Current Trends in the Development of FTIR Imaging for the Quantitative Analysis of Biological Samples

C. Petibois^{*}, K. Wehbe, K. Belbachir, R. Noreen and G. Déléris

Université de Bordeaux, CNRS UMR 5084, CNAB

146 rue Léo Saignat, F-33076 Bordeaux Cedex, France

Fourier transform infrared imaging instrumentation has come of age for the rapid acquisition of biosample IR images, and thus now allows developing analytical methods based upon the molecular information of samples contents. For the biomedicine field, Fourier transform infrared imaging should be able to play a role in the molecular characterization of cells and tissues where no other analytical method provides quantitative information. Now, a compromise may be obtained between an acceptable acquisition time of IR images, which should never exceed a few tens of minutes, and the necessary spatial resolution (down to the diffraction limit, although limited for biology), spectral resolution (2 to 8 cm⁻¹ for biosample analyses), and signal-to-noise ratio level, to provide a diagnostic answer to clinicians during the surgery time. Here, we will discuss the potential of Fourier transform infrared imaging in face to major pathologies (myopathies, brain tumors, metabolic diseases) for which current imaging methods remain unable to provide sufficient information for a precise diagnosis. Quantitative molecular information may be extracted from infrared images of samples as soon as samples volume/thickness is controlled as well as absorption and absorptivity of the molecules to analyze may be isolated from other absorbing compounds. In this context, metabolic parameters appear as the main targets for providing critical information about the physiological status of a biosample due to their characteristic infrared spectra. As examples, it will be shown how to isolate and quantify glucose, lactic-acid, urea, etc. absorptions from complex biosample infrared images.

PACS numbers: 82.80.-d

1. Introduction

Fourier transform infrared (FTIR) spectroscopy was first developed for analytical chemistry purposes in 1905 by William Coblentz, who acquired the very first spectra database of molecules and released tables of wavelengths at which various materials absorbed IR light. From these pioneer studies, FTIR spectroscopy has been long considered as a unique analytical resource to probe the structure of small molecules. This vibrational technique is still considered in chemistry labs as an essential tool in the panel of chemical information required to characterize any purified molecule. The most convenient way to probe a molecular structure was to prepare KBr pellets with small amount of a chemical species for a spectrum acquisition in transmission. Now, after a century of instrumental developments, the essentiality of this technique has moved from the single molecule analysis towards the understanding of complex biological species [1]. In the 90's, FTIR spectroscopy instrumentation turned to other applications with the release of spectrometers able to collect spectra on sample holders dedicated to quantitative, thus reproducible, measurements. As an example, the Bruker IFS 28/B spectrometer was one of the first to offer the opportunity of spreading biofluids [2] of cell lysates [3] in

delimitated sample holders (ZnSe) for transmission measurements. In this way, a given volume of sample could be deposited on a standardized surface dimension (7 mm diameter circle in the case of IFS 28/B spectrometer). Studies on biofluids could show that triplicate measurements allowed to obtain high reproducibility in spectra acquisitions with no more than 1% of change in absorptions on average [4]. Substrates holding 96 to 484 delimitated sample locations have now been released, allowing high throughput measurements as can be performed in most modern biochemical labs of hospitals. Another major advance in FTIR instrumentation for biological applications was the development of microscopes and now imaging systems able to work in both the visible and IR modes [5]. FTIR imaging systems have the advantage of a fast spectra acquisition since the microscope is coupled to an array of IR detectors, which may be linear (16 or 32) detectors) or focal plane $(64 \times 64 \text{ to } 256 \times 256 \text{ detectors})$. The reduction of time for spectra acquisition [6] permits applications to large tissue areas and thus opens IR to the biomedical field. Cell imaging applications are also possible if we consider a few analytical requirements to fit hot questions in biology.

