Norwegian University
of Life Sciences

Master's Thesis 2023 30 ECTS
Faculty of Science and Technology

Investigating subcellular structures in micro-organisms using infrared spectroscopy-based techniques – comparison between the latest micro and nano infrared technologies

Stine Lossius
Environmental Physics and Renewable Energy

Preface

Ås, May 15th 2023

I want to give a huge thanks to my supervisors Achim Kohler and Johanne Heitmann Solheim. I consider myself very lucky to have had both of you as my supervisors, it has been very inspiring and motivating to work with you. I appreciate that you have given me the opportunity to do measurements in Paris which has been a great experience, and I want to thank Ariane Deniset for inviting me to your lab at Paris-Saclay University. I also want to thank all the members of the Biospec group for making me feel so included.

Thanks to my amazing classmates, all our lunches have without a doubt been the highlight of the day during these last months. Thanks to my roommate Mari Valle Kjelby, I will miss having you as my roommate. The five years at NMBU have been amazing and I am truly grateful for all the friendships. Thanks to Nina Zeh for letting me eat your pizza slice five years ago, I knew we would become good friends.

And of course, I want to thank my family for always supporting me and for never being more than a phone call away.

Stine Lossius

Abstract

Infrared spectroscopy is a widely used method to investigate biological samples. The approach enables chemical information of molecules by evaluating the wavenumber dependent absorption of the sample. Infrared microscopic techniques such as Fourier transform infrared microspectroscopy (FTIR) achieve chemically rich spectra with a spatial resolution at the diffraction limit. Nano infrared spectroscopy techniques such as Optical-probed hotothermal induced tnfrated microspectroscopy (O-PTIR) and Atomic Force Microscopy-based infraied spectroscopy (AFM-IR) circumvent the diffraction limit of infrared radiation, and the techniques are promising for obtaining high quality subcellular spatially resolved spectra.

The aim of this thesis is to compare FTIR microscopy and nano infrared techniques for their ability to achieve subcellular resolved spectra in the analysis of biological cells.

For FTIR microscopy, the newly developed approach of deep-learning empowered 3D infrared diffraction tomography is used. The method takes advantage of the scattering features that arise in infrared microspectroscopy. Scattering effects change the absorbance spectra of biological cells considerably, making it hard to analyze the chemical signatures in the sample. Thus, a lot of effort has been made to remove the scattering effect, such as extended multiplicative signal correction, and further by deep learning algorithms, which are less computationally expensive. Instead of looking at the scattering as a disturbing effect, deep-learning empowered 3D infrared diffraction tomography solves the inverse scattering problem and retrieves the pure absorbance spectrum of the cell wall and cell interior of biological cells. This is possible by taking advantage of scattering chemical features that are characteristic for the chemical and physical properties of the sample. The approach makes it possible to obtain subcellular information from structures of biological samples that have sizes below the diffraction limit of microspectroscopic instruments such as FTIR. We show in this thesis that with intact cells with a diameter size bigger than 5 µm such as *Phaffia rhodozyma* it is possible to get high chemical absorbance spectra of the cell wall and cell interior using the deep-learning empowered 3D infrared diffraction tomography.

O-PTIR is theoretically limited by a green laser of 532 nm and the AFM-IR is only limited by the tip radius of the cantilever, giving a spatial resolution down to 50 nm. In order to obtain chemical information from the cell wall and cell interior, biological samples were sectioned. The cells were embedded in epoxy, which is necessary to make nanometer sections. The epoxy shows strong signals in the infrared absorbance spectra, leading to highly contaminated spectra. We show that it is impossible to correct for the epoxy and to obtain chemical information from the cell wall and cell interior from O-PTIR spectra. The AFM-IR sectioned cells were also strongly affected by epoxy, especially in the region between 1200 cm^{-1} and 1000 cm^{-1}. However, we were successfully able to distinguish the cell wall and cell interior due to chemical differences in the spectra.

