A proteomic analysis of mammalian preimplantation embryonic development

Mandy G Katz-Jaffe, Donald W Linck, William B Schoolcraft and David K Gardner

Colorado Center for Reproductive Medicine, 799 E Hampden Ave, Suite 520, Englewood, Colorado 80113, USA

Correspondence should be addressed to M Katz-Jaffe; Email: mkatz-jaffe@colocrm.com

Abstract

Genetic studies on the mammalian preimplantation embryo are providing a wealth of information regarding gene expression. However, changes in the transcriptome do not always reflect cellular function or the complexity and diversity of the mammalian proteome with post-translational modifications or protein–protein interactions. To elucidate embryonic cellular function, a detailed understanding at the protein level is necessary. The aim of this study was to generate protein profiles of mammalian embryos throughout development, and to investigate the effects of oxygen concentration on the embryonic proteome. A protocol was developed to analyse small groups of embryos (n=5) by time-of-flight mass spectrometry. F1 mice zygotes were cultured in G1/G2 sequential media with recombinant albumin (2.5 mg/ml) in 6% CO2 and O2 concentrations of either 5% or 20%. In vivo-developed embryos were flushed from the reproductive tract (day 4). Protein profiles were generated for all embryonic samples and statistical analysis revealed 32 potential proteins/biomarkers with significant changes (P<0.05).

Embryos generated under 5% O2 more closely resembled in vivo-developed embryos. Under 20% O2 conditions, embryos showed down-regulation of 10 proteins/biomarkers (masses between 4 to 20 kDa) (P<0.05) confirming the pathological effects of oxygen during embryonic development. These data demonstrate for the first time the complexity of the mammalian preimplantation proteome. The unique protein profiles of in vivo-developed embryos and a panel of selected biomarkers represent optimal cellular function, against which comparisons can be made to facilitate improvements in mammalian assisted reproduction techniques procedures.

Reproduction (2005) 130 899–905

Introduction

Gene expression analysis is providing valuable information about the activity of the mammalian genome. Nevertheless, genes exert their effect through proteins after transcription and translation, thus genomics do not readily predict protein abundance or provide information about protein function at the biochemical level. Due to the central role that proteins play in biology and physiology, it is paramount to investigate protein abundance and activity. However, unlike nucleic acids, proteins are diverse and complex, displaying a variety of chemistries, affinities and specificities, as well as associations with other proteins and post-translational modifications.

To date, research has concentrated on identifying and localising individual proteins; however physiological processes predominantly involve more than a single biochemical pathway or molecular interaction. The classic proteomics approach involves 2D polyacrylamide gel electrophoresis, requires large amounts of starting material, is complex, labour-intensive and not capable of high throughput analysis (Latham et al. 1992, Shi et al. 1994). In addition, proteins with low or high molecular masses or that are very acidic, basic or hydrophobic can be under-represented in 2D gels. Recent developments in mass spectrometry have been revolutionary, utilising expression profiling and peptide sequencing to identify the proteins that are expressed and that function in a biological system (Hale et al. 2003, Shanker et al. 2005). Comparative protein profiling has been developed for the detection of specific protein expression patterns reflecting varying biological states (Shau et al. 2003, Rocken et al. 2004). One method, surface-enhanced laser desorption/ionisation time-of-flight mass spectrometry (SELDI-TOF MS), involves affinity-based mass spectrometry whereby proteins are absorbed to a chemically modified surface (e.g. cationic or anionic) or to a biochemical molecular surface (e.g. receptors or ligands). These different chip surfaces allow the different classes of proteins (hydrophobic, hydrophilic, acidic or basic) to be captured for analysis. Laser activation irradiates the sample and desorption/ionisation liberates gaseous ions to a constant final energy.
Time-of-flight mass spectrometry then accurately determines the mass-to-charge ratio (m/z) of the bound protein by the time it takes for the ions to pass through the long flight tube from the laser pulse to the detector (Merchant & Weinberger 2000). SELDI-TOF MS has been shown to capture, detect and analyse proteins directly from crude biological fluids, such as serum (Seibert et al. 2004, Buhimschi et al. 2005, Xiao et al. 2005). It is a highly sensitive, high throughput and cost-effective method that has been used to identify biomarkers/changes in protein expression from varying disease states including early detection for ovarian cancer (Zhang et al. 2004), cervical cancer (Wong et al. 2004) and kidney cancer (Junker et al. 2005). Biomarkers are defined as candidate proteins or peptides that are either down- or up-regulated in response to different physiological states.

