# Biochemical applications of FT-IR spectroscopy

Arthur M. A. Pistorius

*Department of Biochemistry, University of Nijmegen, PO Box 9101, NL-6500 HB Nijmegen, The Netherlands.*

## Introduction

One of the most versatile analytical chemical techniques is infrared (IR) spectroscopy. Since the introduction of Fourier transform (FT) infrared spectroscopy, this technique is now increasingly used besides nuclear magnetic resonance (NMR) spectroscopy for quantitative and qualitative analyses in many diverse applications. Among these are the analysis of pharmaceuticals, biomembranes, biopolymers and microbiological applications.

Although the information-content of an IR spectrum is the same, whether obtained from an older dispersive or a FT-IR spectrometer, exploiting the improved performance of the latter (enhanced frequency accuracy, high signal-to-noise ratios and high data acquisition speed) combined with modern computational methods, has opened up ways to investigate biochemical systems in more detail. Moreover, the development of alternative sampling techniques (reflection measurements, microspectroscopy), facilitates the analysis of less common or small or otherwise "difficult" samples.

In this article, some of the many facets of the infrared analysis of the construction elements of life are presented. The emphasis on proteins and lipids reflects the major application areas of FT-IR spectroscopy in the life sciences.

## Vibrational spectroscopy

Unlike NMR spectroscopy, which probes transitions between nuclear spin states, IR and Raman spectroscopy are used to detect transitions between molecular rotational or vibrational energy levels. The principal difference between these two types of vibrational spectroscopy is that IR spectroscopy detects vibrations, during which the electrical dipole moment changes, while Raman spectroscopy is based on the detection of vibrations, during which the electrical polarisability changes. The change in polarisability can be pictured by realising that, owing to certain vibrations, the volume of the molecule is changed. This results in a change in the electron density distribution of the molecule and hence, a different response to an electrical field is observed. A classical example is given by $CO_2$, a linear, non-polar molecule (Figure 1). Simultaneous stretching of both C=O bonds (symmetric stretching vibration) does not result in a change of the dipole moment (which is zero) and will not result in a band in the IR spectrum, although for this vibration a band in the Raman spectrum is observed (changing molecular volume). Stretching of one C=O bond and simultaneously shortening of the other C=O bond (anti-symmetric stretching vibration) creates a dipole and hence, this vibration is IR-active but will not appear in the Raman spectrum.

Most absorption bands, associated with stretching and bending vibrations in organic molecules, including materials from biological origin, occur at more or less discrete positions in the mid-IR spectral range (roughly between $2.5 \times 10^3$ and $25 \times 10^3$ nm or, in wavenumbers, 4,000–400 $cm^{-1}$). Owing to the fact that the band pattern of the spectrum depends both on the sample composition and its environment, this spectral region can yield much information concerning molecular structure. Some bands which are of particular biochemical importance are the so-called amide I and amide II vibrational bands between 1700 and 1500 $cm^{-1}$, representing the amide backbone of peptides and proteins. Bands around 2900, 1740, 1400 and 1230 $cm^{-1}$ represent (phospho-)lipids

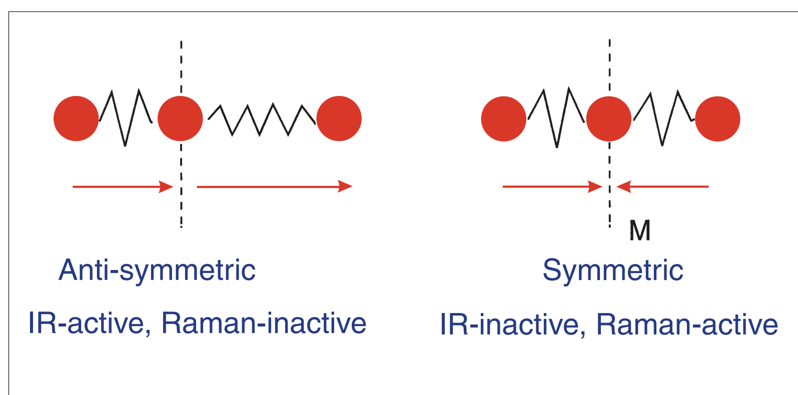


**Figure 1. Stretching vibrations of $CO_2$. Anti-symmetric stretching results in a change in the dipole moment of the molecule and hence, it is IR-active. Due to the preserved symmetry (mirror plane, M, in the right figure) the dipole moment remains unchanged (zero) during the symmetric stretching vibration. Consequently, this vibration is not observed in the IR spectrum of $CO_2$. The arrows indicate the direction of movement of the oxygen atoms at a certain time-point, relative to the central carbon atom.**

