Application Note

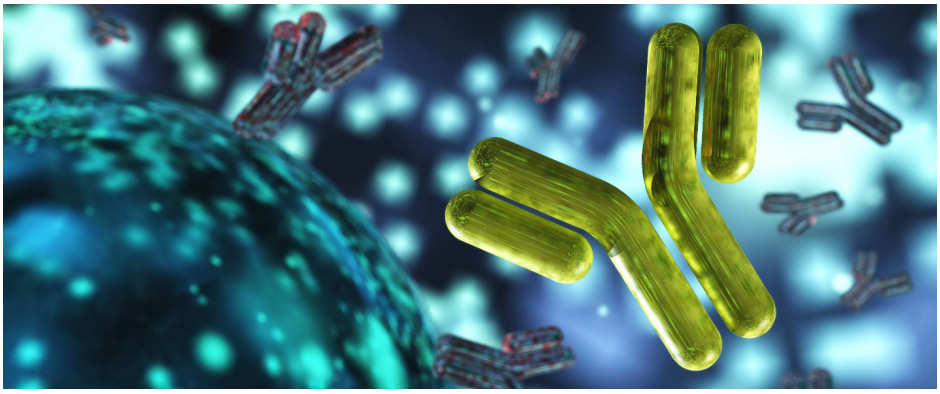
Life Science Research

Agilent

Trusted Answers

Protein Secondary Structure Estimation Using the Agilent Cary 630 FTIR Spectrometer

A rapid and flexible method for evaluating protein secondary structure by FTIR



Authors

Ravindra Gudihal and
Wesam Alwan
Agilent Technologies, Inc.

Abstract

Proteins carry out a wide variety of molecular functions, due in part to their ability to fold into almost unlimited conformations. Thus, assessing the extent of α-helix, β-sheet, turn, and random coil protein secondary structure is essential for understanding protein function. This application note applies Fourier-transform infrared (FTIR) spectroscopy, using the Agilent Cary 630 FTIR spectrometer with attenuated total reflection (ATR) and Agilent MicroLab Expert software, to quickly estimate the secondary structure of proteins in solution. Comparison of FTIR results with literature values derived from X-ray crystallography demonstrates the utility of this approach as a rapid and flexible method for estimating protein secondary structure in solution.

Introduction

Some of the earliest studies of large peptides and protein secondary structure were performed using infrared spectroscopy.[1] Circular dichroism (CD) and X-ray crystallography have also proven essential for assessing protein structure[2,3], but both techniques have drawbacks that limit their use. X-ray analysis requires the protein to be in a crystalline form, necessitating complex and often time-consuming sample preparation, whereas CD can only be applied within a limited optical density range, which limits sample concentration to roughly 0.1 to 1 mg/mL.

FTIR techniques place fewer such restrictions on the sample, and FTIR spectra exhibit absorption bands that carry valuable information about the vibrational modes available to protein backbone amide groups, thus, the secondary structure present.[1] The amide I band frequency is a critical parameter for determining secondary structures. The amide I band has proven to be sensitive and popular for estimating the protein secondary structure. This region of the infrared spectrum, typically found between 1,600 and 1,700 cm^{-1}, is primarily governed by the stretching vibrations of the C=O and C-N groups, which are indicative of the protein's secondary structure. The characteristic frequency ranges for α-helices and β-sheets within this band allow for the assignment of these structures. Though FTIR lacks the high resolution of X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy, it offers a more rapid response, is easy to operate, and is more broadly applicable to secondary structure estimation.

Experimental

Instrumentation

In this study, the Cary 630 FTIR spectrometer with ZnSe optics was used with the single reflection diamond ATR module, which features a short pathlength that allows analysis of relatively highly concentrated protein samples, including viscous samples (Figure 1). Data acquisition was carried out with the MicroLab Expert software using the parameters shown (Table 1).

Figure 1. The Agilent Cary 630 FTIR was used with the single reflection diamond ATR module.

Table 1. Experimental parameters for the Agilent Cary 630 FTIR.

Parameter	Value
Spectral Range	4,000 to 650 cm^{-1}
Sample/Background Scans	140
Resolution	4 cm^{-1}
Zero Fill Factor	None
Apodization	Triangular
Phase Correct	Mertz
Sampling Technology	ATR

Samples: Bovine serum albumin (05470), lysozyme (chicken egg white, L-6876), papain (P4762), β-casein (C6905) and myoglobin (M1882) were purchased from Sigma Aldrich (St. Louis, MO, USA). Monoclonal antibody (mAb) was bought from a local pharmacy distributor in Singapore. The mAb sample was concentrated using Vivaspin 500 centrifugal concentrator spin columns (10 kDa MWCO; Sartorius).

