Fiber evanescent wave spectroscopy using the mid-infrared provides useful fingerprints for metabolic profiling in humans

Marie-Laure Anne[†]
UMR-CNRS 6226
Campus de Beaulieu/Equipe Verres et Céramiques
University of Rennes 1
Sciences chimiques de Rennes
Rennes, 35042, France

Caroline Le Lan[†]
University Hospital Pontchaillou
Liver Disease Unit
    and
UMR-INSERM U991
CHRU Pontchaillou
Rennes, 35033, France

Valérie Monbet[†]
Applied Statistics Laboratory
Bretagne Sud University
Vannes, 56000, France

Catherine Boussard-Plédel
UMR-CNRS 6226
Campus de Beaulieu/Equipe Verres et Céramiques
University of Rennes 1
Sciences chimiques de Rennes
Rennes, 35042, France

Martine Ropert
Department of Biochemistry
    and
UMR-INSERM U991
University Hospital Pontchaillou
Rennes, 35033, France

Olivier Sire
Laboratoire d'Ingénérie des Matériaux de Bretagne
Université Européenne de Bretagne Sud
Vannes, 53000, France

Michel Pouchard
Christine Jard
Health Appraisal Center
Caisse Primaire d'Assurance Maladie
Rennes, 35000, France

Jacques Lucas
Jean Luc Adam
UMR-CNRS 6226
Campus de Beaulieu/Equipe Verres et Céramiques
University of Rennes 1
Sciences chimiques de Rennes
Rennes, 35042, France

Pierre Brissot
University Hospital Pontchaillou
Liver Disease Unit
    and
UMR-INSERM U991
CHRU Pontchaillou
Rennes, 35033, France

Bruno Bureau*
UMR-CNRS 6226
Campus de Beaulieu/Equipe Verres et Céramiques
University of Rennes 1
Sciences chimiques de Rennes
Rennes, 35042, France

Olivier Loréal*
UMR-INSERM U991
University of Rennes 1
IFR 140
CHRU Pontchaillou
Rennes, 35033, France

Abstract. Fiber evanescent wave spectroscopy (FEWS) explores the mid-infrared domain, providing information on functional chemical groups represented in the sample. Our goal is to evaluate whether spectral fingerprints obtained by FEWS might orientate clinical diagnosis. Serum samples from normal volunteers and from four groups of patients with metabolic abnormalities are analyzed by FEWS. These groups consist of iron overloaded genetic hemochromatosis (GH), iron depleted GH, cirrhosis, and dysmetabolic hepatosiderosis (DYSH). A partial least squares (PLS) logistic method is used in a training group to create a classification algorithm, thereafter applied to a test group. Patients with cirrhosis or DYSH, two groups exhibiting important metabolic disturbances, are clearly discriminated from control groups with AUROC values of 0.94±0.05 and 0.90±0.06, and sensibility/specificity of 86/84% and 87/87%, respectively. When pooling all groups, the PLS method contributes to discriminate controls, cirrhotic, and dysmetabolic patients. Our data demonstrate that metabolic profiling using infrared FEWS is a possible way to investigate metabolic alterations in patients. © *2009 Society of Photo-Optical Instrumentation Engineers.* [DOI: 10.1117/1.3253319]

[†]Authors contributed equally to this work.
*Address all correspondence to: Olivier Loréal; Doctor, INSERM U522, University Hospital Pontchaillou, 35033 Rennes, France. Tel. 33.2.99.54.37.37; Fax: 33.2.99.54.01.37; E-mail: olivier.loreal@univ-rennes1.fr; and Bruno Bureau, Professor, Equipe Verres et Céramiques, UMR-CNRS 6226 Sciences chimiques de Rennes, Université de Rennes 1, Campus de Beaulieu, 35042, Rennes, France. Tel: 33.2.23.23.65.73; E-mail: bruno.bureau@univ-rennes1.fr

1 Introduction

Numerous diseases are characterized by metabolic abnormalities that can be detected at plasmatic or serum levels through

the quantification of biochemical components such as glucose, cholesterol, triglyceride, globulins, albumin, transferrin, enzymes, and ions. This explains the development of numerous biochemical methods aimed at quantifying, and more rarely at characterizing, specific serum components. However, most of the methods developed to date yield specific information on a particular component, whereas the determination of one component is frequently not sufficient, physicians needing a combination of biochemical parameters, in addition to clinical and imaging findings, to base their diagnosis.[1,2] For example, the diagnosis of a metabolic syndrome is based on the evaluation of a combination of positive criteria, including glucose, cholesterol, triglyceride, uricemia measurements, and a large number of other criteria to exclude other diseases or associated complications.[3] Therefore, methodologies that provide an extensive view of the serum biochemical species levels from a unique and limited serum sample would certainly help in establishing an early diagnosis in addition to classical clinical, biological and imaging findings.[4,5] Ideally, the goal is to identify serum fingerprints useful for both diagnosis and prognosis.[6]

The mid-infrared infrared (MIR) domain contains most biological molecule spectral signatures through their functional groups.[7,8] This explains the development of Fourier transform infrared spectroscopy that enables the investigator to have true fingerprints of the fundamental functional chemical groups represented in biological samples.[7–15] The recent development of chalcogenide glass optical fiber by our group is an essential step to remotely collecting MIR signals. Their unique and original tapered shape is also decisive to recording informative spectra from biological complex samples.[16] The principle of the method is based on the general concept of fiber evanescent wave spectroscopy (FEWS), and remote spectroscopy using IR glass fibers gives the opportunity to have access to real-time spectra.

Our hypothesis was that FEWS using IR glass fibers, combined with adequate statistical analysis, could give spectral fingerprints useful to perform a rapid classification of patients with diseases having metabolic consequences, and thus to obtain relevant information for the diagnosis.

To demonstrate the concept, we chose to study sera obtained from patients exhibiting chronic liver diseases known to induce various alterations of a number of parameters related to both the etiology and the severity of the disease. We show that the implement of our highly sensitive optical fibers, such as an MIR remote sensor associated with modern statistical methods of spectral data could be a way to investigate metabolic alterations in patients.

2 Patients, Materials and Methods

2.1 Patients

The control group was constituted by 18 to 70 years old healthy volunteers ($n=140$). They were all included in the study during a systematic medical examination in a health care center (Centre d'Examens de Santé, Caisse Primaire d'Assurance Maladie de Rennes). They were all of Caucasian origin, devoid of known acute and/or chronic diseases as well as of excessive alcoholic consumption, and with a normal body mass index less than 25 kg/m^2.