2. Analytical requirements for bioimaging

Molecular bioimaging has imposed a new way to analyze cells and tissues. From vibrational spectroscopies (mainly infrared, and Raman) to particle-based tech-

^{*} corresponding author; e-mail: cyril.petibois@u-bordeaux2.fr

niques (using energy of X-rays, γ -rays, etc.), a wide range of techniques now fit some of the analytical features required for biological investigations [7]. The first feature is about the spatial resolution, which is determined by the physical principle of a technique [8]. However, the spatial resolution achieved by a technique also defines its field of applications. Tissue investigations require to obtain data only at the millimeter scale for morphology analysis (anatomic imaging), but the micrometer scale may be required to distinguish between subcomponents of an organ, such as connective tissue, cell assembly and organization, vascular network, etc. The most popular techniques for *in vivo* routine biomedical imaging, magnetic resonance imaging (MRI) and positron emission tomography (PET), have reached a 1–2 mm spatial resolution in the best situation and thus allow determining the location of pathological tissues without the need for any surgery. However, these techniques are not suitable for single cell investigations and remain limited for tissue volumes below 3–5 mm³. Single cell analyses require to reach the micrometer scale or below, depending on the requirement for attaining the subcellular components. To date, at such spatial resolution, vibrational spectroscopy--based imaging techniques has remained unable to provide sample information with reasonable quality since the spectra obtained usually present a low signal-to-noise (S/N) ratio, making difficult any further spectral data treatment (Fig. 1). On the other side, several particle--based techniques, although potentially destructive for the sample, are able to provide information within a wide nm-mm scale.

A second major feature to consider in biological imaging is the temporal resolution, which will determine the applications of techniques for time-resolved analyses. This field covers analytical situations going from samples stabilized in a given state (crystalline, soluble, gasified, cryofixed, lyophilized, included, etc.) towards molecular dynamics of still living specimens. In stabilized states, there is usually no time limit for data acquisitions. The main issue to address is to maintain the sample in a stable state during data acquisitions. Of course, all techniques are able to provide information about a sample in such a situation. However, a stable state may be required for analytical techniques unable to fit with the living environment, such as for infrared spectroscopy-based techniques, which can hardly operate if water is present in the sample and thus requires specific instrumentation for such a purpose [9]. Therefore, here again, the physical properties of imaging modalities will determine their applicability on time-resolved analyses in biology. Most of modalities using accelerator-based sources are not designed for analyzing living specimens due to the utilization of high energy beams. However, molecular dynamics may be determined out from the living environment, such as for the analysis of proteins folding or conformation under stress conditions (pH, temperature, etc.), where high energy may be required for time-resolved data acquisitions within the ms-ns range of time [10]. The third main



Fig. 1. Example of FTIR cell imaging using different spatial resolutions. (A) $6.25 \times 6.25 \ \mu m^2$ spatial resolution obtained with a Perkin-Elmer Spotlight 300 FTIR imaging system (detector dimension $25 \times 25 \ \mu m^2$ with a $4 \times$ numerical magnification). 64 scans were required to obtain high S/N (= 1324), which is sufficient for further data treatments. (B) $2.66 \times 2.66 \ \mu m^2$ spatial resolution obtained with a Bruker Hyperion 3000 imaging system (detector dimension $40 \times 40 \ \mu m^2$ with a $15 \times nu$ merical magnification; synchrotron radiation IR source from Frascati-Roma facility, Daphne-Light laboratory, Italy). 128 scans were required to obtain sufficient S/N (= 527). (C) The same instrumentation as for (B) but with a $36 \times$ numerical magnification. S/N obtained was only 21 and does not permit further spectral data treatment. All FTIR images were normalized in intensity on a common scale (a.u. of spectrometer) after integration of the $1700-1600 \text{ cm}^{-1}$ spectral interval (amide I) for every pixel.

feature to consider for determining the applicability of imaging modalities to the biological field is the sensitivity. In this domain, it is clear that accelerator-based sources play their most significant role for enhancing analytical properties of imaging modalities that may be implemented on synchrotrons. The high energy, brilliance or brightness obtained from accelerator-based sources offer unique opportunity for imaging modalities to reach nmol/l to ppm concentrations, i.e., up to trace element analyses, that conventional sources cannot. The modulation of spatial resolution and beam intensity may be considered as the main methodological aspect to manage for optimizing a given sample analysis. In other words, the use of the higher spatial resolution leads to a decrease in beam intensity, and conversely. As a consequence of their physical properties, every imaging modality has advantages and limits over the other techniques that determine a quite well defined field of applications for biological investigations.

