Chapter 9

Identification of Biochemical differences between different forms of male infertility by nuclear magnetic resonance (NMR) spectroscopy

Introduction

It is estimated that 10–15% of couples are infertile, and of those, nearly half will involve male factors (Thonneau and Spira, 1991). Despite improvements in the diagnosis and treatment, management of male factor infertility still remains a challenging task. This could be due to the non-availability of a rapid, non-invasive test to assess the quality of semen. Although, today, we have a battery of tests to assess the male fertility potential, the clinical utility of these markers in the prediction of gamete quality and embryo viability is still controversial, often leading to poor success rates when ART is employed (Deepinder et al., 2007).

There has been growing interest in the field of metabolomics which provides useful information about the phenotype an organism through analysis of various biomarkers in body fluids (Lynch et al., 1994; Rohr et al., 1995; Segalen et al., 1995; Tomlins et al., 1998; Schiller et al, 2000; Beckonert et al., 2007). This cost effective technology provides for assessment of metabolic changes within an organism, even when standard clinical chemistry markers are within normal limits (Pauling et al., 1971).

Although different body fluids have been popular targets of proteomic analysis, seminal plasma has not received much attention in this regard. Seminal plasma is the liquid component of semen
and is a mixture of secretions from several male accessory glands, including prostate, seminal vesicles, epididymis, and Cowper's gland. At pH 7.35-7.50, it has buffering properties, protecting spermatozoa from the acidic environment of the vagina. It contains a high concentration of fructose, which is a major nutrient for spermatozoa during their journey in the female reproductive tract. The average protein concentration of human seminal plasma ranges from 35 to 55 g/l. This relatively high abundance of proteins in seminal plasma makes it an ideal candidate for proteomic analysis (Pilch and Mann, 2006).

Human seminal fluid dilutes and transports spermatozoa to the ovum for fertilization and provides a metabolic support to the cells bathed in it. Immediately following ejaculation, a series of reaction occur, initiated by semenogelin (Sg) I and II aggregation to form a gelatinous mass, which is then cleaved by prostate specific antigen (PSA) after 5-20 minutes, facilitating liquefaction (Jonsson et al., 2005; Pampalakis and Sotiropoulou, 2007). Subsequently, the further reactions over the next 6-8 hours results in hyperactivation and capacitation of the spermatozoa (Marieb and Hoehn, 2007).

Disruptions to the biochemical and biophysical reactions following ejaculation are a major cause of infertility in men (Sharma and Agarwal, 1996; Hafez et al., 1997). Since seminal fluid has important roles in spermatozoon survival and overall fertilization success, its impairment can be directly connected to infertility. It is therefore logical to assume that the characterization of different biomarkers in this liquid component could have valuable implications, both in the diagnosis and treatment of male infertility (Pilch and Mann, 2006).
In the present study, an attempt has been made to analyze the seminal plasma of patients with idiopathic / male factor infertility and healthy controls with proven fertility by NMR spectroscopy, with a hope of establishing difference in biomarker profiles, if any between the two groups.

**Materials and Methods**

**Subjects:**
Patients visiting the infertility clinic of Kasturba Medical College and fertile controls, known to have fathered a child within 12 months, participated in the study. Idiopathic infertility included infertile patients, with normozoospermic semen parameters, who had a history of infertility of at least 2 years and normal female partners i.e., normal reproductive history, normal day 2 follicular stimulating hormone (FSH) and luteinizing hormone (LH) levels, normal ovulation (by follicular ultrasound study), and tubal patency (by hysterosalpingogram). Care was taken to ensure that the samples in the control group and experimental groups were age matched. 103 subjects participated in the study, after provision of a written, informed consent. All subjects were asked to provide semen samples after 3-5 days of ejaculatory abstinence. Semen specimens were produced by masturbation directly into a sterile plastic container, in a room specially provided for this purpose and located adjacent to the laboratory. After liquefaction, semen processing and analysis was performed according to the recommendations of the WHO (WHO, 1999). A part of the sample was used for routine andrological examination. Seminal plasma was separated for NMR spectroscopy.
Semen Analysis:

Seminal volume was determined in a graduated tube and sperm concentration was assessed by conventional method using Makler counting chamber (Sefi Medical Instruments, Israel) and expressed in millions/mL. The sperm motility was assessed in at least 100 sperm and expressed as percent of motile sperm (sum of rapid progression plus slow progression sperm). Sperm morphology was assessed by Shorr staining and sperm viability by Eosin-Nigrosin stain. The reference values of semen variables and nomenclature for some semen variables used throughout the thesis has been presented in Chapter 3.

