Investigation of Spermatozoa and Seminal Plasma by Fourier Transform Infrared Spectroscopy

O. BARČOT, M. BALARIN, O. GAMULIN,* D. JEŽEK, P. ROMAC, and J. BRNJAS-KRALJEVIC

University of Zagreb, Medical School, Šalata 3, Zagreb, 10000, Croatia (O.B.); University of Zagreb, Medical School, Dept. of Physics and Biophysics, Šalata 3, Zagreb, 10000, Croatia (M.B., O.G., J.B.-K.); University of Zagreb, Medical School, Dept. of Histology and Embryology, Šalata 3, Zagreb, 10000, Croatia (D.J.); and Vuk Vrhovac University Clinic for Diabetes, Endocrinology and Metabolic Diseases, University of Zagreb, Medical School, Dugi dol 4a, Zagreb, 10000, Croatia (P.R.)

Fourier transform infrared (FT-IR) spectra of human spermatozoa and seminal plasma were recorded and analyzed. The procedure that was established for sample preparation enabled acquisition of reproducible spectra. The parameter $I_{1600}/I_{600}$ for controlling spectra reproducibility was defined. The assignment of bands was carried out using an empirical approach and the origin of the “sperm specific doublet”, the bands at 968 cm$^{-1}$ and 981 cm$^{-1}$, was determined. The principal component regression (PCR) algorithm was used to define the specific spectral regions correlating to characteristics of spermatozoa, such as concentration, straight-line velocity (VSL), and beat cross frequency (BCF). Then, simple spectral parameters, such as band intensities and band ratios, were tested to determine which one best correlates to characteristics of spermatozoa. The region of the amide I band, between 1700 cm$^{-1}$ and 1590 cm$^{-1}$, was defined as a specific spectral region that correlates to the concentration of spermatozoa. The parameter that gave the linear dependence to the concentration of spermatozoa was the intensity of the amide I band. For VSL, the bands between 1119 cm$^{-1}$ and 943 cm$^{-1}$ were defined as the specific spectral region. The relative amount of nucleic acids with respect to proteins showed linear dependence on the straight-line velocity of spermatozoa. BCF showed the best correlation to the bands between 3678 cm$^{-1}$ and 2749 cm$^{-1}$, which largely represent lipids and proteins. These results suggest that FT-IR spectroscopy can serve as an adjunct to conventional histopathology studies.

Index Headings: Fourier transform infrared spectroscopy; FT-IR spectroscopy; Human semen; Spermatozoa; Seminal plasma; Principal component regression; PCR.

INTRODUCTION

Semen is an organic fluid that consists of two parts, the cellular part (spermatozoa and in some instances leucocytes and epithelial cells) and the noncellular part (seminal plasma). Spermatozoa are essential for human reproduction and seminal plasma provides a nutritive and protective medium to spermatozoa during fertilization. Male fertility is realized as an important contributor to the conception potential of a couple, and hence it has become a subject of intense research. Routine semen analyses include a close examination by light microscopy and computer assisted semen analysis (CASA) of the conventional spermatozoa characteristics such as morphology, motility, forward progression, and concentration.

Some other studies include ultrastructural analysis of spermatozoa by transmission electron microscopy (TEM), the use of functional assays to distinguish between fertile and infertile men, the use of flow cytometry studies of semen samples, etc. There is still a lack of data regarding molecular aspects of male infertility that can be employed as a diagnostic test.

We think that Fourier transform infrared (FT-IR) spectroscopy might be the method of choice since it has become an accepted tool for the characterization of complex building blocks of biological systems such as proteins, nucleic acids, lipids, and carbohydrates. It also provides an insight into chemical composition of major constituents of cells and tissues. Based on the premise that every pathological process is caused by the chemical change that precedes the morphological or symptomatic manifestation, FT-IR spectroscopy has proven to have a diagnostic potential. Several research groups reported its use for identification, differentiation, and classification of microorganisms. Differentiation of normal from cancerous cervical tissue, cirrhotic and cancerous liver tissue, normal and cancerous colon tissue, normal and malignant lymph cells and tissue, and investigation of brain cancer were reported as well. Only a few studies have been reported to have used FT-IR spectroscopy for studying the molecular characteristics of spermatozoa structure both for humans and animals.

