Fourier-Transform Midinfrared Spectroscopy for Analysis and Screening of Liquid Protein Formulations, Part 1

Understanding Infrared Spectroscopy of Proteins

Patrick Garidel and Heidrun Schott

Sir William Herschel discovered infrared light in 1800, over two centuries ago. The first spectra of the molecular vibrations of organic liquids were registered in 1881 by Abney and Festing. Most applications at that point were of academic interest. In the early 1930s, researchers at BASF (Badische Anilin und Soda Fabrik) in Ludwigshafen, Germany, realized the importance of infrared spectroscopy for industrial purposes. In 1937, the first infrared (IR) instrument with a modulated beam was built by Lehrer (**1**). Five years later, the first nondispersive infrared analyzer was presented (**2**).

By 1950, Elliot and Ambrose showed that IR spectroscopy provides information on the secondary structure of proteins (**2, 3**). But the utility of this technique for structural analysis of proteins was not widespread. That changed with the rediscovery of Fourier transform analysis — as described by Jean-Baptiste Joseph de Fourier (**4**) in 1822. It led to the development of Fourier-transform infrared (FTIR) spectroscopy in the early 1980s (**5**).

The transition from IR to FTIR spectroscopy allowed strong improvements in spectral quality. The "Advantages" box summarizes major benefits of FTIR. Additionally, outstanding developments in analytical technology (e.g., mathematical tools for data processing and handling and new possibilities in computer science) enabled significant progress for data analysis, allowing extraction of information on protein conformations from IR spectra (**1, 2**).

Only since the past decade has FTIR spectroscopy become an accepted and powerful technique for development of protein formulations. The relation between protein structure (Figure 1) and bioactivity was long ago recognized. Changes in bioactivity were believed to result from alteration in protein structure and organization. Several methods were developed — and now are commonly used — for determination of secondary and tertiary structure of proteins. For obtaining structural information at high resolution, two methods are mainly used: X-ray crystallography and multidimensional nuclear magnetic resonance (NMR) spectroscopy. However, high-resolution studies of proteins are not always feasible (**6**).

(WWW.PHOTOS.COM)

X-ray crystallography requires well-ordered single crystals of high-quality. For many proteins, that is not possible. And it is unclear whether the relatively "static" structure of a single crystal adequately represents protein conformation in a complex and dynamic liquid environment. Multidimensional NMR spectroscopy presents an alternative to X-ray crystallography, offering somewhat better flexibility in studying the structures of proteins in solution (their natural environment). However, data evaluation and interpretation of NMR spectra are very complex and

PRODUCT FOCUS: RECOMBINANT PROTEINS IN LIQUID FORMULATIONS

PROCESS FOCUS: PRECLINICAL TESTING AND PRODUCT FORMULATION

WHO SHOULD READ: R&D, ANALYTICAL, MANUFACTURING, AND FORMULATION STAFF

KEYWORDS: SPECTROSCOPY, INFRARED, ANALYTICAL METHODS, FORMULATIONS

LEVEL: ADVANCED

Advantages of FT Technical Improvements

Time-Saving: All spectral elements are measured at the same time (Fellgett or multiplex advantage).

Better Signal-to-Noise Ratio: Method provides high optical throughput (Jacquinot advantage).

High Accuracy and Reproducibility: HeNe laser serves as an internal frequency standard (Connes advantage).

No Spectral Artifacts: Fourier transform method strictly excludes other frequencies (e.g., stray light).