The first disease group comprised 29 patients exhibiting systemic iron overload related to untreated genetic hemochromatosis (GH). The diagnosis was assessed by the presence of a homozygous C282Y mutation of the HFE gene. In addition, all the patients presented a phenotypic expression characterized by a serum transferrin saturation level above 50%, and a liver iron concentration estimated by magnetic resonance imaging[17] (MRI) or biochemical determination over 36 μmol/g of dry weight liver. The blood samples were collected before the beginning of the iron depletive treatment.

The second group of patients included 37 iron-depleted HFE hemochromatotic patients after the completion of intensive phlebotomies of the induction phase (DGH). They presented serum transferrin saturation levels below 45% and ferritinemia levels below 100 μg/L.

The third patient group was composed of 58 alcoholic cirrhotic patients (AC). The diagnosis was assessed on the clinical history of excessive alcohol consumption and the absence of other recognized liver disease etiology. All the patients were grade B or C with the Child-Pugh classification.[18]

The fourth group included 67 patients with dysmetabolic hepatosiderosis (DYSH). They presented a dysmetabolic liver siderosis not linked to C282Y HFE homozygozity and characterized by increased of serum ferritin, unexplained increase of hepatic iron concentration ($>$36-μmol/g dry weight liver) as determined by MRI examination, and an insulin resistance defined by one or more of the following findings: body mass index $>$27, abnormal glucose metabolism, dyslipidemia, or increased blood pressure.

The study was approved by the local ethic committee (CCPPRB Comité Consultatif de Protection des Personnes se prêtant à la Recherche Biomédicale).

2.2 Laboratory Procedures

All biological tests mentioned in Table 1 were performed in the biochemistry and haematology laboratories of the Pontchaillou Hospital in Rennes according to currently in force clinical guidelines. Peripheral venous blood was collected using plain tubes (Vacutainer Tube, Becton Dickinson, Grenoble, France) after an overnight fast, and all parameters were analysed on fresh samples. For infrared spectroscopic analysis, serum samples were stored at -80 °C in vials from VWR (West Chester, Pennsylvania) (low temperature freezer polypropylene, silicone O-ring seal, sterile) until the analysis, which was performed twice.

2.3 Infrared Spectroscopy

Principles and materials used for this study were previously described.[16,19] Briefly, an analysis of serum was performed using a tapered optical fiber made with Te$_2$As$_3$Se$_5$ glass (TAS) with large optical transparency in the MIR range from 850 to 4000 cm^{-1}. The diameter of the fiber, equal to 400 μm, was reduced in the sensing area. The experimental setup consisted of a Fourier transformed infrared spectrometer (Bruker V22, Bruker Corporation, Billerica, Massachusetts) coupled with the fiber, and a cooled mercury-cadmium-telluride detector. A special kit provided by Bruker was connected to the spectrometer to focus the infrared beam from the black body to the input of the fiber. When chemical species are brought into contact with the fiber, the IR optical rays are

Table 1 Bioclinical features of control and patient groups. Results are expressed as mean±SD. N=normal values; GH=untreated hereditary hemochromatosis; DGH=treated hereditary hemochromatosis; DYSH=dysmetabolic hepatosiderosis; AC=alcoholic cirrhosis. sTfR=soluble transferrin receptor; ALT=alanine amino transferase; AST=aspartate amino transferase; US CRP=ultrasensitive C reactive protein; p=comparison with the control group; and NS=non significant.

	CONTROL N=140	GH N=29	p^A	DGH N=37	p^A	DYSH N=67	p^A	AC N=58	p^A
Age (years)	46±12	46±11	NS	49±10	NS	55±11.1	<0.001	55±9	<0.001
Sex-ratio M/W	1.69	2	NS	2.81	NS	7	<0.001	2	NS
Albumin (g/l) (N: 35–50)	47.2±4.4	41.3±6.4	<0.001	41.6±3.7	<0.001	43.4±3.4	<0.001	30.1±6	<0.001
Transferrin (g/l) (N: 1.7–3.5)	2.4±0.4	1.9±0.2	<0.001	2.2±0.3	0.037	2.4±0.4	NS	1.9±0.8	<0.001
Alpha 1 globulines (N:0.8–2.6)	1.4±0.3	1.3±0.4	NS	1.4±0.4	NS	1.4±0.2	NS	1.9±0.6	<0.001
Alpha2 globulines (N:6–10)	6.8±0.9	6.3±1.24	0.001	6.6±1	NS	6.5±0.8	0.001	5.7±1.3	<0.001
Beta globuline (g/l) (N:6–10)	7.1±0.9	6.3±0.9	0.001	6.6±0.9	0.027	7.1±1.2	NS	6.9±1.6	NS
Gamma globulines (g/l) (N:6.4–13)	8.8±1.9	7.9±2.2	0.02	9±2.5	NS	8.2±1.7	0.042	19.6±8.1	<0.001
Glycemia (mmol/l) (N:3.9–6.1)	4.9±0.5	5.5±0.6	0.001	5.4±0.9	0.003	6.1±1.6	<0.001	6.7±3.2	<0.001
Cholesterol (mmol/l) (N:3.75–6.2)	5.8±1	5.3±1.3	NS	5.4±0.9	0.004	5.6±1.1	NS	4.1±1.3	<0.001
Triglycerides (mmol/l) (N:0.6–1.7)	1±0.5	1.3±0.6	0.015	1.1±0.6	NS	1.5±0.9	<0.001	1.1±0.6	NS
HDL (mmol/l) (N:0.9–2)	1.8±0.4	1.5±0.4	0.003	1.5±0.4	0.001	1.4±0.3	<0.001	1.1±0.5	<0.001
LDL (mmol/l) (N:3–4)	3.5±0.9	3.2±1	NS	3.4±0.8	NS	3.5±0.9	NS	2.6±1	<0.001
Creatinin (mmol/l) (N:62–115)	86±12	84±12	NS	88±10	NS	91±14	0.009	96±49	NS
AST (IU/l) (N: 0–50)	24±8	31±16	0.002	24.1±8	NS	32±16	<0.001	60±33	<0.001
ALT (IU/l) (N:0–60)	28±15	46±30	<0.001	25±11	NS	47±28	<0.001	35±26	NS
GGT (IU/l) (N:0–55)	29±23	32±21	NS	34±37	NS	72±64	<0.001	130±127	<0.001
ALK P (IU/l) (N:100–290)	131±33	147±38	NS	136±42	NS	134±49	NS	310±115	<0.001
Total bilirubin (μmol/l) (N:5–18)	12±5	11±5	NS	11±5	NS	13±5	NS	57±57	<0.001
Na (mmol/l) (N:135–145)	141±2	141±2	NS	141±2	NS	141±2	NS	134±6	<0.001
Ca (mmol/l) (N:2.1–2.55)	2.4±0.1	2.3±0.1	0.002	2.3±0.1	<0.001	2.4±0.1	0.001	2.1±0.2	<0.001
Serum iron (μM) (N: 12.5–25)	19.6±5.8	35.9±6.5	<0.001	20±9.1	NS	23±7.5	0.003	17.8±10.3	0.035