Spermatozoa have a highly specialized structure. Their DNA is densely packed in the nucleus of the head surrounded by the acrosome full of enzymes, and the tail is packed with mitochondria and microtubules that serve for movement. Based on the structure, spermatozoa seem to be suitable for infrared (IR) spectroscopy studies. These studies may provide additional insight into the molecular basis of conception problems, and they serve as an adjunct to conventional histopathology studies.

In this paper we present the results of the FT-IR study of human spermatozoa and seminal plasma. We have first defined the proper sample preparation procedure and then analyzed the IR spectra of spermatozoa and seminal plasma separately. In addition, we have established the relationship between characteristic spectral features and spermatozoa concentration and motility, since these two properties are the best predictors of fertility potential. To our knowledge, this is the first time that the conventional human spermatozoa properties have been correlated to the characteristic spectral features in the IR spectra of spermatozoa.

MATERIALS AND METHODS

The subjects included in this study were male partners of couples having problems with spontaneous conception who presented to the Vuk Vrhovac University Clinic for Diabetes, Endocrinology and Metabolic Diseases between October 2003 and February 2005 for semen analysis. Each man produced a single sample by usual male fertility test routine after four days of advised abstinence. Semen samples were incubated at 37 °C
for 20 min to achieve liquefaction and then analyzed by computer-aided semen analysis (CASA) using the Hamilton Thorne IVOS 10 Analyzer (Hamilton Thorne Research, Beverly, MA). In a CASA system, a series of videotaped microscopic images are digitalized, and computer software programs are used to evaluate spermatozoa concentration, motility, and morphology. 5,39 This first semen analysis provided only motility data. Semen samples were then diluted in physiological solution in a 1:1 volume ratio and centrifuged at 1000 g for 20 minutes. This centrifugation enabled us to separate spermatozoa from seminal plasma. Spermatozoa occupied the bottom part of the cuvette, and seminal plasma the upper part. Seminal plasma was then extracted and placed into another cuvette. The residue was re-diluted in physiological solution in a 1:10 volume ratio to purify the spermatozoa samples. Both cuvettes, the one with seminal plasma and the one with spermatozoa, were centrifuged again at 1000 g for 20 minutes. The supernatant over spermatozoa specimen was discarded. The residue containing only spermatozoa was analyzed by CASA to obtain concentration data. From this stock, 20 μL was taken for spermatozoa FT-IR measurements. The supernatant in the cuvette with seminal plasma was partially removed to another cuvette. It was analyzed by CASA to verify that no residual spermatozoa were left and to define the concentration of other cells in it, such as leucocytes, epithelial cells, or erythrocytes. From this stock, 20 μL was taken for seminal plasma FT-IR measurements.

The infrared spectra were recorded with a Perkin-Elmer GX spectrometer, equipped with a DTGS detector, in absorption mode. For each measurement, 20 μL of sample was spread evenly onto an optical grade silicon wafer (resistivity was 250 Ωcm) to fill a bordered circular area of 10 mm in diameter. Before recording, the samples were dried in vacuum (150 Pa) for 10 minutes. A total of 200 scans were coadded at a resolution of 4 cm⁻¹. These macroscopically acquired spectra represent an average of the spectral features of all cells in the infrared beam.13 However, since the spermatozoa samples were carefully purified, they are presumed to be highly homogenous in the sense that no other cells are present in the sample. Hence, the acquired spectra can be considered spermatozoa representative.

Prior to the analysis, the spectra were baseline corrected using Spectrum v.5.0.1. software supplied with the instrument. The integrated areas were calculated using the procedure built in the same software after the starting and the end baseline points were defined for all analyzed bands.

Infrared methods for the analysis of spermatozoa concentration and motility parameters were calibrated using a principal component regression (PCR) algorithm incorporated in Spectrum Quant+ software provided with the instrument. The PCR algorithm enabled us to determine the correlation between conventional properties of spermatozoa and recorded spectra. The program was filled with the recorded IR spectra and the appropriate property data, either concentration or VSL or BCF. In the first principal component analysis (PCA) stage, the PCR algorithm seeks to account for variation in the spectral data and express it in as few terms as possible. Then, in the next multiple linear regression stage, it correlates these variations with the property data. Analyzing different spectral regions with the PCR algorithm, we determined the intervals of recorded spectra that correlate the best with the property data. Then, from that specific spectral region a simple spectral parameter such as band intensity or band ratio was analyzed to determine which one best correlates with the conventional spermatozoa properties.

