selected for further investigation. Using in situ hybridization techniques, the authors demonstrated the cell-specific spatiotemporal expression and differential hormone regulation of FBKP52. FBKP52 expression was detected in both stromal and epithelial cells on days 1 and 4. However, on day 5 when the trophocoeptoderm attaches to the uterine epithelium and decidualization begins, Hoxa10 and FBKP52 were detected only in the stromal cells. The progesterone-enhanced stromal expression of FBKP52 required the presence of Hoxa10 and nuclear progesterone receptor, while epithelial expression was mediated by oestrogen. Stromal FBKP52 expression in the secondary decidual zone was observed to continue after implantation, suggesting a role for this protein in decidualization.

The investigation of the cellular mechanisms associated with the invasion and migration of the trophoblast cells into the uterine tissue is another area that generates research interest. The precise regulation of trophoblast invasion into the uterine epithelium and the extracellular matrix of the underlying uterine stroma (Aplin 1991), and trophoblast differentiation into the villous and extravillous pathways (Malassine & Cronier 2002) are integral to normal placentation and a favourable pregnancy outcome. These processes are mediated by the actions of inflammatory cytokines (Benyo et al. 1997), angiogenic growth factors (Lash et al. 1999) and adhesion molecules (Zhou et al. 1997, Goldman-Wohl & Yagel 2002). Normal extravillous trophoblast invasion is restricted to the inner third of the uterine myometrium (von Rango et al. 2003) and is associated with the remodelling of the maternal spiral arteries (Craven et al. 1998). Abnormalities at this stage can result in a number of pathologies including those linked to placental dysfunction. Uncontrolled invasion causes gestational trophoblast disease and gestational trophoblast tumours while shallow invasion is associated with fetal growth restriction and pre-eclampsia (Goldman-Wohl & Yagel 2002, Anin et al. 2004). The latter phenotype exhibits the narrow, untransformed spiral artery formation, responsible for the reduced uteroplacental blood flow and oxidative stress observed in pre-eclampsia and fetal growth restriction (Takagi et al. 2004).

Lowered oxygen tension is known to affect trophoblast differentiation (Genbacev et al. 1997) through the transcription factor hypoxia-inducible factor-1 and transforming growth factor β3 (Caniggia et al. 2000). In an attempt to investigate the processes associated with abnormal cytotrophoblast differentiation at the protein level, 2DGE was performed on a cell culture model that replicated the first trimester placental cytotrophoblast phenotype in vitro (Hoang et al. 2001). Lowering the oxygen tension to reflect the hypoxic condition of pre-eclampsia effected a change in less than 3% of the proteins expressed on the gel, indicating operational safeguard mechanisms that prevent oxidative stress. The expected downregulation of antioxidants and increase in glycolytic enzymes was accompanied by a novel finding. There was a significant increase in the level of annexin II, a protein associated with proliferation (Menaa et al. 1999) and fibrinolysis, under hypoxic conditions (Rao et al. 1994). Functional proteomic studies have identified annexin II as a protein that binds to placental protein 13 (PP13)/galectin 13, whose dimerization may have a role in affecting oxygen changes in the placenta (Than et al. 2004).

A more recent comparison of the expression patterns of placental trophoblasts from pre-eclamptic women has also shown differences in abundance in proteins related to hypoxia and oxidative stress (Jin et al. 2004). Another study investigating the regulation of hypoxia in term placental cytotrophoblasts demonstrated the effect of neurokinin B, a placental peptide the levels of which are substantially increased in pre-eclamptic women, on the cytotrophoblast protein profile (Sawicki et al. 2003). Neurokinin B was observed to suppress expression of proteins that counteract the effect of hypoxia including annexin II. In addition, the inhibition of other functionally related proteins brought to light the role of neurokinin B in intravascular coagulation, inflammatory response to infection, apoptosis and other cellular processes. This simultaneous display of the overall effect of a factor in the context of a
physiological condition emphasizes the advantages of the proteomic approach over the more conventional protein detection methods.