In a recent study, in-depth measurements of intact biological cells were performed with O-PTIR. The study claims that it is possible to obtain in-depth chemical information about intact cells. However, there are a lot of uncertainties with this technique. With a limited understanding of the optical instrumentation, it is impossible to be sure that the chemical information obtained by in-depth measurements originates from that focus point. Most likely, the signal comes from the whole laser path of the probe laser and not only from the measured depth of the sample. Therefore, it is not possible to use this approach to obtain subcellular chemical information. However, the technique allows high-quality signals from different depths, giving an approximation of the thickness of the cell.

Sammendrag

Infrarød spektroskopi er en mye brukt metode for å undersøke biologiske prøver. Tiln¬ærmingen muliggjør kjemisk informasjon av molekyler ved å evaluere absorpsjonen av prøven som en funksjon av bølgetall. Infrarøde mikroskopiske teknikker som *Fourier transformerer infrarød* (FTIR) mikrospektroskopi oppnår kjemisk rike spektra med en romlig oppløsning ved diffraksjon grensen. Nano infrarød spektroskopi teknikker som *Optical-probed PhotoThermal induced InfraRed microspectroscopy* (O-PTIR) og *Atomic Force Microscopy-based InfraRed spectroscop* (AFM-IR) omgår diffraksjonsgrensen for infrarød stråling, og teknikkene er lovende for å oppnå høykvalitets subcellulære romlig oppløste spektra.

Målet med denne avhandlingen er å sammenligne FTIR-mikroskopi og nano-infrarøde teknikker og deres evne til å oppnå subcellulære oppløste spektra i analysen av biologiske celler. For FTIR-mikroskopi, den nyutviklede *deep-learning empowered 3D infrared diffraction tomography* metoden vil bli brukt. Metodene utnytter spredningseffektene som oppstår i infrarød mikrospektroskopi. Spredningseffekter endrer absorbansspektrumet til biologiske celler betydelig, noe som gjør det vanskelig å analysere de kjemiske signaturene i prøven. Dermed har det blitt gjort mye arbeid for å fjerne spredningseffekten, som metoden *Extended multiplicative signal correction* og nye metoder ved bruk av dyp læring, som er mindre tidskrevende. I stedet for å se på spredningen som en negativ effekt, løser *deep-learning empowered 3D infrared diffraction tomography* det inverse spredningsproblemet som resulterer i rene absorbansspektrum for celleveggen og innsiden av cellen til biologiske celler. Dette er mulig ved å utnytte spredningseffektene som er karakteristiske for de kjemiske og fysiske egenskapene til prøven. Tilnærmingen gjør det mulig å oppnå subcellulær informasjon fra strukturer av biologiske prøver som har størrelser under diffraksjonsgrensen for mikrospektroskopiske instrumenter som FTIR. Vi viser i denne oppgaven at med intakte celler med en diameterstørrelse større enn 5 µm som Phaffia rhodozyma, at det er mulig å få høy kjemisk absorbansspektra av cellevegg og av innsiden av cellen ved hjelp av *deep-learning empowered 3D infrared diffraction tomography*.

O-PTIR er teoretisk begrenset av en grønn laser på 532 nm og AFM-IR er bare begrenset av radiusen til tuppen av cantileveren, og som gir en romlig oppløsning ned til 50 nm. For å oppnå kjemisk informasjon fra cellevegg og innsiden av cellen ble de biologiske prøvene snittet. Cellene var injisert med epoksy, som er nødvendig for å lage snitter med tykkelse på nanometer. Epoksyen viser sterke signaler i det infrarøde absorbansspektrene, noe som fører til svært forurensede spektra. Vi viser at det er umulig å korrigere for epoxyen og å skaffe kjemisk informasjon fra celleveggen og innsiden av cellen fra O-PTIR-spektra. AFM-IR-snittet celler var også sterkt påvirket av epoksy, spesielt i området mellom 1200 cm-1 og 1000 cm-1. Vi var derimot i stand til å skille cellevegg og innsiden av cellen på grunn av kjemiske forskjeller i spektrene.