Figure 1: Chemical structure of pectine lyase with representations of different secondary structure elements

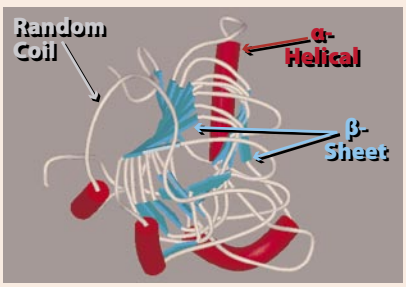
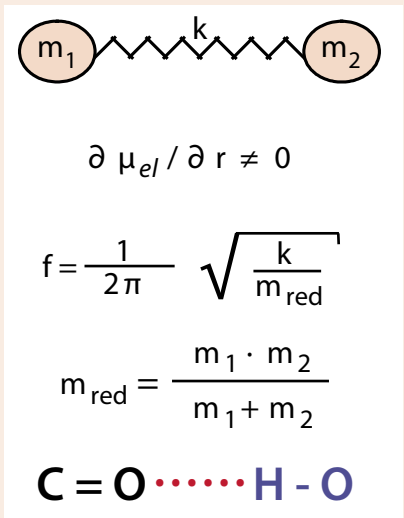


Figure 2: Mechanism of adsorption and the Hooke's law



extremely time consuming, especially for molecules that are larger than ~30 kDa. So NMR applications are limited to small proteins (~15–30 kDa).

Those restrictions have led to development of alternative methods including circular dichroism (CD) and vibrational (infrared and Raman) spectroscopy (**1, 7**). They are often referred as "low-resolution" techniques because they provide only global insight into the overall secondary structure of proteins. They offer no way to establish the precise three-dimensional location of individual structural elements. However, there are a number of advantages to using Fourier-transform infrared (FTIR) spectroscopy.

FTIR measurements can be made using extremely small protein samples (~1 mg/mL) in a number of environments (e.g., solid, semisolid, liquid, or adsorbed to a surface). Here we focus mainly on liquid sample analysis. A major benefit of FTIR is the ease and rapidity of acquiring high-quality spectra (an IR spectrum registers in just a few seconds). This allows a rapid screening of various protein formulations to determine the best conditions for optimal protein stability. Using IR spectroscopy eliminates problems associated with background fluorescence, for example, as detected by Raman or light scattering.

All these facts have made FTIR spectroscopy practical for studying biological systems. By contrast with, for example, electron spin resonance spectroscopy, it does not require additional probe molecules. FTIR is a noninvasive technique. Water absorption, omnipresent especially in the spectral range between 4000 and 400 cm^{-1}, can be subtracted mathematically, which makes study of hydrated (H_2O) biomolecular samples almost custom (**8–11**). Another approach is using D_2O. Mathematical methods are now available and routinely used to separate subcomponent bands that overlap in the spectra of proteins.

The infrared region is subdivided into three regions: near infrared (NIR: 12,500–4000 cm^{-1}), midinfrared (MIR: 4000–400 cm^{-1}), and far infrared (FIR: 400–10 cm^{-1}). The useful unit is the wave number, which is expressed in cm^{-1}. Herein we focus on information derived from MIR spectroscopy.

The Basis of Infrared Spectroscopy

The atoms within a molecule constantly oscillate around an equilibrium position, *r*. Consequently, bond lengths and angles change. In the classical model, two atoms are fixed by a spring, with a spring/force constant, k (Figure 2). The frequency of such motion is within the infrared region, between 1 and 100 µm, and IR radiation of such energy can excite vibrational motions. Thus, infrared spectroscopy is a vibrational spectroscopic method (**8**) that measures the wavelength and intensity of infrared light absorbed by a sample (dry, wet, or gaseous).

The frequency *f* of the vibration between two atoms of mass m_1 and m_2 depends on k and the reduced mass m_{red} (Figure 2 shows the exact relation). Changes in k or m_{red} induce a change in *f*. The equation in Figure 2 describes the frequency of vibrations for a diatomic molecule. Usually the frequency *f* of a vibration is given in Hz (s^{-1}). However, in vibrational spectroscopy it is common to express frequencies in wave number units (waves per unit length), which is the reciprocal of wavelength (**12**).