In order to elucidate embryonic cellular function, a detailed understanding at the protein level is necessary. The major technical issues when it comes to studying the mammalian embryonic proteome include the constraint of requiring large amounts of sample and the difficulty in obtaining sufficient biological material to identify low-abundance proteins by mass spectrometry. To date, studies have focused on investigating individual proteins using large numbers of embryos. Wang et al. (2005) reported the analysis from groups of over 100 mice embryos of the expression of stress-activated protein kinase/Jun kinase (SAPK/JNK) phosphoproteins and p38 mitogen-activated protein kinases (MAPKs) by Western blotting. Another study identified the expression of a major vault protein (MVP) in porcine zygotes cultured in the presence of a specific proteosomal inhibitor using matrix-assisted laser-desorption ionisation-time-of-flight (MALDI-TOF) peptide sequencing and Western blotting (Sutovsky et al. 2005). MVP was also observed to accumulate in poor quality human oocytes and porcine embryos that failed to develop in vitro. In addition, two insulin-responsive glucose transporter isoforms (GLUT4 and GLUT8) and the insulin receptor proteins were confirmed by Western blotting. Another study identified the expression of a major vault protein (MVP) in porcine zygotes cultured in the presence of a specific proteosomal inhibitor using matrix-assisted laser-desorption ionisation-time-of-flight (MALDI-TOF) peptide sequencing and Western blotting (Sutovsky et al. 2005). MVP was also observed to accumulate in poor quality human oocytes and porcine embryos that failed to develop in vitro. In addition, two insulin-responsive glucose transporter isoforms (GLUT4 and GLUT8) and the insulin receptor proteins were confirmed by Western blotting as being present in rabbit blastocysts (Navarrete Santos et al. 2004). However, the only work published describing a proteomics approach in mammalian in vitro fertilisation (IVF) is a study analysing porcine oocyte protein patterns and their variations during in vitro maturation. Up to 600 porcine oocytes were used to separate and visualise proteins of interest by 2D gel electrophoresis. Identification of potential oocyte proteins involved peptide profiling by MALDI-TOF and peptide sequencing by liquid chromatography-tandem mass spectrometry (Ellederova et al. 2004).

In this study, the objective was first to develop a proteomics approach for the generation of protein profiles from small numbers of mammalian embryos, followed by an analysis of protein patterns throughout all stages of preimplantation development to identify proteins that are stage specific. Furthermore, the effects of oxygen concentration (a known regulator of embryonic development) during in vitro development on embryonic cellular function at the protein level were analysed. A comparative analysis to identify potential proteins/biomarkers that are differentially expressed was subsequently undertaken.

**Materials and Methods**

**Mouse embryo culture**

Zygotes were collected from 4-week-old F1 female mice (C57BL/6 × CBA/Ca). Superovulation was induced by an i.p. injection of 5 IU pregnant mare’s serum gonadotrophin (Sigma, St Louis, MO, USA) followed 48 h later by 5 IU human chorionic gonadotrophin (hCG; Pregnyl, Organon Inc., West Orange, NJ, USA). Zygotes were collected at 21 h post-hCG and cultured in groups of 10 in 20 μl medium G1 (Gardner & Lane 2003) at 37°C, 6% CO2 for 48 h, after which they were cultured for a further 24 h to day 4 in medium G2 (Gardner & Lane 2003). The oxygen concentration varied during the experiment – 5% (low) or 20% (high).

**Cation and anion exchange chromatography**

Groups of 5 embryos were extracted in 5 μl lysis buffer (9 M urea/2% CHAPS; Sigma), pulse vortexed and stored at −80°C until further processing. Cationic (CM10) and anionic (Q10) protein chips were washed several times with binding buffers, 0.1 M sodium acetate, pH 4.0, and 50 mM Tris-Cl, pH 9, respectively. The samples were allowed to react with the surface of the protein chip for 30 min at room temperature. The unbound sample was discarded and each spot was washed three times with the same binding buffer for 5-min intervals to remove non-specific binding. The chips were then quickly rinsed with distilled water before air-drying. Sinapinic acid, an energy absorbing molecule (Ciphergen Biosystems, Freemont, CA, USA) was prepared as a saturated solution in 50% acetonitrile/0.5% trifluoroacetic acid and each spot was loaded with 2 × 1 μl of the prepared solution. The protein chips were air dried again and immediately subjected to mass spectrometry.