 © Spectroscopy Europe 1995

**Table 1. Some infrared vibrational bands of biochemical importance. The given frequencies (wavenumbers) are only indicative of the range in which the mentioned bands occur.**

| Frequency ($cm^{-1}$) | Type of vibration | Mainly observed in |
|---|---|---|
| 2900 | C–H stretching | lipids |
| 2850 | C–H stretching | lipids |
| 1740 | C=O stretching in esters, carboxylic acids | lipids, amino acid side chains |
| 1650 | C=O stretching in amides, (amide I) | proteins |
| 1540 | N–H bending in amides, (amide II) | proteins |
| 1460 | C–H bending | lipids |
| 1400 | C–O stretching in carboxylates | amino acid side chains, lipids |
| 1235 | P=O stretching in phosphate esters | lipids, nucleic acids |
| 1100 | C–O stretching | carbohydrates |

and bands around 1100 and below 1000 $cm^{-1}$ may reflect carbohydrates or nucleic acids (Table 1). The wide variety of bands resulting from polar functional groups, which can be observed in the mid-IR range makes IR spectroscopy the most important vibrational spectroscopic technique for studying biomaterials.

# Peak assignment strategies

IR spectra of bio-macromolecules consist of many bands which, in principle, contain information about structural aspects of the molecule under study. In order to retrieve this information, the observed spectral bands must be assigned to distinct components of the molecule or supramolecular structure (i.e. biomembranes). In many cases, comparison with IR spectra of analogues or smaller model compounds is necessary to assign unequivocally the spectral features. In addition, enzymatic methods give the opportunity to make selective modifications, which certainly will give rise to changes in the IR spectrum, e.g. proteolysis, oxidation/reduction, phosphorylation, isomerisation or conjugation. A more sophisticated method of protein modification is site-directed mutagenesis, either by molecular biological techniques or by (semi-)synthesis. Although these techniques are very time-consuming, relative to the acquisition and analysis of FT-IR spectra, only the combined efforts result in detailed information of structure–function relationships within the protein under study.

The most fundamental way, however, to identify absorption bands in an IR spectrum is to make use of non-radioactive isotope substitution. The vibration, exerted by two atoms may be approximated by a harmonic oscillator. According to classical mechanics, the vibrational frequency, ν, is:

$$\nu = \frac{1}{2\pi}\sqrt{\frac{k}{\mu}}$$

with $k$ a force constant, determined by the chemical bond strength and $\mu$, the reduced mass of the 2 atoms ($\mu = m_1 \times m_2/(m_1 + m_2)$). Assuming that the bond strength, which is an electronic property, does not change, converting a $C{-}^1H$ bond into a $C{-}^2H$ bond results in a $C{-}^2H$ stretching vibration with a frequency which is 1.37 times lower than the corresponding $C{-}^1H$ stretching frequency. An example of isotope labelling by solvent exchange is the estimation of the secondary structure of proteins in aqueous solution. By dissolving the protein in heavy water ($^2H_2O$), the amide hydrogens are exchanged by deuterium atoms. This results in a decrease in the corresponding amide I and amide II frequencies, except for the less accessible trans-membrane α-helices. Another example is the use of perdeuterated lipids to separate the C–H signals from lipid and protein in protein–lipid and lipid–lipid interaction studies. Other substitutions e.g.