A linear molecule has $3n - 5$ degrees of vibrational freedom (n = the number of atoms present); a nonlinear molecule has $3n - 6$. For proteins, that can be several thousands, if not more, possible vibrational motions. These motions may not only be excited to the first energy level, but also to higher levels.

However, for three main reasons, not all such excitations are observed individually in a spectrum. As mentioned above and shown in Figure 2, IR radiation is absorbed by valence electrons (bonds) (**1, 2**).

• For infrared absorption, the transition dipole moment (μ_e) must change: $\partial\mu_{el} \div \partial r \neq 0$ (Figure 2). So a molecule such as nitrogen N_2 is not IR active because during its vibrational motion, there is no change in transition dipole moments. By contrast, carbon dioxide (CO_2) is IR active, and an IR spectrum is obtained.

• Transition probability dictates the intensity of an observed IR absorption. If that probability is too small, a transition is too weak to be observed. The transition probability of a vibrational excitation is proportional to the change in dipole moment

during the vibration.

• When vibrations have similar excitation frequencies, they cannot be resolved or identified.

Most IR spectroscopic investigations of proteins use the MIR spectral region. Within this region, first-order excitations of the vibrations are detected for small molecule parts or groups. Those are called *functional group vibrations,* and they include methyl and methylene stretching vibrations as well as stretching of carbonyl bonds. The carbonyl group is a functional part of the amide groups in proteins and peptides. They give rise to well-defined vibrations (well-defined IR bands) and thus are observed in characteristic regions and frequency positions of a spectrum. Figure 3 shows an infrared spectrum of a protein molecule with its absorptions in the spectral region between 4000 and 750 cm^{-1}. Table 1 summarizes the most prominent infrared bands observed for biomolecules such as proteins and lipids (**6, 13–15**). The most important bands for analyzing protein formulations are the amide bands, especially amide I and II.

IR Spectroscopic Experiments

An infrared spectrometer consists of a Michelson interferometer, a light source, a sample chamber, and a detector. The continuum source producing light over a broad range of infrared wavelengths is usually based on silicon carbide (e.g., Nernst Globar). Such IR sources provide high-energy beams at a maximum intensity of 1500–1800 cm^{-1}.

The IR light is split into two beam paths using a half-silvered mirror, then reflected from two mirrors back onto a beam splitter, where both light beams are recombined (a Michelson interferometer). One mirror is fixed, and the second is movable. A laser is used to determine the exact position of the movable mirror. If the distance from the beam splitter to the fixed mirror is not exactly the same as that from the beam splitter to the second mirror, then when the two beams are recombined, there will be a small difference in the phase of their light. Because of the "superposition principle," constructive and destructive interference exists for different wavelengths depending on the relative distances of those two mirrors from the beam splitter. Additionally, the laser used for detecting the mirror position provides a trigger signal for data acquisition.

A data point is collected at each zero crossing — e.g., of the laser interferogram, which corresponds to the quantity measured by the detector (intensity of the combined IR beams as a function of the moving mirror displacement). If the intensity of light is measured and plotted as a function of the movable mirror's position, the resulting graph can be shown to be the Fourier transform of the intensity of light as a function of wave number. In FTIR spectroscopy, the light leaving an interferometer is directed onto a sample, and its intensity is measured using an infrared detector. The intensity of light striking the detector is measured as a function of mirror position. Next, the results are Fourier-transformed to produce a plot of intensity as a function of the wave number.

IR experiments can be performed in transmission or reflection mode (Figures 4 and 5), for which corresponding sample cells are available (**15**). Each sample cell must be a closed chamber that can be rapidly purged with nitrogen or dry air. Both water and CO_2 have strong absorption bands and so must be removed. Water vapor bands especially appear in the spectral region between 1700 and 1600 cm^{-1}, which overlaps with the amide I vibration. In most commercially available IR spectrometers, the sample chamber can hold only one sample. Therefore, background standards and sample spectra are measured sequentially rather in parallel.