**SELDI-TOF MS**

Mass spectrometry (MS) was performed using the PCS-4000 Series SELDI-TOF MS (Ciphergen Biosystems). Mass accuracy was calibrated to <0.1% with the all-in-one peptide molecular mass standard for the mass range of <20 kDa (Ciphergen Biosystems). Time-of-flight data were collected from averaged 530 laser shots per spot capturing peptides and proteins <20 000 Da. Spectra were generated by Ciphergen Express Software (Ciphergen Biosystems).
**Bioinformatic analysis**

Protein profiles were analysed for differences in expression with Proteinchip Software Biomarker Edition, version 3.1 (Ciphergen Biosystems). Peaks with a signal-to-noise ratio higher than 6 were first selected, profiles were then normalised to the total ion current and hierarchical clustering was performed to group samples with similar proteins of mass-to-charge ratios (m/z). Significant differences ($P < 0.05$) in peak height between proteins from different groups were calculated using a Mann-Whitney non-parametric test. To test between-chip and between-run variation, one sample was analysed in triplicate and the coefficient of variance of individual peaks in the replicate spectra was calculated by dividing the standard deviation by the mean peak height multiplied by 100%.

**Results**

Development of the optimal procedure for protein profiling of preimplantation embryos involved the empirical analysis of numerous variables including the number of embryos ($n = 5$), sample preparation, lysis optimisation, on-chip sample processing and MS protocol development. Upon optimisation, a single protocol for the entire procedure was followed throughout the study. Protein profiles were generated for all of the embryonic samples using both the anionic (Q10) and cationic (CM10) protein chip types. The spectra generated from the different surfaces were then used to characterise each developmental stage. Twenty-four in vitro samples from day 2, day 3 and day 4 of embryonic development were analysed under both 5% and 20% oxygen conditions. In addition, 24 samples of in vivo-developed day 4 blastocysts were also analysed. The m/z range that was analysed was between 1000 and 20000 Da. To determine the reproducibility of the procedure, one representative sample was processed in triplicate. The mean correlation of variance (CV) of replicates was calculated to be $< 0.15$, a respectable difference in the field of proteomics. The intensity%CV ($\text{intensity s.d./intensity average} \times 100$), a measure of variability of biomarker peak intensity for individual peaks, was calculated for replicates between 2 and 20%.

Figure 1 shows the gel and line plot data of CM10 data enhanced around the 6000-7000 Da range for in vitro-produced (5% oxygen) embryos. There appear to be clusters of proteins only produced after day 2 of

**Figure 1** Protein profiling spectra on the CM10 chip of each developmental day in the m/z range from approximately 6000 to 7000 Da. Data is shown as gel and line plots. Y-axis; relative intensity, O$_2$; Oxygen.
development, as well as other proteins that appear only in the spectra of day 2 embryos. For the Q10 data, a similar situation is observed in Fig. 2 showing gel and line plots of data enhanced around the 8000–12 000 Da range.

After normalisation of the data, univariate statistical analysis was performed to determine the differences between developmental stages of the in vitro-developed (5% oxygen) embryos. Hierarchical clustering was able to correctly segregate all samples according to their developmental stage; the CM10 heat map is shown in Fig. 3. Approximately 40 potential proteins/biomarkers with significant changes were revealed with m/z between 2000 and 19 000 Da. (P < 0.05). The intensity%CV [(intensity s.d./intensity average) × 100] was calculated for samples at > 30, higher than observed between replicates. Peaks with an m/z < 2000 Da are potentially ion noise from the matrix and were therefore excluded. Scatter plots of peak intensities for two of these candidate biomarkers, 12 150 Da (A) and 3460 Da (B) protein are displayed in Fig. 4.

Further analysis of in vitro-produced embryos cultured under different oxygen concentrations also revealed numerous changes to the embryonic proteome. Embryos generated under high oxygen conditions (20%) were distinguished from embryos generated under low oxygen conditions (5%) by 10 proteins/biomarkers of masses between 4000 and 20 000 Da (P < 0.05). Two examples are shown in Fig. 5, an m/z peak of 4835 Da that was down-regulated under high oxygen concentrations (A) and an 18 530 Da m/z peak that was up-regulated under high oxygen concentrations (B). Interestingly, the protein profiles of blastocysts cultured under low oxygen conditions were more comparable to the in vivo-developed blastocysts than blastocysts cultured under high oxygen conditions (Fig. 4).