Sample preparation: Bovine serum albumin (BSA), mAb, and papain were used at 10 mg/mL. Lysozyme, β-casein, and myoglobin were used at 25 mg/mL. 10 µL of each protein sample was used for measurement. Fresh ultrapure water was obtained from a Milli-Q Integral system equipped with a 0.22 µm membrane point-of-use cartridge (Merck Millipore). The FTIR spectrum of pure water was subtracted from the FTIR spectrum of the protein solutions, and the resulting spectrum was then analyzed.

Data analysis

MicroLab Expert software (version 1.1.0.1) was used for the data analysis:

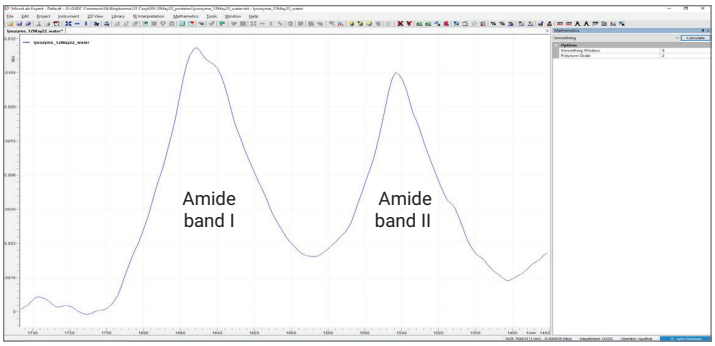
- The averaged spectra of three measurements were used for data processing.
- A blank spectrum (e.g., water or formulation buffer) was subtracted from the averaged spectra using the Subtract Spectra function of the Spectrum Arithmetic under the Mathematics tab.
- The blank subtracted spectra were merged using Merge View of the 2D View tab.
- Further smoothing was done using a smoothing window of 9 and a polynomial order of 3 on the Mathematics tab.
- For the second order derivative spectra, a Savitzky-Golay second-order derivative with a smoothing window of 9 under the Mathematics tab was applied.

Secondary structure estimation using curve fitting procedure with MicroLab Expert Software

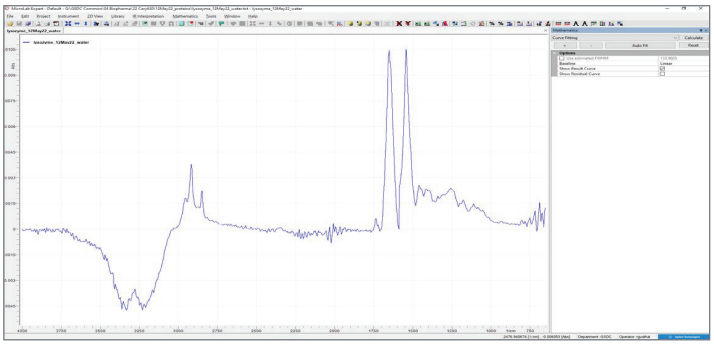
From the Curve Fit menu command of MicroLab Expert Software, the following actions were performed:

1. Perform a baseline correction for amide band I (1,600 to 1,700 cm^{-1}) before applying the curve fit.
2. Activate the 2D data object you would like to fit.
3. From the Mathematics menu, select the **Curve Fit** command.
4. Click the **Plus** button to add a new peak to the list of estimated peaks.
5. Adjust all necessary parameters in the Mathematics tab (FWHM, Maximum X and Y position).
6. Repeat from step 3 until the optimal number of peaks is defined.
7. Click the **Auto Fit** button.
8. Click the **Calculate** button to create a new table with detailed peak information of the estimated peaks.

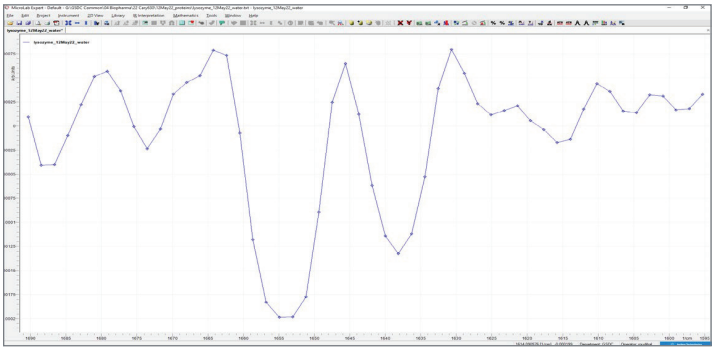
An example of curve fitting using lysozyme is shown in Figure 2.