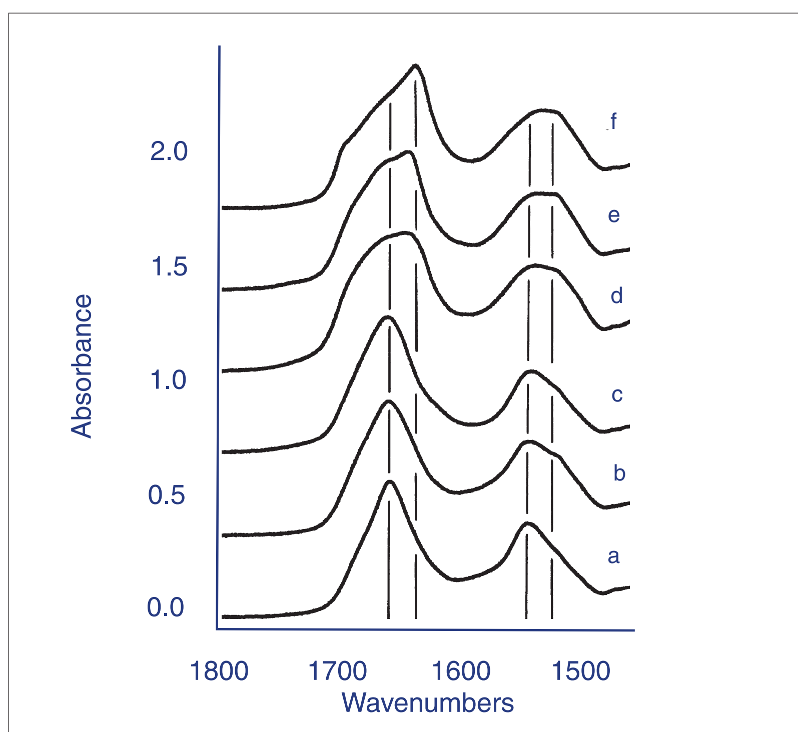


**Figure 2. The secondary structure of proteins determine amide I and amide II envelopes. A set of reference proteins with a known secondary structure, ranging from mainly α-helical (bottom) to mainly β-type structure (top). (a) myoglobin; (b) cytochrome C; (c) lysozyme; (d) carbonic acid anhydrase; (e) trypsin; (f) concanavalin A.**

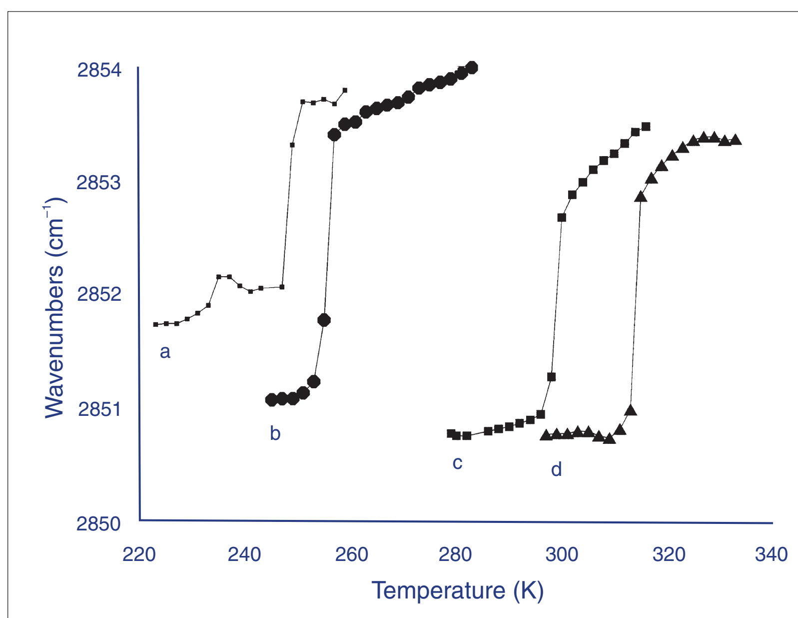


**Figure 3. Typical temperature dependencies of the absorbance maximum of the symmetric stretching band of $CH_2$ groups in liposomes of selected phosphatidylcholines (PC). The discontinuities represent the gel to liquid crystalline phase transition. (a) 16 : 1,16 : 1-PC; (b) 18 : 1,18 : 1-PC; (c) 14 : 0,14 : 0-PC; (d) 16 : 0,16 : 0-PC. (Numbers refer to the fatty acid chainlength and to the number of unsaturated bonds).**

$^{12}C$ to $^{13}C$, $^{14}N$ to $^{15}N$ or $^{16}O$ to $^{18}O$ show less dramatic but equally useful frequency shifts, which allow the IR spectroscopist to utilise as many spectral components as possible.