3. Spectral, spatial, and step-size resolutions for FTIR bioimaging

According to the above mentioned requirements allowing fitting to actual hot questions in biological sciences, a $1 \times 1 \ \mu m^2$ spatial resolution is required for bioimaging, at least for cell studies. This spatial resolution is not available in the true physics of IR spectroimaging if a confocal-like setup is not available on the microscope. One may note also that FTIR imaging system holding linear arrays of detectors are surely not designed to offer a confocal-like geometry. Thus, with the current instrumentations available for FTIR microscopy and imaging, there is no solution proposed to reach the desired spatial resolution on biosamples. The spectral resolution is diffraction limited, i.e., the spectral information collected from a sample results from a sample area of $10 \times 10 \ \mu m^2$ at 1000 cm^{-1} (where carbohydrates and their by-products absorb in biological samples) or $3 \times 3 \ \mu m^2$ at 3000 cm⁻¹ (where fatty acyl chains absorb). This means that the $1 \times 1 \ \mu m^2$ spatial resolution remains absolutely unavailable to any molecular content of biological sample in the mid-infrared spectral interval [11]. One possibility would be to reduce the data collected by turning back to interferograms and modifying the information included to resolve only the $1 \times 1 \ \mu m^2$ area. Such a process has not yet been developed for reconstructing images containing hundreds to thousands of spectra, which might look like a tour de force for people involved in post-processing data treatments!

As a consequence, the current situation makes that every biologist searching for a $1 \times 1 \ \mu m^2$ spatial resolution in FTIR imaging applications must accept an oversampling procedure allowing a step-size resolution. From the point of view of the physicist, this procedure does not offer a spatial resolution of this dimension since the limit of diffraction is the sole parameter to be considered on this issue. Nevertheless, from the point of view of the biospectroscopist, disclosing the spectral and spatial resolutions issues is not an aberration as soon as the two parameters remain considered while interpreting the data obtained from the sample. In fact, it all depends on the biological question to address. If a $3 \times 3 \ \mu m^2$ spatial (and spectral) resolution is sufficient to study lipids in cell membranes, a better spatial resolution would just induce unnecessary roots of errors and criticism to the measurements. But, in many cases, the $1 \times 1 \ \mu m^2$ spatial resolution is necessary; otherwise the biological question is not addressed. One simple example may illustrate this position: if we consider a series of cellular diseases that imply an abnormal metabolic coordination between the



Fig. 2. Simulation of the spectral information obtained using an oversampling procedure for $1.1 \times 1.1 \ \mu m^2$ step-size resolution. Left: the cell is imaged with a $6.25 \times 6.25 \ \mu \text{m}^2$ spatial resolution to analyze the interface between membrane and extracellular matrix. The spectrum obtained at such a resolution offers averaged information about these different components. Right: an illustration of the $1.1 \times 1.1 \ \mu m^2$ spatial resolution obtained by oversampling with a FPA magnified by a $36 \times$ objective on the microscope. The pixels selected were centered on the membrane and on the extracellular matrix. If the molecular absorptions found in these spectra remain diffraction limited, there is circumstantial evidence that the two pixels will cover different contributions from the subcellular components, allowing discrimination between them.

cytosol and the mitochondria inside the cell, it is required to be able to distinguish between common metabolic parameters (at least pyruvic acid, ATP, and ADP) at the interface of these two subcellular areas, i.e., in a 1–3 μ m diffusion distance around mitochondria [12]. First, we must consider that a normal mitochondrion has a 3D dimension of 1 μ m. Second, adenosine and phosphates (from ATP and ADP), but also glucose, have their most characteristic absorptions in the $1200-900 \text{ cm}^{-1}$ spectral interval. Thus, the spectral resolution obtained considering the diffraction limitation is roughly $10 \times 10 \ \mu m^2$ at this spectral interval. As a consequence, we must consider that the problem posed by the utilization of a $10 \times 10 \ \mu m^2$ spatial resolution is that no specific information might be obtained to distinguish between the cytosol and mitochondria. However, if we accept a stepsize that is $\approx 1 \times 1 \ \mu m^2$ using a focal plane array (FPA) with $40\times40~\mu\mathrm{m}^2$ detectors and a $36\times$ magnification, we collect spectra which have an intrinsic information that is diffraction limited, but spectra give new information every $1 \times 1 \ \mu m^2$. The remaining question is about the ability to ensure that spectral information belonging to mitochondria and cytosol may be distinguished in this way. In principle, the answer should be positive since we are in a configuration where we can select a spectrum covering a mitochondrion (of course also with the cytosol