Clinical proteomics and pregnancy disorders

From the cellular context, the changes specific to a disease may be confined to a very small fraction of its protein repertoire. Identification of these proteins can be of use in elucidating the molecular basis of the pathology. There may also be suitable disease markers that, if detected at the early stages or prior to clinical presentation, can indicate a prophylactic strategy to be implemented or improve the efficacy of an intervention. Those detected at the later stages are likely to be more specific and may be directly or closely related to the phenotype of the disease. The discovery of biomarkers is therefore an area that has seen significant activity with respect to application of proteomic technologies (Carrette et al. 2003, Petricoin & Liotta 2003). One of the techniques that has generated considerable interest with its clinical potential is proteomic pattern diagnostics that does not require protein identification (Petricoin & Liotta 2003). The clinical advantage of this method is that minute amounts of biological fluids can be assessed for a protein profile generated by mass spectrometry that may be distinctive to the pathology under consideration. Patterns are validated against algorithms specific to the particular disease and change is examined in the proteomic context rather than in isolation. Measurement of the entire set of functionally altered proteins allows the representation of interactions that are specific to the particular pathology and the resulting spectrum is assumed to be unique to the disease. From this viewpoint, evaluation of the overall profile seems a superior assessment of the physiological response to the disease. However, one of the challenges is that of pattern reproducibility in the face of experimental and individual variations. The second concern is the low sensitivity of current mass spectrometers that make them less suitable as tools in routine clinical analysis when compared with, for example, immunoassays (Diamandis 2004, Garber 2004).

Investigation of gestational tissues and bio-fluids and secretions is expected to generate a broader picture of the physiology and pathophysiology of pregnancy. Pregnancy related disorders such as preterm labour, pre-eclampsia, fetal growth restriction and gestational diabetes mellitus (GDM) are major contributors to maternal, perinatal and/or neonatal morbidity and mortality (Schmidt et al. 2001, Masse et al. 2002, Lumley 2003, Garite et al. 2004). Each of these disorders has a prevalence of approximately 5 to 10% in a general obstetric population (WHO 1987, King 1998, de Onis et al. 1998, Lumley 2003), they have multifactorial aetiologies (Friedman et al. 1991, Romero et al. 1993, Lin & Santolaya-Forgas 1998, Ben-Harouch et al. 2004) and they are very often difficult to predict. DNA sequence variations and polymorphisms caused by exogenous or endogenous factors can cause functional differences at the protein level, limiting the ability of genetic tests to predict the risk of multifactorial disorders of pregnancy such as pre-eclampsia and preterm labour (Shimizu & Bryant-Greenwood 2004). Several proteins have been observed to have significant associations with these pathologies (Cooper et al. 1993, Masse et al. 2002, Page et al. 2002, Goldenberg et al. 2003). However, thus far they have not demonstrated the sensitivity, specificity or predictive values required for accurate detection of women at risk for these disorders. Over half of the women in preterm labour have no identifiable clinical risk factor. This inability to accurately predict those who are at risk of delivering preterm can result in unnecessary treatment. The lack of specific diagnostic markers confounds the diagnosis of pre-eclampsia and gestational diabetes whose clinical signs and symptoms are similar to that of gestation independent hypertension and diabetes.

The search for novel/candidate biomarkers using proteomic approaches is an avenue that is now being actively exploited in pregnancy research. Proteins present in biological fluids that can be accessed as non-invasively as possible are ideal disease biomarker candidates. Efforts are underway to identify serum markers of ectopic pregnancy (Gerton et al. 2004). Novel and specific proteolytic fragments of insulin-like growth factor-binding protein 1 (IGFBP-1) and calgranulin B have been proposed as biomarkers of intra-amniotic infection (IAI) based on results obtained by the use of proteomic methods on amniotic fluid samples collected from women and primates with subclinical chorioamnionitis (Gravett et al. 2004). This study identified regions with differences in peak intensity in the protein profile using a rapid low resolution surface-enhanced laser desorption ionization (SELDI) method coupled with mass spectrometry, followed by liquid chromatography–tandem mass spectrometry (LC-MS/MS) identification of proteins in tryptic digests obtained from one-dimensional gels of the samples and validation with Western blotting. Intrauterine infection shows a strong association with spontaneous preterm labour but is often clinically undetectable and intervention is futile when clinical presentation is delayed (Goldenberg et al. 2000). Amniotic fluid levels of IGFBP-1 have been observed to increase at the beginning of the second trimester when the amnion fuses with the chorion–decidua (Wathen et al. 1993) and the presence of amniotic fluid isoforms of IGFBP-1 in the cervicovaginal fluid has been observed to be predictive of intrauterine infection, premature rupture of the fetal membranes (Rutanen et al. 1993) and puerperal infectious morbidity (Kekki et al. 2001). Calgranulin has not been linked to intra-amniotic infection, but is an inflammatory protein (Roth et al. 2003). The study by Gravett et al. (2004) also demonstrated the feasibility of using these proteins as serum markers of IAI, although this needs to be validated with large prospective studies.