# Analysis of overlapping bands

An IR spectrum of an acyclic and non-linear molecule, containing a total number of *N* atoms, theoretically consists of 3*N*–6 spectral bands, representing the vibrational degrees of freedom in this molecule. For a spectrum of a typical phospholipid molecule this number already becomes about 400. Not all these modes are separately observed in the spectrum but it is obvious that in spectra of biomacromolecules, overlap of bands seriously impairs the analysis and interpretation of a spectrum. In order to reduce this problem, several mathematical techniques have been adopted to separate overlapping bands, i.e.:

1. derivative spectroscopy
2. Fourier self-deconvolution (FSD)
3. curve fitting.

Derivatisation as well as FSD transform the absorption bands to lineshapes with narrower width, thereby resolving overlapping components. Often these methods are referred to as resolution enhancement techniques, but one has to keep in mind that only the graphical resolution (i.e. the appearance of the spectrum) is improved; the actual optical resolution is still determined by the spectrometer settings.

Curve fitting is the process of regenerating a measured spectrum by mathematically co-adding a pre-defined number of bands with known peak positions (e.g. obtained from deconvoluted spectra), bandwidths and lineshape functions. Subsequently, these parameters are iterated until the theoretically generated spectrum and the measured spectrum coincide to a high degree. In biochemical IR spectroscopy these methods are especially being used in the analysis of amide I band profiles of proteins. Hereto, in the last five years, several methods have been proposed to circumvent the disturbing effect of the absorption band of water around 1650 cm$^{-1}$. Owing to different hydrogen bonding patterns in secondary structural elements, the amide I band consists of several components, representing the relative

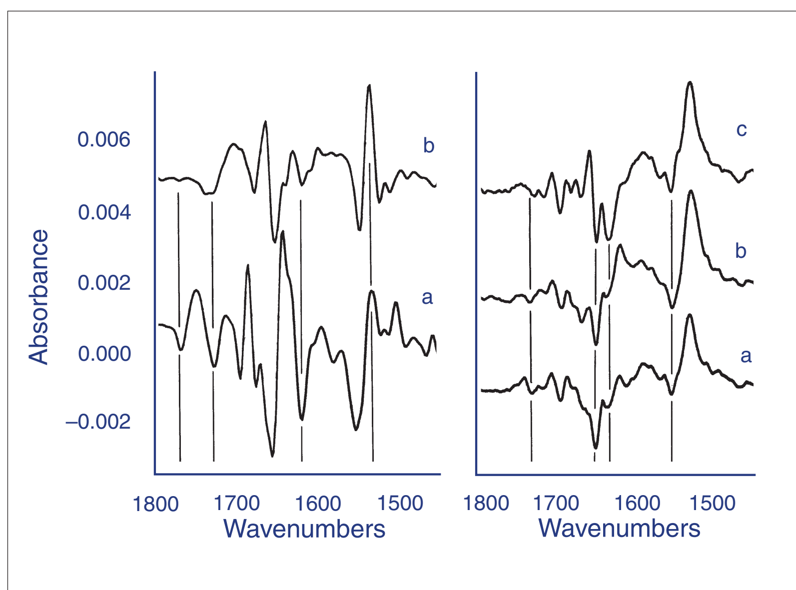


**Figure 4 (left). IR difference spectroscopy can reveal small spectral changes. Difference spectra are obtained by subtraction of the spectrum of the unilluminated photoreceptor membrane from the spectrum obtained after illumination. (a) rhodopsin to metarhodopsin II transition at 10°C; (b) rhodopsin to metarhodopsin I transition at –28°C. (spectra are based on 120 µg protein). (*Courtesy of F. DeLange, Department of Biochemistry, University of Nijmegen*). (right). Detail of isotope-edited IR difference spectra of the rhodopsin to metarhodopsin I transition at room temperature. (a) normal difference spectrum; (b) carbon atom at position 12 of the chromophore replaced by $^{13}C$—no significant effect in this spectral region; (c) carbon atom at position 15 replaced by $^{13}C$ (position of attachment of chromophore to protein) showing shift of C=N band, originally at 1657 cm$^{-1}$ to 1634 cm$^{-1}$, thereby cancelling the positive band at 1626 cm$^{-1}$.**