surrounding the mitochondrion) and the spectra closer to the mitochondrion, but without information belonging to this latter. Other major cell biology issues require high spatial resolution, namely interfaces between subcellular components, such as between membranes and extracellular matrix (Fig. 2). This makes an important difference with a $10 \times 10 \ \mu m^2$ step-size resolution since a $10 \ \mu m$ distance between two spectra corresponding respectively to the mitochondrion and to the cytosol is larger than the space of diffusion $(1-3 \ \mu m)$ for the molecules involved in the abnormal metabolism regulation between these two subcellular elements. The $10 \times 10 \ \mu m^2$ spatial resolution would be just unsuccessful to provide any information in this case.

4. Extracting molecular information from biosample spectra

Now, the question to address is about the extraction or isolation of relevant molecular information for a quantitative analysis. This question lies with a field named chemometrics, a wide range of spectral data treatment methods that covers mainly two kinds of applications: 1 -statistical treatments for classifying spectra on the basis of their absorptions; 2 — mathematical treatments for extracting one or more absorptions from spectra for further comparisons. If we consider only the biological applications, the aim of classification is to find spectral data that will allow distinguishing between spectra representing different physiological situations; healthy vs. cancerous as an example. Usually, only a qualitative result is provided [13]. On the other side, mathematical treatments may find their legitimacy if a quantitative analysis is required. In this case, a correlation must be found between spectra IR absorptions and sample molecular contents. In biosample spectra, it has been shown that IR absorptions may be found characteristic of individual molecules or families of molecules (Fig. 3). If we consider the main biomolecules that may be found in biological samples, there is a small variety between biological species if we exclude a few examples [14]. For cells and soft tissues, water is the main component, with 60 to 80% in volume. From the remaining, 60-70% belongs to proteins, 15-25% to lipids, 1-10% to carbohydrates, 1–3% to nucleic acids, 0.9–1.5% to salts, 0.1-0.2% to metals, and less than 2% for the remaining species. Taking into account that salts, metals, cations, and trace species have only weak or no absorptions in the mid-IR interval, only lipids, proteins, and carbohydrates may be analyzed in biosamples. Thus, considering these biomolecular components, fatty acyl chains $(3020-2800 \text{ cm}^{-1})$ and proteins $(1700-1500 \text{ cm}^{-1})$ may be analyzed as families of molecules and only global information will be obtained. Alternatively, absorptions belonging to carbohydrates (glucose, glycogen) and their by-products (pyruvic acid, lactic acid) are available with a good characterization in the $1200-900 \text{ cm}^{-1}$ spectral interval (Fig. 4). Actually, the best way to extract



Fig. 3. Typical spectrum and band assignment for most of biological samples. (a) — rocking, (b) — asymmetrical stretching, (c) — scissoring, (d) — symmetrical stretching, (e) — wagging, (f) — wagging, (g) — twisting.