RESULTS AND DISCUSSION

Spectra Analysis and Band Assignment. As a first step in the spectral analysis, recorded spectra were visually filtered to remove those that were obviously inadequate for the band assignment. These included, for example, spectra from samples that had insufficient cellular material in the field of view and the one that showed gross contamination. Furthermore, since all biological samples are very complex, and numerous factors can influence the spectra, it was important to define a parameter that will control proper sample preparation. Although all spermatozoa samples were carefully purified, several experiments, not shown here, showed spectral contamination, most likely from seminal plasma and other cells such as leucocytes or epithelial cells. Because the most prominent deviations in the spectra appeared in the region around the band at 1087 cm⁻¹, we calculated the ratio of the integrated areas of the bands around 1087 cm⁻¹ (I₁₀₈₇, straight line baseline between points 1150 cm⁻¹ and 987 cm⁻¹) and the band at 966 cm⁻¹ (I₉₆₆, straight line baseline between points 987 cm⁻¹ and 940 cm⁻¹). Since these bands are the markers for nucleic acids,19,40 we expected the ratio I₁₀₈₇/I₉₆₆ to be constant. Linear regression, shown in Fig. 1, resulted in I₁₀₈₇/I₉₆₆ = 10.7 ± 1.4 with a regression coefficient β = 0.9962. We have postulated that the spermatozoa samples that have a constant ratio I₁₀₈₇/I₉₆₆ in the spectra are properly prepared and can be considered reproducible. The samples, whose spectra were excluded, based on this criterion, most probably had residual epithelial cells, erythrocytes, or leucocytes, which altered nucleic acids vibrational bands.

Moreover, since spermatozoa have a DNA very different from other human cells, we have presumed that the value of this ratio might be spermatozoa specific. Comparison of spectra, available in the literature, of various other cells, such as blood components26 and cervical cells,24 indicated that the value of the I₁₀₈₇/I₉₆₆ ratio differs not only from the one for
The band at 1547 cm\(^{-1}\) is assigned to the amide II absorption of cell proteins (mostly the C–O stretching vibrations of the peptide backbone) with contributions from C–O stretching modes of nucleic acids, and the band around 1547 cm\(^{-1}\) assigned to the amide II absorption of cell proteins (N–H bending coupled to C–N stretching mode).\(^8,19\) The bands around 1450 cm\(^{-1}\) and 1400 cm\(^{-1}\) arise mainly from asymmetric and symmetric CH\(_2\) bending modes of methyl groups of proteins, respectively,\(^24,30\) with some contribution of the vibrations of fatty acids.\(^31\) The band at 1236 cm\(^{-1}\) and the broad multicomponent band around 1087 cm\(^{-1}\) are generally assigned to the asymmetric phosphate (\(\nu_C\)-PO\(_2^\ast\)) and symmetric phosphate (\(\nu_S\)-PO\(_2\)) stretching mode of nucleic acids, respectively.\(^29\) The contribution to these bands by the phosphate residues present in membrane lipids is negligible. It is because the peak intensity ratio between the \(\nu_C\)-O band (around 1740 cm\(^{-1}\)) and the \(\nu_C\)-PO\(_2^\ast\) band (1236 cm\(^{-1}\)) of phospholipids is greater than 1,\(^24,29\) and in the spectra of spermatozoa is much smaller than 1 (\(\nu_C\)-O is only a weak shoulder). This indicates that phospholipids do not contribute significantly to the intensity of the \(\nu_C\)-PO\(_2^\ast\) band in the infrared spectra of spermatozoa. Although the band at 1236 cm\(^{-1}\) could theoretically also have a contribution from the amide III band of proteins, we think that it is not the case here for the reason indicated by Wong et al.,\(^24\) namely, protein conformation in human spermatozoa is predominantly \(\alpha\)-helix (amide I peak around 1655 cm\(^{-1}\)).\(^33\) In that case, the location of the amide III band should be at 1260 cm\(^{-1}\) to 1290 cm\(^{-1}\) and therefore does not contribute to the intensity of the PO\(_2\) band at 1236 cm\(^{-1}\).\(^24\)

The band at around 966 cm\(^{-1}\) is assigned to C–O and C–C stretching vibrations of nucleic acids.\(^19\)