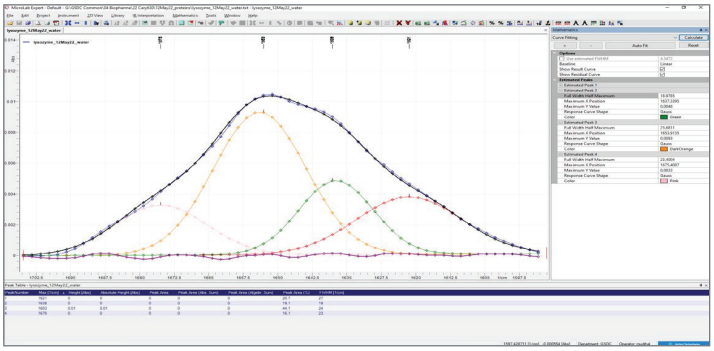
1 Result of subtracting the average lysozyme spectra from the water spectra



2 Perform baseline correction for amide band I (1,600 to 1,700 cm^{-1})



3 Use a Savitzky-Golay filter with a smoothing window of 9 to apply a second-order derivative in order to identify concealed peaks



4 Perform curve fitting on the spectrum obtained in step 2, focusing on the spectral range of 1,600 to 1,700 cm^{-1}

Figure 2. Estimation of protein secondary structure using the curve fitting procedure in Agilent MicroLab Expert software. Component contributions to the secondary structure are estimated by calculating the band area of each component peak.

Component contributions to the protein secondary structure were calculated using the band area of each component peak. Secondary structures were assigned according to Table 2. Due to the different hydrogen bonding interactions between proteins and the different environments in which they exist, all assignments are shown as a range. Please note the "Full width at half maximum" for future use with the same sample.

Table 2. Correlation of secondary structure with amide I band frequency.[4]

Structure	Amide 1 Frequency (cm^{-1})
α-Helix	1,648 to 1,660
Random	1,640 to 1,648
β-Sheet	1,625 to 1,640
Cross-β Structure	1,610 to 1,628
Turns	1,660 to 1,699

A component centered between 1,648 and 1,660 cm^{-1} has been assigned to the α-helix secondary structure. Bands in the region of 1,625 to 1,640 cm^{-1} are usually assigned to the β-sheet structure (Table 2). Random coil conformation exhibits IR bands around 1,643 and 1,649 cm^{-1}.

Results and discussion

The infrared spectra of six native proteins characterized by different known secondary structure: myoglobin (helical), immunoglobulin G (sheet), lysozyme (helical), bovine serum albumin (mostly helical), papain (mixed), and β-casein (disordered) were investigated. Peak fitting was performed on water-subtracted spectra according to the procedure described in the data analysis section (Figure 2). An amide I band provides little insight into the underlying secondary structure components, because the widths of the different component bands are greater than the separations between their peaks. To identify hidden peaks, the second derivative of the spectrum was used as a band narrowing/peak sharpening method. Figure 3 shows an overall curve fitting for six proteins studied. There are several discrete peaks in the amide I region of the spectrum.

The area under the fitted curve in an infrared spectrum, corresponding to the amide I band, is used to assign specific secondary structures, such as α-helices and β-sheets. This assignment is based on characteristic frequency ranges, which are referenced from Table 2. By analyzing the area under the fitted curve in the amide I band, the relative amounts of each secondary structure present in the protein are estimated.