molecular information seems to be the curve-fitting of the considered spectral intervals, but several conditions are required (Fig. 4): 1 — the second derivative procedure must be used to find out the absorptions bands contributing to the global spectrum absorption in the spectral interval; 2 -thus, a high S/N is required to avoid the formation of absorption bands in the second derivative spectrum, at least for spectral intervals where the intensity may be weak for biological samples $(1200-900 \text{ cm}^{-1} \text{ as an example}); 3 - \text{band shape (with }$ Gaussian and Lorentzian contributions for most of mid--infrared absorption bands) and limits (wave numbers) must be carefully fixed to avoid errors in the prediction of band areas, which will be the final result to monitor. For the first point, there is only one parameter to fix carefully, which is the sensitivity of curve shape changes to assess absorptions bands. A 1% sensitivity or below is usually convenient to ascertain the positioning of every band contributing to the spectral interval but, if S/N is poor, 2 to 5% may be applied, with the risk that existing bands will not be revealed. In fact, in most cases, a poor S/N prohibits the use of this procedure and curve fitting thus becomes impossible to perform. The third point, which is to fix band shape and limits, there is still an intense debate in the spectroscopists community. In spectral intervals where only global information about families of molecules can be obtained (lipids and proteins), the debate is about the assignment of the bands rather than their position or shape. As an example, for determining protein secondary structure, it has become common to use the curve fitting of the amide I spectral interval. Depending on the presence or not of collagen in the samples, which exhibit a α -type helix considered as representative of the triple helix conformation of the protein, 7 to 8 absorption bands will be revealed from the second derivative spectrum of the sample. There is a certain consen-



Fig. 4. Typical curve-fitting results obtained from biological samples. (A) The second derivative spectrum of the sample is used to determine the position of every band contributing to the spectral interval (here, a skeletal muscle tissue sample containing collagens, i.e., a triple helix band is found at 1648 cm^{-1}). (B) Curvefitting of the same spectral interval after positioning the bands for their maximal intensity according to second derivative results in (A). (C) Curve-fitting of the 3050–2800 $\rm cm^{-1}$ spectral interval for determining the ν (C=CH), ν_{as} (CH₂), and ν_{as} (CH₃) absorption bands of fatty acyl chains. (D) Curve-fitting of the 1200- 900 cm^{-1} spectral interval for determining the characteristic absorptions of glucose (1033 cm^{-1}), glycogen (1011 cm^{-1}) , and lactic acid (1127 cm^{-1}) in a skeletal muscle tissue sample. Band assignation was assessed according to the curve-fitting of pure product spectra from these biomolecules.

sus in the community about the position of these bands. However, what is still debated is the assignment of these bands (α -helix, β -sheets, β -turns, unordered structure, triple helix). For other spectral intervals, the debate may turn on the positioning of the bands and possibly the number of bands that contribute to the spectral absorption. In this case, the main problem is about the composition of the sample, which must be known in advance to allow positioning the bands of every molecule. To do so, it is also necessary to proceed to the curve-

fitting of the pure product spectra of these molecules in order all their absorption bands are taken into account in the curve-fitting of the biological sample. In a perfect world, this would ensure that all contributions are considered. In the true practice however, it is obvious that the exact molecular composition of a sample cannot be known and that a range of these chemical species are not available for pure product spectra (enzymes, hormones, membranes-bound molecules, DNA, mRNA, etc.) [15]. As a consequence, a compromise must be accepted and only the main components of biological samples, if available, must be considered. If we focus on metabolic parameters absorbing in the $1200-900 \text{ cm}^{-1}$ spectral interval, the most concentrated species are glucose (\approx 5 mmol/l), lactic-acid (\approx 2 mmol/l), pyruvic- acid $(\approx 0.1-0.5 \text{ mmol/l})$, glycerol ($\approx 0-1 \text{ mmol/l})$, and possibly glycogen in some tissues (
 \approx 2–5 mmol/l). As soon as these molecular species may be distinguished in a given spectral interval (Fig. 4D), their concentration may be determined according to the absorptivity at their most characteristic absorption band: concentration = absorptivity (mmol l^{-1} a.u.⁻²) × absorbance (a.u.⁻²) × sample volume (1), where a.u. $^{-2}$ are arbitrary units of band area and sample volume takes into account the dimension of the sample that was analyzed. In the case of single cells, UV ellipsometry is suitable to determine sample volume. In the case of tissue, the volume only depends on the dimension of the pixel and the thickness of the slice. As an example vol = pixel size $(1.1 \times 1.1 \ \mu m^2) \times$ tissue thickness $(14 \ \mu m) = 16.94 \ \mu m^3$. By the way, in theory, it becomes possible to calculate the concentration of a given chemical species at a spatial resolution that allows determining its biodistribution at the subcellular level. However, several issues must be considered before to interpret such result: $1 - \text{the } 1.1 \times 1.1 \ \mu \text{m}^2$ spatial resolution is still an aberration if we consider the diffraction limit; thus the absorption of a biomolecule in the $1200-900 \text{ cm}^{-1}$ spectral interval still represents a $\approx 10 \times 10 \ \mu m^2$ sample area; 2 — the quality if the curve-fitting procedure is critical and errors in bands positioning or shape may induce major changes in concentration prediction; 3 — there is no alternative method able to provide a concurrent result for validation at such a spatial resolution of untreated samples (maybe UV fluorescence, but the technique is limited in the number of molecular probes utilizable at the same time). Although actually in its development phase, the coherent anti-Stokes Raman spectroscopy (CARS) technique will probably become a strong alternative to FTIR imaging. Thus, cross-validation of FTIR and CARS imaging results should become feasible in the near future.