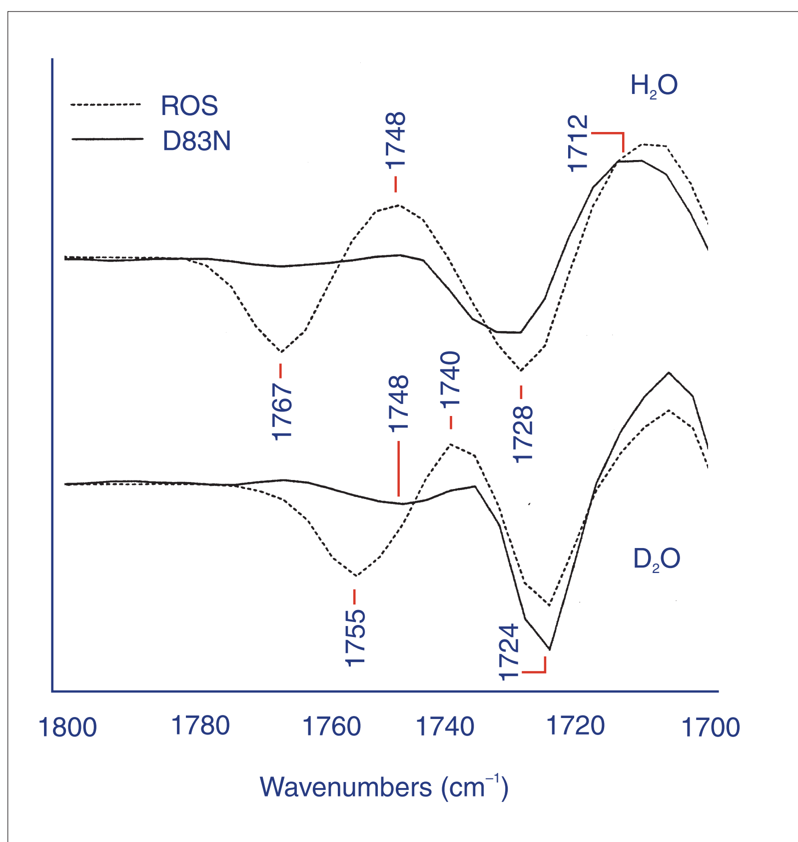


**Figure 5. Detail of the difference spectrum of the rhodopsin to metarhodopsin II transition, showing the effect of the replacement of a single amino acid (aspartic acid at position 83 is replaced by asparagine or a carboxylic acid group is replaced by a primary amide). In an aqueous environment, the signal at 1767 $cm^{-1}$ is abolished by the substitution (upper trace). Repeating the experiment in $^2H_2O$ results in a shift of the complete pattern to lower wavenumbers (lower trace). The dashed line represents the normal difference spectrum and the solid line represents the difference spectrum obtained from the modified photoreceptor. (Reprinted with permission from P. Rath *et al. Biochemistry* 32, 10277–10282 (1993). Copyright 1993 American Chemical Society.)**

amount of e.g. α-helical or β-type secondary structure present. Regrettably, the outcome of these studies is sometimes a matter of debate.[1] By choosing deconvolution parameters unwarily, artifacts are introduced, which can easily be interpreted as "real" spectral components. Moreover, using deconvolution and curve fitting software from different vendors appears to produce different results, which makes comparison and calibration of results from different laboratories quite difficult.

The amide I components in the spectrum of a protein with an unknown secondary structure can also be related to the amide I components of a set of reference spectra, originating from proteins with a known secondary or tertiary structure (Figure 2). This procedure relies on the partial least squares (PLS) method for resolving the relative contributions of α-helical, β-type or turn structures. Although this is an elegant method, requiring no elaborate and time consuming curve fitting procedures, the application is restricted, owing to the, contrary to popular belief, still inadequately understood relationship between the amide bands in an IR spectrum and the secondary structure of a protein. With the ever-increasing amount of protein structural data, it is to be expected that a neuronal network will give the best results, since by training, this approach can interpret more than just a limited number of elementary secondary structures.

## Environmental effects

Another way to study biochemical systems, using IR spectroscopy, is to look at the spectral response to changes in the environment, e.g. acidity, enzyme inhibitors, temperature, influence of light or solvent. Superfusion of an immobilised sample, mounted on an ATR (attenuated total reflection) IR accessory, is a good approach to make controlled changes in the environment.[2] Since in an ATR experiment the IR beam only partially penetrates the sample, spectral analysis is less affected by subtraction artifacts originating from vibrational bands of water: a problem normally encountered when performing IR experiments under physiologically relevant conditions, especially around 1650 $cm^{-1}$ (interference with amide I band).