Two main types of infrared detectors are used in MIR: mercury cadmium telluride (MCT) and deuterated triglycine sulfate (DTGS) detectors. The quality of a detector depends on its noise level, detectivity (defined as the reciprocal of its noise equivalent power, which describes the incident radiant power for a signal-to-noise ratio of 1 within a given bandwidth of 1 Hz at a given

Table 1: Band assignments of the main infrared active vibrations of biomolecules

Frequency Range (cm^{-1})	Assignment
3490 and 3280	Asymmetric and symmetric H–O–H stretching
3250–3300	Amide A (N–H stretch in resonance with amide II overtone)
3080	Amide B
3010	=C–H stretching of alkenes
2957	Asymmetric CH_3 stretching
2920	Asymmetric CH_2 stretching
2872	Symmetric CH_3 stretching
2851	Symmetric CH_2 stretching
1738	C=O stretch
1600–1700	Amide I (mainly C=O strech)
1645	H–O–H bending
1480–1575	Amide II (N–H bend in plane and C–N stretch)
1468	CH_2 scissoring
~1395	C=O stretch of COO^-
1378	CH_3 symmetric bend
1343	CH_2 wagging
1230–1330	Amide III (N-H bend in plane and C-N stretch)
1240	Asymmetric PO_2^- stretch
1170	Ester C–O asymmetric stretch
1080	Symmetric PO_2^- stretch
1047	C–OP stretch
980	Choline asymmetric stretch
625–770	Amide IV (mainly O=C-N deformation)
720	CH_2 rocking

wavelength), and the spectral detection cut-off (**1, 2**). Another important issue is detector response time, which determines how fast the mirror of a Michaelson interferometer can be moved.

The most common MIR detectors are DTGS detectors, which are very stable but have the drawback of relatively slow response times. MCT detectors offer higher sensitivities. Their response times are about 10 times faster than for DTGS detectors, so spectra can be acquired much faster. The linear detector range is higher for DTGS than for MCT, and DTGS detectors come "ready-to-use," whereas MCT detectors must be cooled with liquid nitrogen to work properly.

Figure 4 shows schematically a sample set-up in transmission mode. The intensity of the beam passing a sample follows the Lambert–Beer law. The transmission T is defined as the ratio of the beam I passing the sample and the initial beam I_0. T spectra can easily be transformed to absorption (A) spectra by the relation shown. This enables IR spectroscopy to be used for determining protein concentration. Attenuated, transmitted IR radiation can be displayed as a function of wavelength, which provides an IR spectrum.

Transmission cells used in MIR spectroscopy typically consist of NaCl or KBr windows for nonaqueous media and CaF_2 or BaF_2 windows for aqueous solutions. Table 2 characterizes the most widely used IR windows, and their properties must be considered in preparing an IR experiment. Analysis of aqueous solutions by MIR spectroscopy is possible using thin sample cells with an optical path length below ~25 µm. This is because increasing the path length renders opaque the region around 1640 cm^{-1} (the bending vibration of water), where the amide I vibration is detected. An alternative is switching from H_2O to D_2O. Commercially available systems for transmission experiments require only very small volume of ~50 µL, with a path length of ~7 µm. Such sample cells come with a temperature control system that enables registration of IR spectra as a function of temperature.

Spectra also can be registered in reflection mode using horizontal attenuated total reflection (hATR) techniques. In ATR, a sample comes into contact with the surface of an internal reflection element (IRE) that has a high refractive index n_1 (Figure 5). Radiation is totally reflected at the boundary between two media of higher (n_1) and lower (n_2) refractive indices (e.g., when a sample is placed on top of the crystal) if it hits this boundary at an incident angle greater than the critical angle

$$\theta_c = \text{arc sin } (n_2/n_1)$$

Reflected radiation energy penetrates the boundary as a so-called *evanescent wave*. Penetration depth d_p is the thickness within which intensity decreases to 1/exp of the intensity at the boundary. As Figure 5 shows, d_p is a function of the refractive indices n_1 and n_2, the incident angle θ, and the wavelength λ. Absorption is a consequence of the IR beam's penetration into a sample. Such hATR set-ups are available for one or multiple path cells, which enhance sensitivity.