5. Conclusion

FTIR imaging instrumentation and spectral data treatment methods are now sufficiently developed to allow quantitative applications on biological samples. In the case of single cells, which present a naturally varying morphology, the limit for a quantitative approach is to determine the volume of sample covered by every pixel of the image [16]. The combination with a second technique is required on that point, which should not be destructive for the sample, UV-ellipsometry as an example. However, for tissues, the volume is easy to determine after microtomic sectioning. In this case the roots of error belong to data treatments, curve-fitting procedures as an example, and the lack of alternative method able to confirm the results obtained. Nevertheless, as soon as the methodology is fixed to allow a quantitative approach of biological sample contents by FTIR imaging, this technique will be able to play a role in the panel of modern imaging modalities for the biomedical field.

References

- D. Naumann, D. Helm, H. Labischinski, *Nature* **351**, 81 (1991).
- [2] C. Petibois, G. Cazorla, A. Cassaigne, G. Deleris, *Clin. Chem.* 47, 730 (2001).
- [3] C. Petibois, G. Déléris, Analyst 129, 912 (2004).
- [4] C. Petibois, V. Rigalleau, A.M. Melin, A. Perromat, G. Cazorla, H. Gin, G. Deleris, *Clin. Chem.* 45, 1530 (1999).
- [5] L.H. Kidder, V.F. Kalasinsky, J.L. Luke, I.W. Levin, E.N. Lewis, *Nature Med.* 3, 235 (1997).

- [6] P. Innocenzi, T. Kidchob, J.M. Bertolo, M. Piccinini, M.C. Guidi, C. Marcelli, J. Phys. Chem. B, Condens. Matter Mater. Surf. Interfaces Biophys. 110, 10837 (2006).
- [7] C. Petibois, K. Gionnet, M. Goncalves, A. Perromat, M. Moenner, G. Déléris, Analyst 131, 640 (2006).
- [8] E. Levenson, P. Lerch, M.C. Martin, *Infrared Phys. Technol.* 51, 413 (2008).
- [9] H. Fabian, P. Lasch, D. Naumann, J. Biomed. Opt. 10, 031103 (2005).
- [10] M. Cestelli-Guidi, M. Piccinini, A. Marcelli, A. Nucara, P. Calvani, E. Burattini, J. Opt. Soc. Am. A 22, 2810 (2005).
- [11] R. Bhargava, B.G. Wall, J.L. Koenig, Appl. Spectrosc. 54, 470 (2000).
- [12] B.H. Cohen, D.R. Gold, Cleveland Clin. J. Med. 68, 625 (2001).
- [13] C. Petibois, G. Cazorla, H. Gin, G. Deleris, J. Lab. Clin. Med. 137, 184 (2001).
- [14] C. Petibois, G. Déléris, Trends Biotechnol. 24, 455 (2006).
- [15] D.C. Malins, K.M. Anderson, N.K. Gilman, V.M. Green, E.A. Barker, K.E. Hellstrom, *Proc. Natl. Acad. Sci. USA* **101**, 10721 (2004).
- [16] D.A. Moss, M. Keese, R. Pepperkok, Vibrat. Spectrosc. 38, 185 (2005).