Nuclear Magnetic Resonance (NMR) spectroscopy:

The NMR spectroscopy analysis of seminal plasma has been conducted at the National NMR Facility, Tata Institute of Fundamental Research, Colaba, Mumbai. For NMR analysis, samples were centrifuged at 200 x g for 10 minutes to separate the cells from the rest of the fluid. The fluid component (supernatant) was stored at -80°C for NMR spectroscopy, until shipment. As the analysis was conducted in a different laboratory, located at a distance of approximately 1000 km, the samples were shipped in commercial insulated shipping containers. Maximal care was taken and stringent quality control has been applied while shipping of samples to the specified laboratory. Small chunks of dry ice were first placed at the bottom of the shipping container, and then the sample box was placed in the centre of shipping box, followed by placement of dry ice over and around the box. Immediately after shipment, the samples were shifted to a -80°C deep freezer available with the host facility. Just prior to analysis, the samples were thawed at room temperature for 15 minutes and then diluted 2:1 in D₂O containing 50 mg/ml 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS-D₄), as a chemical shift and concentration reference. 600 µl of the sample was transferred immediately to a 5-mm NMR tube and inserted into the spectrometer.
After allowing for temperature equilibration (to 300 K), the sample was shimmed and data were acquired approximately 15 minutes post-thaw. Data were acquired on a Bruker Ultrashield Plus NMR spectrophotometer operating at 700 MHz $^1$H frequency (Figure 9.1).

**1D NMR:**

1 dimensional $^1$H NMR spectra of the samples were acquired, with a relaxation delay of 4 seconds, 16 transients were collected in 32K data points with a spectral width of 20 ppm. For suppression of the water signal, it was continuously irradiated during the relaxation and mixing time. Acquisition time for each spectrum was ~1.14 minutes. The baseline and phases were corrected by using an automated program kindly made available by Dr. Eberherd Humpfer (Bruker, Biospin).

**Multivariate Data Analysis:**

The NMR spectra were taken through the multivariate analysis. All the spectra were referenced to DSS. The spectra were binned in 0.04 ppm window in software AMIX. The working region of the spectra was 0.5 ppm to 9.5 ppm. After binning, the binned data were normalized with respect to the whole spectrum (to take care of the dilutions effect).

Initially, Principal Component Analysis (PCA) was done to find out outliers. It was followed by Orthogonal Partial Least Square- Discriminant Analysis (OPLS- DA). PCA is an unsupervised method, which means no class entity is assigned to the sample set; hence the model displays any hidden pattern that may be present within the data. Also, PCA is useful to find out the outliers. OPLS-DA is a supervised method; therefore, class specificity is assigned to the sample set. This further refines the class segregation. The OPLS-DA models can be judged by using two parameters- $R^2_X$ and $Q^2$(cum), the former shows the fraction of total variance explained by the
model built and the later shows the extent of separation of the two classes, or in other words, the predictability of the model. The $Q^2$ (cum) is calculated using cross validation methods.

Prior to data analysis, the spectra were binned to a spectral width of 0.04 ppm and normalized to the total intensity of each spectra. This process was achieved using AMIX. The data matrices so generated were used for multivariate data analyses. This was performed in Simca-P+ 12.0 (Umetrics Inc. Sweden). The data matrix was mean centered and pareto scaled and analyzed further. PCA, in principle, is a technique of reduction of dimension that extracts the principal components (PC) from the data matrix, which, in turn, explains certain variation present in the data. Each principal component is a linear combination of the variables (bins, for our case). PC1 explains the maximum variation present in the data; PC2 explains the maximum from the rest of the variation and so on. Hence, each PCs are orthogonal to each other. The data is visualized using scores plot, where each point represents a sample. The significant variables contributing to the pattern shown in the scores plot is extracted from the loadings plot where each point is a spectral variable/bin. The OPLS-DA model generates the predictive component which explains the class specific variation and the orthogonal component which explains variations not related to the class definition. Here also, the data is visualized using the scores and loadings plot as earlier. OPLS-DA, in addition, provides loadings S-plot which helps to identify the class of sample where certain group of variable is increased or decreased, hence aiding in the data interpretation.
Reagents

**Sigma Aldrich, St. Louis, MO, USA**

- Nigrosin Cat No N4754
- Deuterium oxide Cat No 19170-1
- 4,4 dimethyl-4-silapentane-1-sulfonic acid (DSS) Cat No 175617

**Merck & Co, USA**

- Shorr stain Cat No UN1993
- Xylene Cat No 1330-20-7
- Ammonia solution (about 25% pure) Cat No 1336-21-6

**Fisher Scientific, UK**

- Methanol Cat No 32407

**BDH Chemicals, UK**

- Eosin Cat No 3419720

**Sisco Research Laboratories, India**

- DPX Mountant for histology Cat No 42848
Results

The mean semen characteristics of the various subjects are presented (Table 9.1). The averaged representative spectra of the different categories of subjects are summarized in Figure 9.2a. Resonances from a number of low molecular weight metabolites can be identified immediately from previously assigned chemical shift, such as choline, glycerophosphocholine (GPC), citrate, fructose, lactate and resonances from amino acids such as alanine, valine, leucine and isoleucine and is presented in Figure 9.2b (Lynch et al., 1994). In addition, broad bands, characteristic of proteins, have also been observed throughout the entire spectrum, the assignment of which is hindered due to broadening. In view of the abundance of proteins such as semenogelin (Sg), PSA and albumin (Alexandrino et al., 2004; Yoshida et al., 2003), it is possible that few of the peaks observed in the spectra of the present study originate and correspond to resonances from these proteins.