**Discussion**

It is estimated that a single gene can code for >3 proteins, due to splicing, post-translational modifications (e.g. phosphorylation) and other chemical entities (e.g. carbohydrate...
interactions) that give proteins different properties. Proteins form complex tertiary and quaternary structures and they can be structural or functional, directing both inter- and intracellular interactions. As such, they perform many of the functions that dictate how a cell will react to both its internal and external environment. A single protein may have numerous different biological functional links. Thus, an in depth understanding of the embryonic proteome should lead to a true indication of cellular function and metabolism during mammalian preimplantation development.

The development of a proteomics approach using SELDI-TOF MS in this report has led to the discovery of biomarkers associated with in vivo mammalian embryonic development. These biomarkers represent optimal cellular function and provide a potential diagnostic platform for improving IVF procedure including the further improvement of in vitro culture conditions, stimulation protocols and cryopreservation techniques. In addition, panels of proteins/biomarkers that are specific to each of the individual developmental stages were successfully identified. Due to the multifactorial nature of embryonic development, a proteomic analysis of embryonic development should lead to a true indication of cellular function during mammalian preimplantation development.

Figure 3 CM10 heat map displaying hierarchical clustering of in vitro developed embryos according to their developmental stage. Each column represents an individual sample and each row an individual protein. (red, up-regulated; green, down-regulated).
development, it is possible that a combination of several biomarkers will be necessary to effectively diagnose developmental competence.

The data also revealed that a specific pattern of 10 biomarkers \((P < 0.05)\) effectively discriminated \textit{in vitro} embryos cultured at low oxygen concentrations (5%) from \textit{in vitro} embryos cultured at high oxygen concentrations (20%), and that embryos cultured in the presence of a reduced oxygen concentration exhibited a more \textit{in vivo}-like profile. Oxygen at a concentration of 20% (atmospheric) has been shown to adversely affect embryonic development compared with embryos cultured at 5% (Quinn & Harlow 1978, Batt \textit{et al}. 1991, Gardner & Lane 1996). These data further support the notion that it is not appropriate to culture the preimplantation mammalian embryo in the presence of atmospheric oxygen.

In order to utilise this proteomic approach accurately it is important that experimental procedures and quality control for the SELDI-TOF MS are well calibrated and standardised and that all biomarkers are validated in large studies.

As well as the obvious clinical applications of improving the management of infertile patients and facilitating the clinical application of potentially new diagnostic technologies, this proteomics approach gives researchers an opportunity to expand our knowledge at the cellular level of mammalian preimplantation development including the maternal–embryonic dialogue at the time of implantation. Future experiments will also involve the purification and identification of a selected number of these candidate biomarkers. Identification of these biomarkers will provide mechanistic insight into the biological processes occurring at the cellular level during preimplantation embryonic development. It is also plausible that the identification of these proteins could assist with the detection of viability. From a clinical perspective, quantification of viability potential will result in an increase in IVF pregnancy rates and live births, while reducing the number of embryos transferred.

![Figure 4](image-url) Two examples of differential expression between \textit{in vivo} and \textit{in vitro} embryos at both low and high oxygen concentrations. (A) A 12 150 Da protein \((P < 0.001)\). (B) A 3460 Da protein \((P < 0.01)\). Y-axis; relative intensity. Horizontal line; mean value points represent individual sample.
Figure 5: An m/z peak of 4835 Da was down-regulated (A) while an m/z peak of 18530 Da was up-regulated (B) when in vitro-produced embryos were cultured under high oxygen (O2) concentrations. Y-axis: relative intensity; Horizontal line: mean value points represent individual sample.

Acknowledgements

The authors would like to thank Anthony Eddington from Ciphergen Biosystems for his expert assistance. The authors declare that there is no conflict of interest that would prejudice the impartiality of this scientific work.

References


Gardner DK & Lane M 2003 Towards a single embryo transfer. Reproductive Biomedicine Online 6 470–481.


Received 24 June 2005
First decision 22 July 2005
Revised manuscript received 28 July 2005
Accepted 19 August 2005