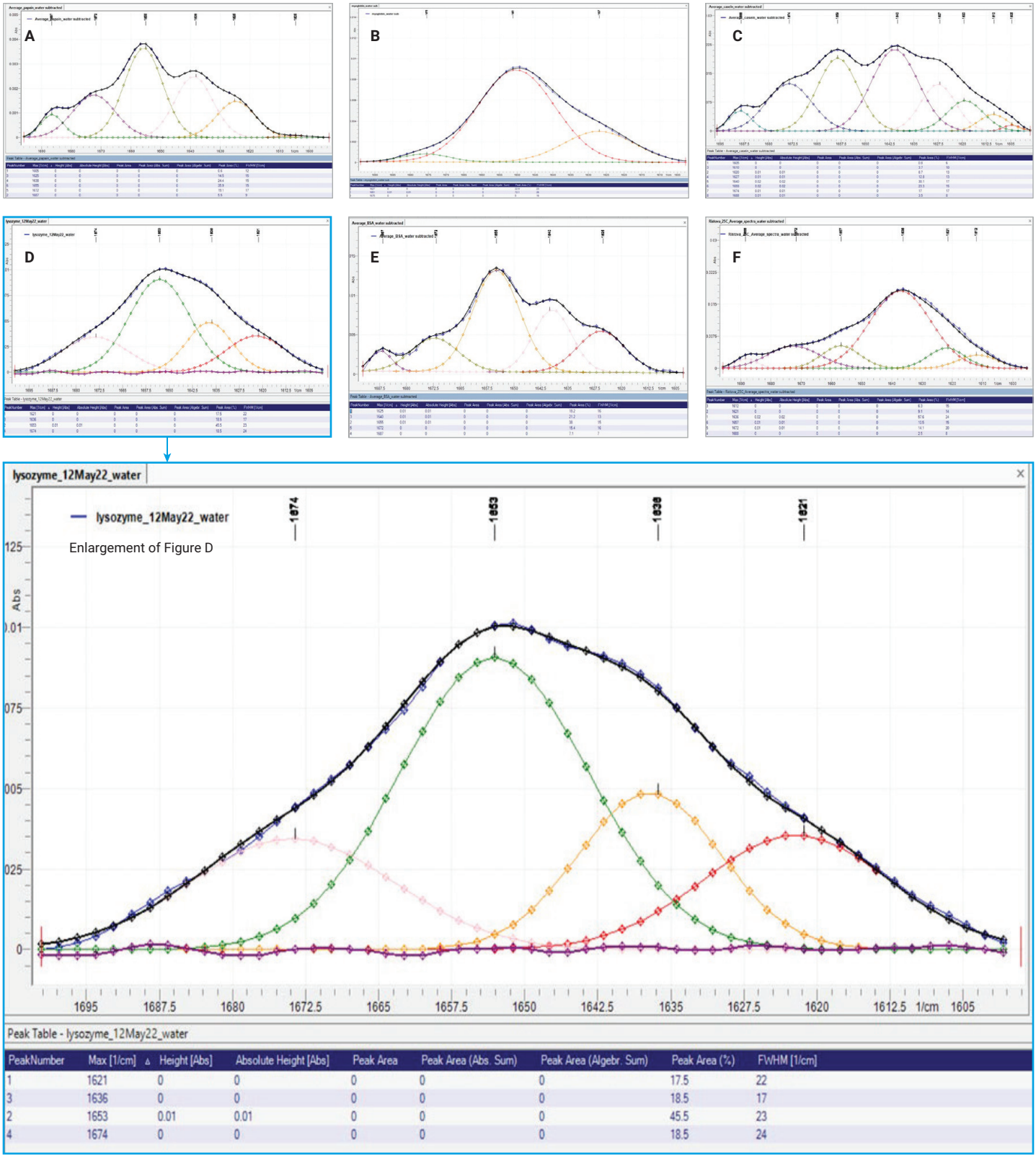


Figure 3. Curve fitting for the six proteins studied: (A) papain (mixed); (B) myoglobin (helical); (C) β-casein (disordered); (D) lysozyme (mixed); (E) bovine serum albumin (mostly helical); (F) immunoglobulin G (sheet). The enlarged lysozyme fitting curve is highlighted in the blue box.

Table 3 summarizes the quantitation results for α-helix, β-sheet, turn, and random coil secondary structures.

Table 3. Relative amount of protein secondary structures estimated by FTIR curve fitting, and comparison with X-ray studies.[5–7]

% Structure								
	α-Helix		β-Sheet		β-Turn		Random Coil	
Proteins	X-Ray	FTIR	X-Ray	FTIR	X-Ray	FTIR	X-Ray	FTIR
BSA	74[7]	38	3.1	18	1.4	23	20.6	21
Papain	25[6]	36	21	39	4	6	50	19
β-Casein	n.d.	23	n.d.	22	n.d.	21	n.d.	34
Lysozyme	45[5]	44	19	19	23	16	13	21*
IGG (Rituximab)	3[5]	10	67	67	18	14	12	9
Myoglobin	85[5]	72	n.d.	23	8	5	7	n.d.