Biochemical effects of environmental factors are predictable to a certain extent, and therefore, the observed spectral changes may be correlated to certain components of the system. Some examples may illustrate what kind of information can be obtained by FT-IR spectroscopy.

1. The IR spectra of biomembranes show band shifts, as a function of temperature, associated with phospholipids (C–H stretching and bending vibrations, C=O stretching in ester bonds, phosphate diester stretching modes). Figure 3 shows some characteristic examples of melting profiles of phospholipids. Clearly, the frequency of the $CH_2$ stretching vibration and the transition temperature depend on the compostition of the phospholipid with regard to acyl chainlength and presence of double bond(s): factors which determine the fluidity of a membrane. These experiments give information about lipid phase transitions, complementary to DSC (differential scanning calorimetry). IR spectroscopy, however, has the advantage that, in addition to the transition temperature, information is obtained about specific interactions, e.g. binding of proteins to the polar lipid headgroups. This approach is used in biochemical and biomedical studies, employing model membranes (e.g. the working-mechanisms of anaesthetics) as well as in botanical *in vivo* studies (effects of different hydration conditions on intact plant cell membranes).

2. By lowering the temperature, the photolytic cascade of the visual pigments, located in the retina of the vertebrates' eye, can be blocked at several distinct stages. Analysis of IR spectra obtained before and after illumination of functional visual pigments, demonstrate that only small structural changes do occur [Figure 4 (left)]. Earlier studies focused upon the chromophore, 11-

*cis*-retinal, located in the core of the photoreceptor protein, which, in the first stage, isomerises to the all-*trans* state. This reaction forces the receptor to adopt a different structure, eventually leading to the generation of a nerve impulse and the perception of light. For these studies, numerous visual pigments containing a chromophore with isotope substitutions on almost every atom have been analysed by FT-IR difference spectroscopy [Figure 4 (right)]. Currently, the role of selected amino acids in the photoreceptor protein is investigated. Hereto, several mutant visual pigments have been prepared by means of molecular biological techniques. Most intriguingly are the polar residues, located in the transmembrane domains of the photoreceptor, since these domains normally consist of hydrophobic amino acid residues (Figure 5).

In the field of membrane receptor research, FT-IR spectroscopy is demonstrating its qualities at its best. The extremely high signal-to-noise ratios enable the detection of difference signals in the order of 0.0002 absorbance units in a matter of a few minutes. In some cases, these signals can be pin-pointed to a single amino acid residue, out of a few hundred. Using conventional FT-IR instruments one can monitor the time-dependent decay of (photo-)intermediates with a time resolution of milliseconds. With the recently developed step-scan FT-IR spectrometers[3] the time resolution can be even better than microseconds.

Currently, the use of caged agonists is explored. These agonists only bind to a receptor after they are released from their cage by a light pulse. Thus, also non-photosensitive receptors can be triggered by a light pulse, without the need to further manipulate the sample in the spectrometer: a prerequisite for obtaining useful difference spectra.

# Conclusion

FT-IR spectroscopy is increasingly being used for the analysis of biomacromolecules. Although it cannot rival NMR for its structure resolving power, it can give more global structural information about larger biochemical or biologial systems where, owing to line broadening effects, the use of high resolution NMR spectroscopy is not possible. By combining IR spectroscopy with modern molecular biological techniques, detailed structure–function relationships on a local level can be revealed.

## Acknowledgements

The author wishes to thank Patricia Groenen, Feico Schuurmans Stekhoven, Clemens Prinsen, Nico Sommerdijk (*University of Nijmegen*) and Wim Wolkers (*Agricultural University of Wageningen*) for their congenial collaboration during the past six years. The research presented in this paper was sponsored by the Dutch National Foundation for the Advancement of Chemical Research (NWO-SON).

# References

1. W.K. Surewicz, H.H. Mantsch and D. Chapman, *Biochemistry* **32,** 389 (1993).
2. P.R. Griffiths and J.A. DeHaseth, in *Fourier Transform Infrared Spectrometry*. John Wiley & Sons, New York, Ch. 14, p. 457 (1986).
3. W. Uhmann, A. Becker, C. Taran and F. Siebert, *Appl. Spectrosc.* **45,** 390 (1991).

CIRCLE 009 FOR FURTHER INFORMATION