Table 1 (*Continued.*)

	CONTROL N=140	GH N=29	p^A	DGH N=37	p^A	DYSH N=67	p^A	AC N=58	p^A
Transferrin saturation % (N: 23–45)	33.7±10.9	77.7±14.5	<0.001	37.1±17.2	NS	39.7±15.4	0.019	44.8±30.9	NS
Ferritin (μg/l) (N: 55–345)	174±169	1168±1373	<0.001	37±76	<0.001	678±263	<0.001	410±545	NS
US-CRP (mg/l) (N<5)	2.1+−4.6	2.2+−2.7	NS	4.1+−10.6	0.018	2.3+−2.5	0.01	16.5+−19.8	<0.001
Hb (g/dl) (N:14–18)	14.6±1.2	14.7±1	NS	14.3±1.2	NS	14.9±1.1	NS	11.5±1.8	<0.001

partially absorbed at the interface following the attenuated total reflection rules. The optical signal is recorded at the output extremity of the fiber, providing the infrared single beam spectra. A reference single beam spectrum was obtained "in air" before each sample analysis. A 10-μl serum sample drop was then placed in physical contact with the fiber. Fiber was then removed from the serum sample and 100 scans were acquired and summarized. Before a new sample analysis, the fiber was carefully washed using distilled water and a new air spectrum was collected. To investigate the impact of the TAS glass on the analysis, we performed two independent analyses on the same sample but using two different glass fibers at two different dates.

2.4 Discrimination Methods

A partial least squares logistic (PLS logistic) method was used to discriminate samples.[20] The PLS logistic is an extension of the multiple linear regression model, allowing the derivation of multiclass decision rules and thus classification of data.[21] This model was described for instance in Ref. 21, and it is such that the probability that a patient p_i belongs to a group g given its recorded spectrum s is related to the linear predictor $s'\theta_g$ by the logistic link

$$P(p_i \in g | S=s) = \frac{\exp(s'\theta_g)}{1 + \sum_{k=1}^{N_g} \exp(s'\theta_k)},$$

where N_g denotes the number of groups and the θ_g are parameters that are estimated by a penalized maximum likelihood method. S denotes the value of the spectra for a set of wave numbers $\{\lambda_1, \ldots, \lambda_p\}$. In practice, no statistical feature selection is done in this analysis. The interpretation of this model is: the larger $s'\theta_g$, the higher the probability for the patient p_i to belong to the group g. Here, $s'\theta_g$ is a linear combination of the values of the absorbance at the different wavenumbers. 3/5 of the patients of each group were randomly selected to create the training group, which is used to estimate the parameter $\theta = \{\theta_{gj}, j=0, \ldots, p, g=1, \ldots, N_g\}$. The PLS was run on the whole transparency domain of the fiber, that is to say: 3800 to 2800 and 1800 to 850 cm^{-1}. The diagnosis given by the physician and resulting from the clinical, biological, and imaging procedures used in a context of good medical practice was used as gold standard.

The 2/5 remaining patients were then considered as a test set to validate the classification algorithm. Each test spectrum was introduced in the fitted PLS-logistic model to return the probability of the corresponding patient to belong to each group. The patient was then classified within the group with the highest probability, and the predicted group is compared to the true group of the patient. To validate regression models, it is usual to perform cross-validation. Two approaches can be considered: the leaf-by-leaf validation and validation on the basis of a test sample that is randomly chosen on the basis of all observations. The advantage of the second approach is that it allows inferring on the robustness of the classification method by repeating the validation with several random test samples. Therefore, the discrimination analysis was repeated ten times with distinct training and test groups that were randomly selected from the different groups and expressed results that correspond to the means.

Training and analysis were performed either with the five groups of patients, or using only two groups: the control group of healthy subjects and one group of diseased patients. The PLS-logistic method was also used to compare males and females. All computations, analyses, and graphics are executed with programs developed in Matlab v6.5 and Excel.

To investigate if some wavelengths have an important weight in the discriminant analysis, we evaluate the discriminant power by multiplying each standardized absorbance s_j by the parameter θ_{1j} of the logistic model. In practice, the weights are close to zero when the absorbance at the corresponding wavelength is not discriminant, and are large (in absolute value) if it is discriminant.

2.5 Statistical Analysis

Results of classical biochemical tests are expressed as mean ± standard deviation (SD). SPSS statistical analysis was performed by using the nonparametric Mann-Whitney test. The performance of the two groups of PLS-logistic classification was analyzed using the construction of receiving operating characteristic (ROC) curves and by the calculation of the area under the ROC (AUROC) using programs developed in Matlab. The ROC curve is a graphical plot of the sensitivity versus (1−specificity) for a binary classifier system as its discrimination threshold is varied. The cutoff of the regression score was identified through the calculation of the

Table 2 Confusion matrix for the five studied groups. The whole population, i.e., the five control and patients groups, were included in a global study that aimed at addressing each patient to the correct group using FEWS analysis compared to the bioclinical standard. Results are expressed in percentage corresponding to the mean of ten runs. The two independent analyses gave similar results which are both presented: the first number corresponds to the first analysis and the second one to the second analysis.

		FEWS classification				
		DGH	GH	DYSH	AC	CONTROL
Bioclinical classification	DGH	30-50	15-10	05-10	20-00	30-30
	GH	25-30	25-20	00-10	15-10	35-30
	DYSH	10-05	05-00	50-75	05-00	30-20
	AC	10-00	00-05	05-10	75-70	10-15
	CONTROL	15-10	05-05	03-05	02-05	75-75

highest Youden index (sensitivity+specificity-1), which gives a global index on the global performance of the test.[22] The performance of the five groups' PLS-logistic classification cannot be represented using ROC curves and is thus illustrated by a confusion matrix of classification as indicated in Table 2.