Two heparin sulphate proteoglycans, agrin and perlecans, identified by 2DGE in a study by Vuadens et al. (2003) were suggested as potential biomarkers of premature
rupture of fetal membranes (PROM), a condition whose pathology is different to that of spontaneous preterm labour (Fortunato & Menon 2001) and which is responsible for a quarter of preterm deliveries. These two proteins have been identified solely in the amniotic fluid and may therefore be able to decrease the microscopic maternal blood contamination encountered when testing for PROM.

The structural integrity of the decidual–fetal membrane–placental unit is crucial to the maintenance of pregnancy and the progressive decline in adhesiveness towards parturition is a carefully modulated event. Several cytokines, growth factors and proteases are involved in the decidual–placental interaction (Carbillon et al. 2000) and decidual lymphocytes have been associated with placental separation from the decidua during parturition (Abadia-Molina et al. 1997). Molecules associated with maternal–fetal adhesion such as the glycoprotein fetal fibronectin, are usually found in the cervicovaginal secretions closer to parturition (Lockwood et al. 1991). The presence of fetal fibronectin in the cervicovaginal fluid between 22 and 34 weeks of gestation is suggested to be indicative of premature disruption at the choriodecidual interface (Lockwood et al. 1991). Despite its relatively low positive predictive value, this is one of the best, currently available, biochemical diagnostic tests for preterm labour, demonstrating its clinical utility with a high negative predictive value that is used to distinguish false positive cases of impending preterm labour (Chuileannain et al. 1998). Immunoassays have established the association of several solitary factors, including modulators of adhesion, and specific combinations of factors, with preterm labour (Goldenberg et al. 2003, Urban et al. 2003, Torbe & Czajka 2004). The protein profile of relevant biofluids can highlight those proteins that might be associated with labour/impending labour. Our own initial 2DGE analysis has revealed a distinct protein profile for cervicovaginal fluid proteins in samples obtained from women presenting with preterm labour (Fig. 1).

Elucidating the molecular basis of placental function is central to understanding the role of the placenta in the pathogenesis of several pregnancy disorders. Changes in placental function have been measured both in vivo and in vitro in pregnancy disorders. For instance, placental glucose transfer and metabolism is altered in GDM (Osmond et al. 2000), abnormal intervillous blood flow is detected in early pregnancy loss (Jauniaux et al. 2003) and placental factors are associated with maternal vascular endothelial cell dysfunction in pre-eclampsia (Cooper et al. 1993). The complexity of the placenta, an organ that contains tissue and blood of two separate individuals, can confound investigations aimed at characterizing the human placental proteome. Bilateral perfusion of a placental cotyledon is therefore an ideal model for study of maternal–fetal transport processes (Osmond et al. 2000), regulation of fetal blood flow (Gude et al. 1998) and secretion of factors into the maternal and fetal circulations (Gude et al. 1991). The characterization of the proteome of the effluents of ex vivo perfused placentae is therefore likely to provide a valuable insight into the repertoire of proteins secreted by the placenta in vivo and into the role of the placenta as an endocrine organ during pregnancy. In our laboratory, analyses of concentrated fractions of maternal and fetal effluents by 2DGE have rendered a complex pattern of protein spots (Fig. 2) adding a further dimension to proteomic studies that involve the use of serum, whole tissue or cultured cells. Comparison of effluents from normal pregnancies with those from pathological pregnancies may enable the identification of novel proteins that may be secreted into either the maternal or fetal circulations during pregnancy disorders.