* Assigned to cross-β structure
n.d. = Not definite

The water-subtracted spectra of immunoglobulin G (Figure 1B) exhibited a very strong amide I band maxima near 1,636 cm^{-1}, reflecting their high β-sheet content confirmed by X-ray studies.[5] The peak at about 1,636 cm^{-1} in immunoglobulin G spectra accounted for 67% of the total amide I band area. The water-subtracted infrared spectra of lysozyme showed a relatively narrow amide I band with a maximum near 1,654 cm^{-1}, which is characteristic of proteins with a large content of α-helical secondary structure.[5] As estimated by curve fitting, lysozyme had a distinct α-helical component of 44% at the 1,654 cm^{-1} band position, with 19% β-sheet component. Myoglobin protein showed a predominately α-helical secondary structure, but also a β-sheet component. Proteins such as BSA, papain, and casein are also assigned to different components of the protein secondary structure. The secondary structure of BSA is sensitive to various experimental conditions, which can lead to a variation in its structural distribution. Factors such as temperature, pH, and the presence of surfactants can significantly influence the protein's secondary structure, including its helical content. The interaction with water molecules and the competition for hydrogen bonds can destabilize the helical structure, highlighting the delicate balance of interactions that maintain protein stability. Crystallization can help preserve the protein's native conformation by removing the solvent, which explains why the most helical form of BSA is typically observed in its crystalline state.[7–9]

Table 3 summarizes the quantitation results in terms of percent α-helix, β-sheet, turn, and random secondary structures.

Conclusions

This application note demonstrates a method for the analysis of protein secondary structure in solution using FTIR. The Cary 630 FTIR with single reflection diamond ATR was used due to the short path length and low sample requirements. FTIR analysis required resolving the amide I into its constituent components, achieved here using MicroLab Expert software. A band curve fitting method was used to estimate the relative amounts of key secondary structural features, and results derived from FTIR were found to broadly agree with literature reports.

The ability of the Cary 630 FTIR to study protein structure in a nondestructive manner, combined with its ease of use and minimal sample preparation, suggest potential use for this technique in the rapid evaluation of gross secondary structural characteristics and as a tool for straightforward screening of experimentally induced structural changes.

References

1. Barth, A. Infrared Spectroscopy of Proteins, Biochimica et Biophysica Acta (BBA). *Bioenergetics* **2007**, *1767(9)*, 1073–1101.
2. Greenfield, N. J. Using Circular Dichroism Spectra to Estimate Protein Secondary Structure. *Nat. Protoc.* **2006**, *1(6)*, 2876–2890. doi: 10.1038/nprot.2006.202. PMID: 17406547; PMCID: PMC2728378.
3. Shi, Y. A Glimpse of Structural Biology Through X-Ray Crystallography. *Cell* **2014**, *159(5)*, 995–1014.
4. Jackson, M.; Mantsch, H. H. The Use and Misuse of FTIR Spectroscopy in the Determination of Protein Structure. *Crit. Rev. Biochem. Mol. Biol.* **1995**, *30*, 95–120.
5. Levitt, M.; Greer, J. Automatic Identification of Secondary Structure in Globular Proteins. *J. Mol. Biol.* **1977**, *114*, 181–239.
6. Kamphuis, I. G.; Kalk, K. H.; Swarte, M. B. A.; Drenth, J. Structure of Papain Refined at 1.65 Å Resolution. *J. Mol. Biol.* **1984**, *179*, 233–256.
7. Majorek, K. A.; Porebski, P. J.; Dayal, A.; Zimmerman, M. D.; Jablonska, K.; Stewart, A. J.; Chruszcz, M.; Minor, W. Structural and Immunologic Characterization of Bovine, Horse, and Rabbit Serum Albumins. *Mol. Immunol.* **2012**, *52*, 174–182.
8. Barreto, M. S. C.; Elzinga, E. J.; Alleoni, L. R. F. The Molecular Insights into Protein Adsorption on Hematite Surface Disclosed by In-Situ ATR-FTIR/2D-COS study. *Sci. Rep.* **2020**, *10*, 13441.
9. Abrosimova, K. V.; Shulenina, O. V.; Paston, S. V. FTIR Study of Secondary Structure of Bovine Serum Albumin and Ovalbumin. *Journal of Physics: Conference Series* **2016**, 769.

Further information

- Agilent Cary 630 FTIR Spectrometer
- Agilent MicroLab FTIR Software
- Agilent MicroLab Expert
- FTIR Analysis & Applications Guide
- FTIR Spectroscopy Basics – FAQs

www.agilent.com/chem/cary630

DE-001141

This information is subject to change without notice.

© Agilent Technologies, Inc. 2024

Printed in the USA, October 1, 2024
5994-7808EN

Agilent
Trusted Answers