3 Results

3.1 Bioclinical Findings

Clinical and biological findings are indicated in Table 1. All the findings were in accordance with those previously reported in the literature and are characteristic of the different groups of patients.

Untreated genetic hemochromatotic patients were characterized, as expected, by an increase of all the serum iron parameters (serum iron, transferrin saturation, and ferritin) compared to healthy volunteers.[23] In addition, they presented: 1. a slight but significant decrease of albumin, transferrin, alpha2 and beta-globulins, cholesterol including both HDL and LDL, and calcemia; and 2. a slight increase of glycemia, triglycerides, AST, and ALT.

Iron depleted genetic hemochromatosis normalized their iron parameters as expected[23] and in addition exhibited persistent decrease of albumin, transferrin, and beta-globulins, as well as of HDL cholesterol and calcemia. The glycemia remained high, and AST was found decreased as compared to the control group. The hypoferritinemia was related to the fact
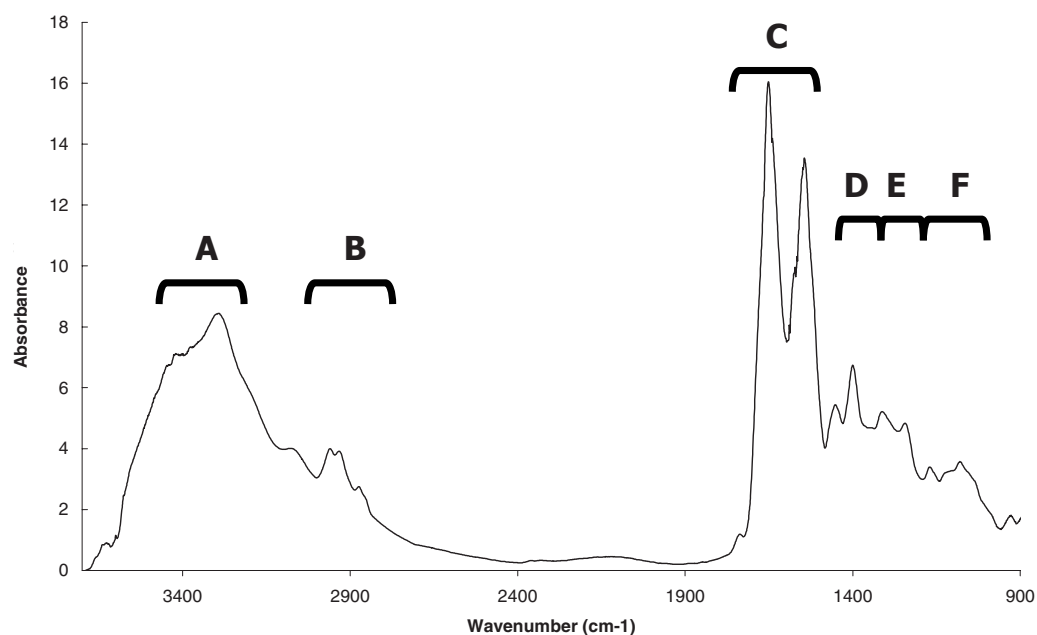

Fig. 1 Typical FEWS spectrum. This spectrum was obtained from 10-μL human normal serum. The figure displays the various frequency domains of absorbance of biomolecules. The overall spectra was used for discriminant analysis. Region A: protein amides A and B bands. Region B: CH_2, CH_3 of lipids. Region C: proteins amide I and II bands. Region D: fatty acids and amino acids. Region E: phospholipids. Region F: polysaccharides.
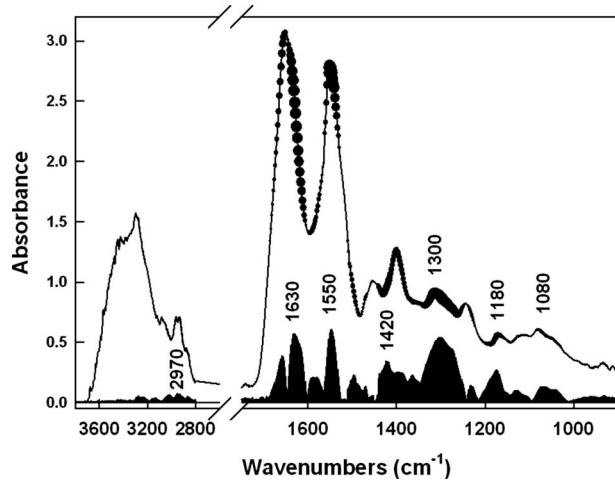
Fig. 2 Discriminating spectral domains. The figure displays a typical MIR spectrum acquired from blood samples. Superimposed to this spectrum, the statistical weights θ_g are featured both by the symbol sizes of the spectral curve and, for better visibility, by an histogram exhibiting the relative values of θ_g. This histogram clearly emphasizes that the discrimination between groups is based on a global signature, which encompasses spectral domains reflecting as a whole lipids, proteins, and sugars.

Fig. 3 ROC curves for patient classification using PLS regression. The line represents the average of ten runs for which the training group (3/5 of patients) and test group (2/5 of patients) were randomly selected. (a) Iron desaturated genetic hemochromatotic patients (DGH) and control patients. (b) Iron overloaded genetic hemochromatotic patients (GH) and control patients. (c) Alcoholic cirrhotic patients (AC). (d) Patients with dysmetabolic hepatosiderosis (DYSH). The corresponding AUROC are indicated on graphs. Standard deviations obtained from the ten runs are indicated on the graphs.

that the completion of iron depletion was judged on a drop of ferritinemia, which reflects the iron stores[24] and therefore the efficiency of the iron depletive treatment.

In the DYSH patient group, the level of a large number of markers of protein, glucose and lipid metabolisms were significantly different from those found in the control group,[25] but the modulation remained moderate compared to those found in the cirrhotic group of patients. In addition, patients presented an increase of ferritin protein, reflecting an increase of iron stores.

The alcoholic cirrhotic patient group exhibited the most important metabolic disturbances in terms of both number and severity of affected parameters. The only parameters that were not distinct from the control group were: beta-globulins, triglycerides, creatinine, ferritin, ALT, and transferrin saturation.