A spectrum of seminal plasma is shown in Fig. 2B. The region of importance for our investigation was between 1300 cm\(^{-1}\) and 900 cm\(^{-1}\). As far as chemical content is concerned, seminal plasma mostly comprises proteins, free amino acids, ribonucleic acid, fructose, different ions, etc.\(^1,2\) In the region of our interest, proteins have either weak or no absorption, while vibrational bands of nucleic acids are present.\(^40\) Comparing this spectrum to the spectrum of spermatozoa, it is obvious that the intensity of the band at 1236 cm\(^{-1}\) is reduced and that the band at 1087 cm\(^{-1}\) is broadened and slightly red shifted to 1081 cm\(^{-1}\). The band at 966 cm\(^{-1}\) is missing and the one at 980 cm\(^{-1}\) appeared. We suspect that these changes are the consequence of very low content of DNA and the presence of RNA in seminal plasma, contrary to spermatozoa, which contain predominantly DNA.\(^2\) RNA is also characterized by the bands\(^40\) at 1157 cm\(^{-1}\), 1117 cm\(^{-1}\), and 1036 cm\(^{-1}\), which together with polysaccharide bands\(^22,26\) at 1024 cm\(^{-1}\) and 1050 cm\(^{-1}\) might be the cause of the red shift and broadening of the band at 1087 cm\(^{-1}\). Wood et al.\(^26\) have recorded the spectra of semen in order to assess the possible contamination of cervical samples. They have defined a “distinctive doublet at 981/968 cm\(^{-1}\)” as a marker for semen identification, which they left unassigned. We believe that the band at 966 cm\(^{-1}\) originates from spermatozoa, probably its DNA.\(^40\) The band at 980 cm\(^{-1}\) originates from the seminal plasma and is as yet unassigned. It might be due to vibrations of RNA superimposed on the vibrations of some free DNA or DNA of other cells present in seminal plasma. These statements have yet to be proved.

In order to estimate the relative amount of nucleic acids with respect to proteins, we used parameters defined by Benedetti
et al.\(^4^0\). They have defined the A1/A2 ratio as a useful parameter. A1 is an integrated area of nucleic acids\(^1^5\) bands with baseline points from 1150 cm\(^{-1}\) to 940 cm\(^{-1}\). A2 is an integrated area of the amide II band with baseline points from 1590 cm\(^{-1}\) to 1480 cm\(^{-1}\). The average value of this ratio for spermatozoa is 1.20 ± 0.36, calculated from 16 spectra previously defined as reproducible. This value suggests an accessed amount of nucleic acids with respect to proteins. Furthermore, in the same study, Benedetti et al.\(^4^0\) have shown that the ratio of optical densities of amide II and amide I bands (DAII/DAI) also changes with the change of nucleic acid/protein ratio. The average value of this ratio for spermatozoa is 0.34 ± 0.05, suggesting that the only nucleic acid present in spermatozoa is DNA. These findings are in accordance with the known biochemical composition of spermatozoa.\(^2\)

**Spermatozoa Parameters versus Spectral Features:**

**Spermatozoa Concentration.** A set of PCR trials was carried out using 16 spermatozoa spectra to determine the specific spectral region correlating to spermatozoa concentration obtained by CASA. These trials revealed that the optimal spectral region for concentration is between 1700 cm\(^{-1}\) and 1590 cm\(^{-1}\). This is the region of the amide I band. Figure 3A displays a scatterplot comparing IR-predicted to reference values of spermatozoa concentration. The best fitting regression line \(Y = Ax + B\) has \(A = 0.998\) and \(B = 0.109\), with a correlation coefficient of \(\beta = 0.991\). Based on this finding we have tested whether the optical density of the amide I band (DAI) can be the representative parameter for concentration and plotted DAI versus concentration values. From Fig. 3B it is obvious that the correlation between these two parameters is linear, with the significant correlation coefficient \(\beta = 0.947\). It is our opinion that, if the samples are properly prepared by using reproducibility criteria above, FT-IR spectroscopy has a potential as a diagnostic tool for spermatozoa counting.