Many commercially available hATR systems come with a mounted Peltier element for temperature control. Sample volumes can be as small as 10–20 µL. That is important especially in the preformulation phase of drug development because of a "chronic" lack of material. With hATR techniques, formulation screening for protein stability can be performed at temperature intervals from 0 to 95 °C. Protein concentration can vary between one and at least 100 mg/mL. For investigations of protein spectra, a buffer spectrum must be registered and subtracted from the protein spectrum. ATR set-ups offer several important advantages:

- easy-to-clean ATR crystals (and cells)
- no sample preparation
- consistent path lengths
- useful in investigating both soft powder and liquid samples.

Figure 3: FTIR spectrum of a protein in the spectral region between 4000 and 750 cm^{-1}, showing various regions in which vibrational motions of the amide backbone occur (spectrum taken in the reflection mode)

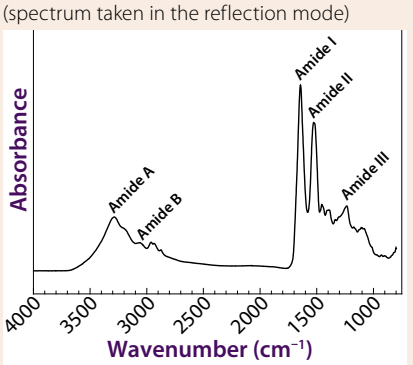


Figure 4: IR experiment in the transmission mode (E = electric field; I = intensity; *f* = frequency; d = optical thickness; μ_e = dipole moment; T = transmission; A = absorption)

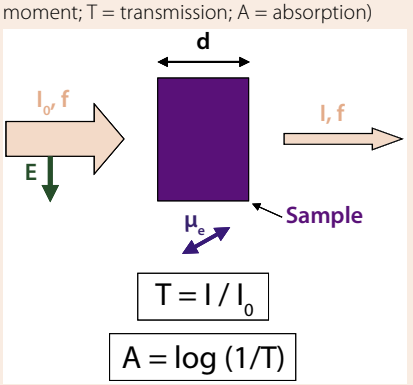
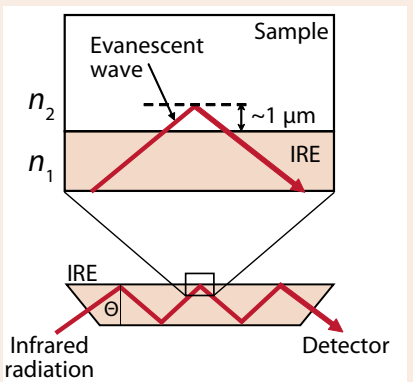


Figure 5: Schematic representation of the ATR set-up in infrared spectroscopy (IR = infrared beam; n_1 = respective index; n_2 = refractive index; IRE = internal reflection element; Θ = angle of incidence; µ = wavelength; dp = penetration depth)



$$d_p = \frac{\lambda}{2\pi n_1 \sqrt{\sin^2 \Theta - n_{21}^2}}$$

$$n_{21} = n_2 / n_1$$

$$n_1 > n_2$$

Table 2: The most widely used infrared window material: some physicochemical properties.