3.2 Fiber Evanescent Wave Spectroscopy Mid-Infrared Spectra

The mid-infrared spectral domain encompasses the 850 to 4000 cm^{-1} range. Attribution of absorbance bands has been previously reported.[8] Absorbance related to the serum samples (Fig. 1) are essentially found between 3600 and 2800 cm^{-1} as well as between 1700 and 900 cm^{-1}. The presence of functional groups present in proteins was revealed between 3100 and 3600 cm^{-1} for NH amide A and 1200 and 1700 cm^{-1} for the amide I, II, and III, respectively. The lipid absorption bands show up at 2800 to 3100 cm^{-1} and 1400 to 1740 cm^{-1} due to the presence of methylene and methyl C–H bonds and the triglycerides and phospholipids C=O ester bonds, The glucids C–O–C vibrations show up between 900 and 1200 cm^{-1}. Despite the possibilities of differences from one spectrum to another, visual analysis of spectra did not permit us to discriminate patients regarding the disease. However, the PLS-logistic model allows discrimi-
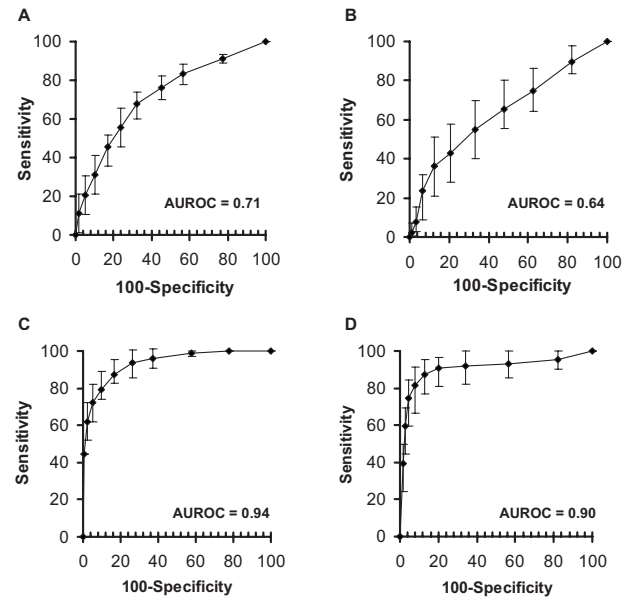
nating of the spectra thanks to the weights θ_g applied to the spectrum's values at the critical wave numbers. Figure 2 shows that the discrimination method between the different groups is based on a global signature encompassing lipids, proteins, and sugar spectral domains. Despite the fact that some wavelengths strongly contribute to the analysis, the best results are obtained using the whole spectra.

3.3 ROC Curves

To express results obtained from the PLS-logistic model, ROC curves were plotted (Fig. 3). The faster the ROC curve climbs toward the upper left-hand corner of the graph, the more the discrimination method is efficient. We observe that the proposed method was particularly efficient for the cirrhotic and DYSH patients, while its performance for the other patients was inferior. The area under ROC curves (AUROC) was computed to assess overall performance of our experiments. AUROC values, averaged from 10 PLS-logistic runs on the same sample, were 0.94 ± 0.05 for cirrhotic patients, 0.90 ± 0.06 for DYSH patients, 0.71 ± 0.10 for iron-depleted hemochromatotic patients, and 0.64 ± 0.11 for untreated hemochromatotic patients. Sensitivity, specificity, and positive and negative predictive values, determined using the optimum Youden index, are reported in Table 3. The second set of analysis gave similar results.

It is noteworthy that we were unable to discriminate between male and female in the validation group.

Table 3 Performance of FEWS, as clinical predictor, evaluated in the test groups. Values are given for the cutoff of the regression score, which corresponds to the highest Youden index. (AC: alcoholic Cirrhosis; DYSH: dysmetabolic hepatosiderosis; DGH: iron depleted genetic hemochromatosis; GH untreated genetic hemochromatosis; Se: sensitivity; Sp: specificity; PV: predictive value; AUROC: area under ROC curve; and CI: confiance interval).

	Cutoff	Se (95%CI)	Sp (95%CI)	Youden index	Positive PV	Negative PV	AUROC (95% CI)
Control versus AC	0.5	86±13	84±9	70	79.5	92	0.94±0.05
Control versus DYSH	0.5	87±13	87±8	74	85	92	0.90±0.06
Control versus DGH	0.5	68±23	68±12	36	40	85	0.71±0.10
Control versus GH	0.6	43±28	80±10	23	40	86	0.64±0.11

3.4 Patient Discrimination between Groups

To investigate the possibility to correctly address the patient in a more complex situation, we used the PLS method on the whole population. Results are reported in Table 2. We found that both controls and cirrhotic patients of the test group were correctly classified with success scores of 70 and 75%, respectively. Accurate classification was found in 55 to 75% of the cases for dysmetabolic hepatosiderosis patients, whereas most of them were not correctly classified in both hemochromatotic patient groups.

4 Discussion

Most systemic diseases lead to the development of metabolic consequences, which in turn expose patients to the development of complications that engage functional or vital prognosis. Such metabolic consequences, which are the direct consequence of cellular metabolism alterations, can be detected in plasma and/or serum.

However, to date, each metabolic abnormality is searched through the quantitative assessment of individual biochemical parameters, including the evaluation of the different glucid, lipid, and protein species.[1,2] This clearly implies that the physician must *a priori* hypothesize that one or more specific biochemical parameters are altered in the patient. In addition, current biochemical procedures explore only the serum components that have been previously characterized, and most of the biochemical tests do not provide qualitative assessment of biomolecules. Thus, during chronic alcoholism in absence of cirrhosis, the transferrin protein level is not altered, whereas an alteration of the glycosylation process, not evaluated by routine test, can be found.[26]

The quantification of combined known biological species quantification has been proposed to classify diseases such as Alzheimer disease[27] and liver fibrosis stage.[28,29] In addition, the development of new methodologies giving wide overview of metabolic profiling could be useful for diagnosis and has been proposed through the use of technologies devoted to metabolome characterization. Thus, surface-enhanced laser desorption ionization time of flight (SELDI-TOF) technology provides new hope. Indeed, this technology gives the opportunity to obtain more global information by providing peptide profiles that are supposed to reflect the pathological state by the modification of ratios between the different peaks (re-

views in Refs. 30 and 31). However, even though this technology is promising, the obtained information is only partial because the profiling is restricted to low molecular weight peptides (hence excluding most proteins) and, in addition, does not reflect the glucid and lipid metabolisms that can provide valuable information. In addition, some pitfalls requiring improvement of the methodology have been identified.[32] Taken altogether, it appears that the acquisition of global metabolic profiling through the SELDI-TOF method is time and cost consuming and not exhaustive.