**Figure 1** Silver stained 2-D gel (pH range 5–9, 4–20% polyacrylamide) display of proteins from cervicovaginal fluid.

**Figure 2** Coomassie Blue stained 2-D gel (pH range 5 to 9, 10 to 20% polyacrylamide) displays of proteins from a fetal effluent sample (top) and a maternal effluent sample (bottom).
Can proteomics deliver?

Over the last few years there has been a shift towards large-scale proteomic research in several areas of biological research. The Human Proteome Organization (HUPO) was established to globalize and integrate proteomic techniques, and within a short time span it has diversified into distinct disciplines for investigating subsets of the proteome (HUPO 2004). New technologies with high throughput have evolved at a rapid pace and are frequently being updated for increased efficiency and simplicity of use, while the reproducibility and robustness of existing methodologies, like 2DGE, is being continuously improved. Limitations of gel-based proteomic methods to display low abundance proteins, together with their expression at detection levels that challenge the sensitivity range of mass spectrometers, paved the way for gel-independent approaches such as protein chip microarrays (Espina et al. 2004). The effectiveness of these techniques is very closely linked to the precision of the bioinformatics tools used to decipher the enormous volumes of data generated.

The application of proteomic methods has yielded significant information relevant to many pathological conditions and has identified several, hitherto unknown, proteins (Carrette et al. 2003, Dumont et al. 2004, Gronborg et al. 2004, Liao et al. 2004, Pitarch et al. 2004). Nevertheless, despite the accessibility of some gestational tissues such as the placenta and fetal membranes, very few published studies so far have incorporated these methodologies in pregnancy research. While a considerable amount of data is available on several aspects of pregnancy including implantation, trophoblast differentiation and parturition, the understanding of molecular connections between the multiple processes occurring during pregnancy is far from complete. A comprehensive examination of tissues at the maternal–fetal interface will be necessary to understand the progressive changes that occur during pregnancy. Incorporation of these methodologies with existing ones will provide a novel perspective on the molecular constitution and dynamics of gestational tissues. The use of tissue-specific gene knockout strategies has been suggested for collection of streamlined data uncomplicated by compensatory interactions/mechanisms (Bernal 2001). Homogenous cell types can be isolated using techniques such as flow cytometry (Vince et al. 1990) and laser microdissection (Craven & Banks 2002). The latter method is compatible with 2DGE as well as with direct MS and may be used with frozen or paraffin-embedded tissues. However, the amount of microdissected material obtained is relatively small and the amount of protein that can be recovered depends on the type of tissue, fixative and stain used (Craven & Banks 2002). The study of specific cell types could be beneficial in comprehending the stage-specific functional variations exhibited by gestational tissues. Cell populations from tissues such as decidua and placenta show gestational age differences during pregnancy. For instance, natural killer (NK) cells constitute around 50–90% of decidual lymphocytes in early pregnancy (Koopman et al. 2003) and are associated with trophoblast invasion, placenta and maternal immunotolerance of the fetus (Burrows et al. 1995, Lanier 1999). Towards term, however, the number of NK cells and T cells decreases drastically and the alteration in the regulatory T cell population has been associated with labour (Sindram-Trujillo et al. 2004). There is increasing evidence that labour is an inflammatory process (Marvin et al. 2002, Osman et al. 2003) closely associated with elevated levels of cortisol, oestrogen and prostaglandins. It may also be influenced by factors from the fetus (Gibb & Challis 2002). The molecular mechanisms behind these processes remain unclear. Insights into the mechanism behind immunotolerance may contribute to understanding of some autoimmune diseases (Giacomelli et al. 2004) as well as transplant rejection. Leukocyte proteomic analysis that has the potential to uncover the mechanism of inflammation (Wang et al. 2004) may be especially important to pregnancy and parturition.