Material	TR	*n*	RL	H	Remarks
NaCl	28,000–700	1.52	~4.5 %	15	Hygroscopic, slightly soluble in alcohol and NH_3
KBr	33,000–400	1.54	~4.5 %	7	Soluble in water, alcohol and glycerine, hygroscopic
AgCl	23,000–400	2.00	~11 %	10	Insoluble in water, soluble in NH_4OH, sensitive to light
CaF_2	66,000–1200	1.40	~2.8 %	158	Insoluble in water, resists most acids and bases, soluble in NH_4 salts
BaF_2	50,000–900	1.45	~3.3 %	82	Low water solubility, soluble in acid and NH_4Cl
Ge	5000–600	4.01	~36 %	550	Insoluble in water, soluble in hot H_2SO_4
ZnSe	20,000–500	2.43	~17 %	150	Soluble in strong acids, dissolves in HNO_3

TR = Transmission range in cm^{-1} — *n* = Refractive index: at 2000 cm^{-1}

RL = Reflectance loss per surface — H = Hardness according to Knopp

Looking for More Detail?

Midrange FTIR spectroscopy is showing promise for use in formulation stability studies, especially in the early preformulation phase, because it enables fast screening of a number of formulations using only small amounts of protein. The protein concentrations used in FTIR range from 1 mg/mL to at least 100 mg/mL.

Next month, Part 2 concludes this article by dealing with the interpretation of protein infrared spectra and presenting various applications and results derived from FTIR screening of protein formulations.

References

1 *Infrared and Raman Spectroscopy.* Schrader B, editor. VCH: Weinheim, Germany, 1995.

2 Günzler H, Heise HM. *IR-Spekroskopie: Eine Einführung.* VCH: Weinheim, Germany, 1996.

3 Elliot A, Ambrose EJ. Structure of Synthetic Polypeptides. *Nature* 165, 1950: 921–922.

4 Fourier JBJ. *Théorie Analytique de la Chaleur.* Paris, France, 1822; reprint available from Editions Jacques Gabay: Paris, France, 1988.

5 Low MJD. Fourier Transform Spectroscopy. *Naturwissenschaften* 75, 1970: 280–287.

6 Winter R, Noll F. *Methoden der Biophysikalischen Chemie.* Teubner-Verlag: Stuttgart, Germany, 1998.

7 Fasman GD, editor. *Circular Dichroism and the Conformational Analysis of Biomolecules.* Springer: Berlin, Germany, 1996.

8 Chalmers JM, Griffiths PR. *Handbook of Vibrational Spectroscopy.* Wiley & Sons, New York, NY, 2002.

9 Dousseau F, Therrien M, Pezolet M. On the Spectral Subtraction of Water from the FTIR Spectra of Aqueous Solutions of Proteins. *Appl. Spectrosc.* 43, 1989: 538–542.

10 *Infrared Spectroscopy of Biomolecules.* Mantsch HH, Chapman D, editors. Wiley-Liss, Inc.: New York, NY, 1996.

11 Smith BC. *Fundamentals of Fourier Transform Infrared Spectroscopy.* CRC Press: Boca Raton, FL, 1996.

12 Griffiths PR, deHaseth JA. *Fourier Transform Infrared Spectrometry.* Wiley InterScience, New York, NY, 1986.

13 Brandenburg K, Seydel U. Infrared Spectroscopy of Glycolipids. *Chem. Phys. Lipids* 96, 1998: 23–40.

14 Hübner W, Blume A. Interactions at the Lipid–Water Interface. *Chem. Phys. Lipids* 96, 1998: 99–123.

15 Garidel P. What Can We Learn from Infrared Spectroscopy for the Development of Protein Formulations? *Formulation and Drug Delivery Strategies for Biopharmaceuticals.* IBC Life Sciences: Munich, Germany, 2004; 1–20.

Corresponding author ***Patrick Garidel*** *(Dr. rer. nat.) is head of pharmaceutical basic development, and* ***Heidrun Schott*** *(Dipl.-Ing. FH) is responsible for the development of infrared techniques in the group pharmaceutical development, Department of Process Science/ Biopharmaceuticals, at Boehringer Ingelheim Pharma GmbH & Co. KG, D-88397 Biberach an der Riss, Germany, +49 7351 54 55 56, fax +49 7351 83 55 56, patrick.garidel@bc.boehringer-ingelheim.com, www.boehringer-ingelheim.com*