Alternatively, the single shot acquisition of a global metabolic profiling could therefore be useful to improve patient care. The use of spectroscopy has been proposed to evaluate fingerprints in biomedia.[8]

Thus, on one hand, near-infrared spectroscopy that explores wavelengths comprised between 12500 to 4000 cm^{-1}, and allows the investigation of harmonic and combinatory vibrations of C–O, C–H, and N–O is proposed to characterize biological samples. NIR spectroscopy is used to measure oxygenation of tissue (review in Ref. 33) and has been reported useful for other applications such as evaluation of water content, tumoral process in diseased skin,[34,35] and of coronary artery plaque vulnerability.[36]

On other hand, the MIR domain (review in Ref. 8) explores the 850 to 4000 cm^{-1} wavenumber domain, and gives more global information due to the fact that the IR absorption in this domain reflects the fundamentals of the vibrations of the biomolecule functional groups present in complex biological fluids. Moreover, the location of specific peaks related to one specific chemical group may be altered by the presence of other molecules and/or functional groups.[7] Therefore, the spectra acquired in the MIR domain give access to multiple peaks representing a large number of functional groups present in proteins, lipids, and glucides. However, the OH vibration peak of water, always found in samples, overlaps with some bands, including amides bands. Despite this potential limitation, and taking advantage of the physical properties of the light diffusion within the TAS optic fiber, we decide to test the usefulness of the MIR FEWS approach to search for discriminating metabolic profiles by analyzing human serum samples with various metabolic diseases.

IR TAS glass fiber light transmission permits us to take advantage of the evanescent wave that propagates all along

the fiber,[16,37–43] and thus to perform FEWS. The general principle of this remote spectroscopy has already been described,[44] but for the first time here it is applied to a large scale medical study. Thus, the present work converts MIR FEWS into a technology with potential interest for the physician. The following points are essential to explain the sensitivity of our FEWS spectra compared to the classical way of collecting MIR spectra, by transmission or using an ATR plate. First, the recent development of tapered glass fiber exhibiting a diameter smaller than 100 μm on the sensing zone drastically increases the intensity of the evanescent wave propagating along the fiber, amplifying then the optical coupling interaction with the serum.[16,19,45] Second, the hydrophobic behavior of the covalent chalcogenide glasses enables us to emphasize the infrared signal absorbed by the biological molecules to the detriment of the invasive signal of water, which is the main medium in the serum.[41]

To validate the interest of this method, we have explored, by comparison with a normal group, patients with chronic liver diseases. Such diseases are known to have a potential impact on serum biochemical parameters, as illustrated in Table 1. This is related to the central role that the liver plays in the control of major metabolisms. Thus, as expected, GH impacts mainly the serum iron parameters but may also have other slight biochemical impacts on other parameters. The serum iron parameters were normalized after iron depletive treatment in the treated GH group. In the cirrhotic group, the major hepatic dysfunction affects most of the biochemical parameters studied. Finally, DYSH is mainly associated with disturbances in lipid (triglycerids and HDL cholesterol) and glucid metabolism, but also with increased iron stores reflected by ferritin increase. Therefore, we were expecting that FEWS-IR spectroscopy could discriminate these situations that differ by the metabolisms involved and the severity of the disease.

The sole inspection of individual spectra did not allow identifying abnormalities that characterize a particular pathological state. Therefore, we used supervised classification based on PLS-logistic regression to discriminate spectra and then patients. The obtained results may be improved in the future if the signal-to-noise ratio (SNR) could be increased. ROC curves demonstrate that applying the algorithm obtained in the training group by comparing each pathological group with the control group allowed us to correctly classify most of the cirrhotic patients as well as those presenting dysmetabolic hepatosiderosis. This is likely related to the presence of significant metabolic alterations found in these two groups of patients, especially in cirrhotic patients for which the number of altered parameters and the intensity of the modulation, compared to normal values, are major. Indeed, alcoholic cirrhotic patients in our study presented severe hepatic dysfunction with multiple biochemical alterations including proteins, glucids, and lipids. Patients with dysmetabolic hepatosiderosis also exhibited metabolism alterations especially involving lipids and glucids parameters, but the relative difference compared to normal values is lower, and a more reduced number of parameters related to proteins is affected (Table 1). At variance, we were not able to correctly classify iron saturated or depleted hemochromatotic patients despite alterations and

which are known to induce oxidative stress toward biomolecules in iron saturated groups.[46]

Whether the small but significant difference of age in cirrhotic and dysmetabolic patients, which are older than those in the three other groups, may have an impact on the FEWS-IR analysis needs to be discussed. Considering that: 1. within the biochemical parameters usually evaluated in serum most of them are not influenced by age, and the normal range is not affected by age in the adult population, except for lipid metabolism, for which normal values are only very slightly different with age in adults,[47] and 2. such modulations are extremely moderate regarding the intensity of alterations found in biochemical parameters in cirrhotic and dysmetabolic groups, they could only have a slight impact on the analysis. In addition, it is noteworthy that the best performance of the analysis is obtained when the entire spectrum is considered, and not only one or some specific wavelength that could only reflect parameters related to a single metabolism. This last point clearly underlines the global investigation that FEWS-IR analysis permits.

The previously described procedure, based on the comparison between one pathological group and one control group, implies that the physician hypothesizes one diagnosis for the patient. However, in most cases, the diagnostic hypothesis is less precise and more than one diagnosis can be evoked. To test the possibility to classify a given patient among a panel of possible diseases, we tested our capacity to correctly identify the cirrhotic and dysmetabolic patients among the five possible groups. Despite that common metabolic abnormalities can be found in the different groups, such as glucid and lipid alterations in both cirrhotic and dysmetabolic patients, we were clearly able to discriminate these patients. This is likely related to the fact that the spectral analysis gives a large number of data, thus performing true metabolic profiling (Fig. 2), characteristic of each situation.