The OPLS-DA analysis revealed significant biochemical differences between the fertile control group and other forms of male infertility – idiopathic infertility, oligozoospermia and azoospermia (Figure 9.3, 9.4 and 9.5, 9.6). Comparison of the control group with either asthenozoospermia or teratozoospermia failed to reveal any significant difference in biomarker profiles (Data not shown). Interestingly, when the idiopathic group was compared against all other forms of infertility, distinct clustering was observed, highlighting distinct differences in biomarker profiles between the groups (Figure 9.6).

The difference in biomarker profiles between the control and the azoospermia group could be due to up regulation of glucose in the former. It is probable that the lack of available
spermatozoa to utilize energy could lead to observation of such a biochemical profile. The same effect is also observed with respect to the control group and oligozoospermia, with up-regulation of an additional biochemical component, citrate. At present, however, it remains to be understood if this probable up regulation of fructose and citrate in oligozoospermic individuals arises from a defect in a specific metabolic pathway. The difference in biomarker profiles between the idiopathic infertility and the rest of the groups combined could originate from either the up-regulation or down regulation of a several compounds, including lysine, arginine, tyrosine, citrate, proline and fructose. However, precise biochemical identification of compounds responsible for the observed differences in biochemical profiles cannot be arrived at until further 2D-NMR experiments are carried out.
Figure 9.1: 700 MHz Bruker spectrometer employed for nuclear magnetic resonance spectroscopy of seminal plasma at National NMR Facility, Tata Institute Of Fundamental Research, Mumbai
Figure 9.2a: Averaged spectra of the fertile and infertile group

Averaged spectra of class 2a (Idiopathic infertility)  Averaged spectra of class 2b (Oligozoospermia)  Averaged spectra of class 2c (Azoospermia)

Averaged spectra of class 2g (Teratozoospermia)  Averaged spectra of class 2e (Azoospermia)  Averaged spectra of class C (Fertile Control)
Figure 9.2b: Typical ¹H NMR spectra of human seminal plasma from fertile group (control)

Key: Val, valine; Ile, isoleucine; Leu, leucine; Lac, lactate; Lys, lysine; Ala, alanine; Arg, arginine; Spa, spermine; Glu, glutamate; Gin, glutamine; Cit, citrate; GPC, Glycylglycine; Cho, choline; Gyl, glycine; Tyr, tyrosine; Uni, uridine; His, histidine; Pha, phenylalanine; Glc, Glucose; Fru, Fructose
Figure 9.3: Biochemical difference between control and Idiopathic Infertility by $^1$H NMR
1: Idiopathic infertility; 2: Control group (fertile)
Figure 9.4: Biochemical difference between control and oligozoospermia by \(^1\text{H}\) NMR
1: Control Group (fertile); 2: Oligozoospermia
Figure 9.5: Biochemical difference between control and azoospermia by $^1$H NMR
1: Azoospermia; 2: Control group (fertile)
Figure 9.6: Biochemical difference between idiopathic infertility & other forms of infertility by $^1$H NMR
1: Oligozoospermia, Aasthenozoospermia, Teratozoospermia, Azoospermia; 2: Idiopathic infertility
Table 9.1: Semen characteristics of fertile controls and infertile subjects who underwent NMR spectroscopic analysis

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Age (Yrs)</th>
<th>Seminal volume (ml)</th>
<th>Sperm count (Millions/ml)</th>
<th>Total sperm motility (%)</th>
<th>Sperm with normal morphology (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>6</td>
<td>36.17 ± 1.60</td>
<td>2.67 ± 0.40</td>
<td>64.50 ± 9.94</td>
<td>71.50 ± 5.74</td>
<td>35.5 ± 3.01</td>
</tr>
<tr>
<td>Idiopathic infertility</td>
<td>17</td>
<td>36.24 ± 0.99</td>
<td>2.12 ± 0.18</td>
<td>60.41 ± 5.37</td>
<td>66 ± 1.97</td>
<td>35.41 ± 1.77</td>
</tr>
<tr>
<td>Oligozoospermia</td>
<td>20</td>
<td>34.53 ± 1.18</td>
<td>3.82 ± 0.66</td>
<td>9.64 ± 0.99</td>
<td>41.50 ± 4.26</td>
<td>15.60 ± 1.85</td>
</tr>
<tr>
<td>Asthenozoospermia</td>
<td>20</td>
<td>37.20 ± 0.97</td>
<td>3.00 ± 0.27</td>
<td>45.85 ± 4.99</td>
<td>42.95 ± 3.29</td>
<td>20.25 ± 1.84</td>
</tr>
<tr>
<td>Teratozoospermia</td>
<td>20</td>
<td>38.15 ± 1.41</td>
<td>3.80 ± 0.32</td>
<td>55.90 ± 0.544</td>
<td>53.3 ± 4.16</td>
<td>17 ± 0.88</td>
</tr>
<tr>
<td>Azoospermia</td>
<td>20</td>
<td>32.85 ± 0.93</td>
<td>2.78 ± 1.30</td>
<td>NA*</td>
<td>NA*</td>
<td>NA*</td>
</tr>
</tbody>
</table>

NA* - Not applicable