Disease proteomics is likely to have a significant impact on clinical research and medicine. Matrix-assisted (MALDI) and/or surface-enhanced (SELDI) laser desorption ionization methods have been used in combination with mass spectrometry to profile various tissues, bio-fluids and tissue sections (Aldred et al. 2004, Chaurand et al. 2004) to discover variations that contribute to the diseased state. Studies are identifying altered expression of proteins that have not previously been associated with the specific condition, for example the cytotrophoblastic expression of annexin II during hypoxia (Hoang et al. 2001), but may have a role in the pathophysiology of disorders associated with hypoxia and oxidative stress such as pre-eclampsia (Takagi et al. 2004). The availability of larger integrated global datasets will facilitate deduction of functional relationships, and as more information about the signalling and cellular communication unfolds, drug targets can be identified. Techniques such as imaging mass spectrometry, where the distribution of several proteins in a tissue section can be detected simultaneously without the need for specific antibodies or protein identities, allow drug response monitoring (Reyzer et al. 2003, Chaurand et al. 2004, Schwartz et al. 2004). The surge in studies identifying proteins through mass spectrometric methods in various pathologies has prompted more stringent validation using complementary methods like Western blotting, immunoassays and other immunological techniques. However, filtering of mass spectrometric results based on <5% peptide identification prior to validation and a necessity for excessive validation can result in loss of significant information (Hancock 2004, Veenstra et al. 2004). For proteins identified as potential biomarkers to have any clinical utility, validations with large-scale prospective clinical studies are required. It is also necessary to be able to find an association between the proteome...
and the clinical phenotype for implementing any intervention (Shimizu & Bryant-Greenwood 2004).

Finally, the challenges facing current proteomic methodologies are numerous, and multiple approaches are required to examine cellular composition, behaviour and function. Artificial intelligence systems have to be honed for better management of data, easy access and tracking, as well as for recognising the multiple proteomic patterns that can arise for a single pathology given the heterogeneous nature of both disease and population. The inability of proteomic patterns to detect protein markers previously identified in the disease is a matter for concern (Garber 2004) and protocols need to be continuously enhanced for better reproducibility. The sensitivity of mass spectrometers is still an issue and the identification of low copy proteins is extremely difficult. Nevertheless, as Perticon said, it would be wrong not to ‘... investigate anything that seems to be able to discriminate disease from non-disease’ (Garber 2004).

Acknowledgements

The production of this review paper has been supported by a scholarship (to R S) and a grant (to F C) from the University of Melbourne as well as research funding (to R S) from the Royal Women’s Hospital. We also gratefully acknowledge Janet Stevenson, May Orgurinovic and Vicki Doherty at the Royal Women’s Hospital, Pregnancy Research Centre for their technical assistance. The authors declare that there is no conflict of interest that would prejudice the impartiality of this scientific work.

References


Bernal AL 2001 Overview of current research in parrition. Experimental Physiology 86 213–222.


Foster LJ, De Hoog CL & Mann M 2003 Unbiased quantitative proteomics of lipid rafts reveals high specificity for signaling factors. PNAS 100 5813–5818.

www.reproduction-online.org


Downloaded from Bioscientifica.com at 12/12/2018 04:30:13AM via free access


Lanier LL 1999 Natural killer cells fertile with receptors for HLA-G. Nature 396 5343–5345.


Lim MS & Elenitoba-Johnson KS 2004 Proteomics in pathology research. Laboratory Investigation 84 1227–1244.

Lim H, Ma L, Ma WG, Maas RL & Dey SK 1999 Hoxa-10 regulates uterine stromal cell responsiveness to progesterone during implantation and decidualization in the mouse. Molecular Endocrinology 13 1005–1017.


Ma WG, Song H, Das SK, Paria BC & Dey SK 2003 Estrogen is a critical determinant that specifies the duration of the window of uterine receptivity for implantation. PNAS 100 2963–2968.


Olsen JV & Mann M 2004 Improved peptide identification in proteomics by two consecutive stages of mass spectrometric fragmentation. PNAS 101 13417–13422.


Rejtar T, Chen HS, Andreev V, Moskovets E & Karger BL 2004 Increased identification of peptides by enhanced data processing of high-resolution MALDI-TOF/TOF mass spectra prior to database searching. Analytical Chemistry 76 6017–6028.