In conclusion, our results suggest that the concept of the metabolic profiling using IR-FEWS could be a way to investigate diseases having metabolic consequences in patients.

Acknowledgments

References

1. H. C. Sox, "The evaluation of diagnostic tests: principles, problems, and new developments," *Annu. Rev. Med.* **47**, 463–471 (1996).
2. A. Tatsioni, D. A. Zarin, N. Aronson, D. J. Samson, C. R. Flamm, C. Schmid, and J. Lau, "Challenges in systematic reviews of diagnostic technologies," *Ann. Intern Med.* **142**(12 Pt 2), 1048–1055 (2005).
3. R. H. Eckel, S. M. Grundy, and P. Z. Zimmet, "The metabolic syndrome," *Lancet* **365**(9468), 1415–1428 (2005).
4. L. K. Schnackenberg, "Global metabolic profiling and its role in systems biology to advance personalized medicine in the 21st century," *Expert. Rev. Mol. Diagn.* **7**(3), 247–259 (2007).
5. L. K. Schnackenberg and R. D. Beger, "Monitoring the health to disease continuum with global metabolic profiling and systems biology," *Pharmacogenomics* **7**(7), 1077–1086 (2006).
6. G. Schlotterbeck, A. Ross, F. Dieterle, and H. Senn, "Metabolic profiling technologies for biomarker discovery in biomedicine and drug development," *Pharmacogenomics* **7**(7), 1055–1075 (2006).

7. M. Jackson and H. Mantsch, "Biomedical infrared spectroscopy," *Infrared Spectroscopy of Biomolecules*, D. Chapman, Ed., pp. 311–340, Wiley-Liss, New York (1996).

8. M. N. Pons, S. Le Bonte, and O. Potier, "Spectral analysis and fingerprinting for biomedia characterisation," *J. Biotechnol.* **113**(1–3), 211–230 (2004).

9. M. Jackson, J. R. Mansfield, B. Dolenko, R. L. Somorjai, H. H. Mantsch, and P. H. Watson, "Classification of breast tumors by grade and steroid receptor status using pattern recognition analysis of infrared spectra," *Cancer Detect. Prev.* **23**(3), 245–253 (1999).

10. D. Naumann, "Infrared spectroscopy in microbiology," in *Encyclopedia of Analytical Chemistry*, pp. 102–131, Wiley and Sons Ltd., Chichester, UK (2000).

11. D. Rohleder, G. Kocherscheidt, K. Gerber, W. Kiefer, W. Kohler, J. Mocks, and W. Petrich, "Comparison of mid-infrared and Raman spectroscopy in the quantitative analysis of serum," *J. Biomed. Opt.* **10**(3), 031108 (2005).

12. S. Spassov, M. Beekes, and D. Naumann, "Structural differences between TSEs strains investigated by FT-IR spectroscopy," *Biochim. Biophys. Acta* **1760**(7), 1138–1149 (2006).

13. N. Toyran, P. Lasch, D. Naumann, B. Turan, and F. Severcan, "Early alterations in myocardia and vessels of the diabetic rat heart: an FTIR microspectroscopic study," *Biochem. J.* **397**(3), 427–436 (2006).

14. P. Lasch and D. Naumann, "Spatial resolution in infrared microspectroscopic imaging of tissues," *Biochim. Biophys. Acta* **1758**(7), 814–829 (2006).

15. M. Beekes, P. Lasch, and D. Naumann, "Analytical applications of Fourier transform-infrared (FT-IR) spectroscopy in microbiology and prion research," *Vet. Microbiol.* **123**(4), 305–319 (2007).

16. S. Hocde, O. Loréal, O. Sire, C. Boussard-Pledel, B. Bureau, B. Turlin, J. Keirsse, P. Leroyer, and J. Lucas, "Metabolic imaging of tissues by infrared fiber-optic spectroscopy: an efficient tool for medical diagnosis," *J. Biomed. Opt.* **9**(2), 404–407 (2004).

17. Y. Gandon, D. Olivie, D. Guyader, C. Aube, F. Oberti, V. Sebille, and Y. Deugnier, "Non-invasive assessment of hepatic iron stores by MRI," *Lancet* **363**(9406), 357–362 (2004).

18. J. Boursier, E. Cesbron, A. L. Tropet, and C. Pilette, "Comparison and improvement of MELD and Child-Pugh score accuracies for the prediction of 6-month mortality in cirrhotic patients," *J. Clin. Gastroenterol.* **43**(6), 580–585 (2009).

19. J. Keirsse, E. Lahaye, A. Bouter, V. Dupont, C. Boussard-Pledel, B. Bureau, J. L. Adam, V. Monbet, and O. Sire, "Mapping bacterial surface population physiology in real-time: infrared spectroscopy of Proteus mirabilis swarm colonies," *Appl. Spectrosc.* **60**(6), 584–591 (2006).

20. J. Zhu and T. Hastie, "Classification of gene microarrays by penalized logistic regression," *Biostat.* **5**(3), 427–443 (2004).

21. G. Fort and S. Lambert-Lacroix, "Classification using partial least squares with penalized logistic regression," *Bioinformatics* **21**(7), 1104–1111 (2004).

22. H. Huang, M. L. Shiffman, S. Friedman, R. Venkatesh, N. Bzowej, O. T. Abar, C. M. Rowland, J. J. Catanese, D. U. Leong, J. J. Sninsky, T. J. Layden, T. L. Wright, T. White, and R. C. Cheung, "A 7 gene signature identifies the risk of developing cirrhosis in patients with chronic hepatitis C," *J. Hepatol* **46**(2), 297–306 (2007).

23. C. Le Lan, O. Loreal, T. Cohen, M. Ropert, H. Glickstein, F. Laine, M. Pouchard, Y. Deugnier, A. Le Treut, W. Breuer, Z. I. Cabantchik, and P. Brissot, "Redox active plasma iron in C282Y/C282Y hemochromatosis," *Blood* **105**(11), 4527–4531 (2005).

24. P. Brissot, M. B. Troadec, E. Bardou-Jacquet, C. Le Lan, A. M. Jouanolle, Y. Deugnier, and O. Loreal, "Current approach to hemochromatosis," *Blood Rev.* **22**(4), 195–210 (2008).

25. R. Moirand, A. M. Mortaji, O. Loreal, F. Paillard, P. Brissot, and Y. Deugnier, "A new syndrome of liver iron overload with normal transferrin saturation," *Lancet* **349**(9045), 95–97 (1997).