**Motility Parameters.** The CASA analysis outcomes included several different motility parameters.

As a measure of progression we have chosen the straight-line velocity (VSL) parameter because it is a direct measurement.\(^3^9\) Performed PCR trials revealed that the optimal spectral region for VSL is between 1119 cm\(^{-1}\) and 943 cm\(^{-1}\). The correlation coefficient \(\beta\) was improved if the region from 1040 cm\(^{-1}\) to 989 cm\(^{-1}\) is excluded from the optimal spectral region. Hence, PCR revealed that the optimal spectral region for VSL is the region of nucleic acids bands. Figure 4A displays a scatterplot comparing IR-predicted to reference values of spermatozoa VSL. The best fitting regression line \(Y = Ax + B\) has \(A = 1\) and \(B = -1.58 \times 10\(^{-5}\)\), with a correlation coefficient of \(\beta = 0.911\). We also tested whether the relative amount of nucleic acids with respect to proteins (A1/A2) can be the representative parameter for VSL, and plotted A1/A2 versus VSL values, as shown in Fig. 4B. It is obvious that the correlation is linear, with the significant correlation coefficient \(\beta = 0.878\). It seems that well-packed DNA with a low protein content in spermatozoa’s head enables better motility.

As an indicator of spermatozoa vigor we have chosen the beat cross frequency (BCF) parameter, which represents the number of lateral oscillatory movements of the spermatozoa head around the mean path expressed in Hertz (Hz). PCR performed on twelve spermatozoa spectra revealed that the optimal spectral region, with the best correlation to BCF, is between 3678 cm\(^{-1}\) and 2749 cm\(^{-1}\). This is the region of OH, \(\text{NH}_2\), \(\text{NH}_3\), \(\text{CH}_2\), and \(\text{CH}_3\) vibrations mostly representing proteins and lipids present in spermatozoa tail.\(^3^0\)\(^3^3\) Since the BCF is well correlated with the tail beat frequency,\(^4^4\) we think that it is plausible to choose this region of the spectra to estimate the BCF. Figure 5A displays the scatterplot comparing IR-predicted to reference values of spermatozoa beat cross frequency. The best fitting regression line \(Y = Ax + B\) has \(A = 0.999\) and \(B = 0.00836\), with a correlation coefficient of \(\beta = 0.962\). Figure 5B shows the difference in spectra of spermatozoa with the highest (max) and the lowest (min) BCF. Although PCR gave a good correlation between reference and IR predicted values, the spectra in Fig. 5B indicate that the parameters of spectra in a specific region differ slightly between max and min BCF. Consequently, it is not easy to establish a straightforward correlation between them and the BCF. Further studies of correlations between spectral parameters are needed in order to find the simple parameter that will best represent the changes of BCF.

**CONCLUSION**

To the best of our knowledge, this is the first attempt to correlate the characteristic spectral features in the IR spectra of...
spermatozoa to the conventional human spermatozoa properties. We have defined the procedure for sample preparation, which enabled reproducible spectra acquisition. The criterion for controlling reproducibility was also defined. Our suggestion is that this criterion may also be used for cell differentiation. The analysis of infrared spectra of spermatozoa and seminal plasma, and the comparison to the previously assigned vibrational bands, led us to the conclusion that the band at 966 cm\(^{-1}\) originates from the spermatozoa, probably its DNA, and that the band at 980 cm\(^{-1}\) originates from the seminal plasma. At present, this band is unassigned. From the spectra of spermatozoa we have estimated the relative amount of nucleic acids with respect to proteins. Calculated values are in the agreement with known chemical composition of spermatozoa. PCR revealed the regions in the spectra specific for conventional spermatozoa properties. Simple parameters, such as band intensity and band ratio, which represent spermatozoa concentration and motility parameters, were defined. This indicates the potential for the utilization of vibrational spectroscopy in the diagnosis of human reproduction dysfunction.

ACKNOWLEDGMENTS

This work was supported by the collaborative project “Biomedical Investigation of Reproduction and Development” and the project No.0108043 of the Ministry of Science, Education and Sports of the Republic of Croatia and was approved by the Ethical Committee of the University of Zagreb Medical School.

35. A. Arav, M. Pearl, and Y. Zeron, Cryo. Lett. 21, 179 (2000).

Fig. 5. (A) The comparison of IR-predicted BCF values to the reference BCF values. (B) The difference in spectra between the highest (max) and lowest (min) value of BCF. Letters correspond to following vibrational bands: (a) stretching vibrations of OH and/or NH groups centered around 3300 cm\(^{-1}\); (b) asymmetric CH\(_2\) stretching vibrations at 2960 cm\(^{-1}\); (c) asymmetric CH\(_3\) stretching vibrations at 2925 cm\(^{-1}\); (d) symmetric CH\(_2\) stretching vibrations at 2875 cm\(^{-1}\); (e) symmetric CH\(_3\) stretching vibration at 2854 cm\(^{-1}\).