26. R. G. Batey and G. Madsen, "The carbohydrate-deficient transferrin test in hospital practice," *Drug Alcohol. Rev.* **17**(1), 105–109 (1998).

27. S. Ray, M. Britschgi, C. Herbert, Y. Takeda-Uchimura, A. Boxer, K. Blennow, L. F. Friedman, D. R. Galasko, M. Jutel, A. Karydas, J. A. Kaye, J. Leszek, B. L. Miller, L. Minthon, J. F. Quinn, G. D. Rabinovici, W. H. Robinson, M. N. Sabbagh, Y. T. So, D. L. Sparks, M. Tabaton, J. Tinklenberg, J. A. Yesavage, R. Tibshirani, and T. Wyss-Coray, "Classification and prediction of clinical Alzheimer's diagnosis based on plasma signaling proteins," *Nat. Med.* **13**(11), 1359–1362 (2007).

28. T. Poynard, F. Imbert-Bismut, M. Munteanu, and V. Ratziu, "FibroTest-FibroSURE: towards a universal biomarker of liver fibrosis?" *Expert Rev. Mol. Diagn.* **5**(1), 15–21 (2005).

29. P. Cales, F. Oberti, S. Michalak, I. Hubert-Fouchard, M. C. Rousselet, A. Konate, Y. Gallois, C. Ternisien, A. Chevailler, and F. Lunel, "A novel panel of blood markers to assess the degree of liver fibrosis," *J. Hepatol* **42**(6), 1373–1381 (2005).

30. M. Kiehntopf, R. Siegmund, and T. Deufel, "Use of SELDI-TOF mass spectrometry for identification of new biomarkers: potential and limitations," *Clin. Chem. Lab. Med.* **45**(11), 1435–1449 (2007).

31. T. C. Poon, "Opportunities and limitations of SELDI-TOF-MS in biomedical research: practical advices," *Expert Rev. Proteomics* **4**(1), 51–65 (2007).

32. M. Dijkstra, R. J. Vonk, and R. C. Jansen, "SELDI-TOF mass spectra: a view on sources of variation," *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **847**(1), 12–23 (2007).

33. T. Hamaoka, K. K. McCully, V. Quaresima, K. Yamamoto, and B. Chance, "Near-infrared spectroscopy/imaging for monitoring muscle oxygenation and oxidative metabolism in healthy and diseased humans," *J. Biomed. Opt.* **12**(6), 062105 (2007).

34. N. S. Eikje, K. Aizawa, and Y. Ozaki, "Vibrational spectroscopy for molecular characterisation and diagnosis of benign, premalignant and malignant skin tumours," *Biotechnol. Annu. Rev.* **11**, 191–225 (2005).

35. N. S. Eikje, Y. Ozaki, K. Aizawa, and S. Arase, "Fiber optic near-infrared Raman spectroscopy for clinical noninvasive determination of water content in diseased skin and assessment of cutaneous edema," *J. Biomed. Opt.* **10**(1), 14013 (2005).

36. J. D. Caplan, S. Waxman, R. W. Nesto, and J. E. Muller, "Near-infrared spectroscopy for the detection of vulnerable coronary artery plaques," *J. Am. Coll. Cardiol.* **47**(8 Suppl), C92-6 (2006).

37. D. Le Coq, K. Michel, J. Keirsse, C. Boussard-Pledel, G. Fonteneau, B. Bureau, J. Le Quéré, O. Sire, and J. Lucas, "Infrared glass fibers for in-situ sensing chemical and biochemical reactions," *C. R. Chim.* **5**, 1–7 (2003).

38. H. Steiner, M. Jakusch, M. Kraft, M. Karlowatz, T. Baumann, R. Niessner, W. Konz, A. Brandenburg, K. Michel, C. Boussard-Pledel, B. Bureau, J. Lucas, Y. Reichlin, A. Katzir, N. Fleischmann, K. Staubmann, R. Allabashi, J. M. Bayona, and B. Mizaikoff, "In situ sensing of volatile organic compounds in groundwater: first field tests of a mid-infrared fiber-optic sensing system," *Appl. Spectrosc.* **57**(6), 607–613 (2003).

39. P. Lucas, D. Le Coq, C. Juncker, J. Collier, D. E. Boesewetter, C. Boussard-Pledel, B. Bureau, and M. R. Riley, "Evaluation of toxic agent effects on lung cells by fiber evanescent wave spectroscopy," *Appl. Spectrosc.* **59**(1), 1–9 (2005).

40. P. Lucas, M. R. Riley, C. Boussard-Pledel, and B. Bureau, "Advances in chalcogenide fiber evanescent wave biochemical sensing," *Anal. Biochem.* **351**(1), 1–10 (2006).

41. P. Lucas, M. Solis, D. Le Coq, C. Junker, M. R. Riley, J. Collier, D. E. Boesewetter, C. Boussard, and B. Bureau, "Infrared biosensors using hydrophobic chalcogenide fibers sensitized with live cells," *Sens. Actuators B* **119**, 355–362 (2006).

42. K. Michel, B. Bureau, C. Boussard-Pledel, T. Jouan, J. Adam, K. Staubmann, and T. Baumann, "Monitoring of pollutant in waste water by infrared spectroscopy using chalcogenide glass optical fibers," *Sens. Actuators B* **101**, 252–259 (2004).

43. X. Zhang, B. Bureau, P. Lucas, C. Boussard-Pledel, and J. Lucas, "Glasses for seeing beyond visible," *Chem.* **14**(2), 432–442 (2008).

44. Y. Raichlin and A. Katzir, "Fiber-optic evanescent wave spectroscopy in the middle infrared," *Appl. Spectrosc.* **62**(2), 32A–72A (2008).

45. S. Hocde, C. Boussard-Pledel, G. Fonteneau, and J. Lucas, "Chalcogens based glasses for IR fiber chemical sensors," *Solid State Sci.* **3**(3), 279–284 (2001).

46. D. Galaris and K. Pantopoulos, "Oxidative stress and iron homeostasis: mechanistic and health aspects," *Crit. Rev. Clin. Lab Sci.* **45**(1), 1–23 (2008).

47. "Plasma lipid distributions in selected North American populations: the Lipid Research Clinics Program Prevalence Study. The Lipid Research Clinics Program Epidemiology Committee," *Circulation* **60**(2), 427–439 (1979).