I en nylig studie ble det utført dyptgående målinger av intakte biologiske celler med O-PTIR. Studien hevder at det er mulig å få kjemisk informasjon av ulike dybder for intakte celler. Det er imidlertid mange usikkerheter med denne metoden. Med en begrenset forståelse av den optiske instrumenteringen, er det umulig å være sikker på at den kjemiske informasjonen som ble oppnådd ved ulike dybdemålinger, stammer fra dette fokuspunktet. Mest sannsynlig kommer signalet fra hele laserbanen til sondelaseren og ikke bare fra den målte dybden av prøven. Derfor er det ikke mulig å bruke denne tilnærmingen for å oppnå subcellulær kjemisk informasjon. På en annen side tillater metoden høy kvalitet på spektra fra forskjellige dybder, noe som gir en tilnærming på tykkelsen til cellen.

Contents

Preface iii

Abstract v

Sammendrag vii

1 Introduction 1

2 Theory 5
- 2.1 Infrared spectroscopy 5
- 2.2 The definition of the pure- and apparent absorbance spectrum 7
- 2.3 Spatial resolution 10
- 2.4 Fourier transformed infrared spectroscopy 10
  - 2.4.1 Michelson interferometer and Fourier transform 10
- 2.5 The photothermal effect based techniques 12
  - 2.5.1 The photothermal effect 12
  - 2.5.2 Optical-probed photothermal induced infrared microspectroscopy 14
  - 2.5.3 Atomic force microscopy-based infrared spectroscopy 15
- 2.6 Spectral preprocessing and data analysis 16
  - 2.6.1 The multiplicative signal correction 16
  - 2.6.2 The extended multiplicative signal correction 17
  - 2.6.3 Principal component analysis 19
  - 2.6.4 Multivariate curve resolution 19
  - 2.6.5 3D infrared diffraction tomography 20

3 Method 23
- 3.1 Biological samples 23
- 3.2 Instrumentation 23
- 3.3 Spectral preprocessing and data analysis 26

4 Results and discussion 31
- 4.1 Spatial resolution of O-PTIR 31
- 4.2 Sectioned cells measured with O-PTIR 31
  - 4.2.1 Removing epoxy from the sectioned cells 33
  - 4.2.2 Sectioned cells measured with AFM-IR 36
- 4.3 depth-resolved O-PTIR measurements of intact cells 39
- 4.4 3D infrared diffraction tomography 48

5 Conclusions and Outlook 55

Bibliography 59

Chapter 1

Introduction

Infra**R**ed (IR) spectroscopy is a frequently used tool for scientists in a wide range of areas, such as biology, medicine and material science [23]. The technique enables detection of the chemical composition of materials in the form of an interpretable chemical fingerprint of the specific sample. An important milestone in the history of infrared spectroscopy was in the 1960s decade, when the Michelson interferometer from 1891 was combined with the **F**ast **F**ourier **T**ransform (FFT) algorithm from 1965 made by Cooley and Tukey [29] [7]. As a result of this combination, **F**ourier **T**ransform **I**nfra**R**ed (FTIR) Spectroscopy was developed. Digilab FTS 14 developed in 1969 the first commercial FTIR instrument, and paved the way for modern FTIR spectroscopy as we know it today [14]. With the technology being commercialized, it was easily accessible to the scientists. In 1975 the Nicolet Instrument Corporation invented the Nicolet model 17199 FTIR spectrometer. The competition between Nicolet and Digilab resulted in the development of the reliable, versatile and sensitive instrument that exists today. The next big milestone of infrared spectroscopy was in the 1980s when the infrared microscope was developed, where an FTIR spectrometer was combined with a microscope with adapted optics to obtain infrared spectra spatially resolved. Combining FTIR spectroscopy and microscopy provides both chemical and physical information from the same sample position. Infrared images can either be obtained by scanning sample areas with an infrared beam that is focused by an aperture and where radiation is collected by a single element detector or by an imaging system, where hyperspectral images are recorded with a **F**ocal **P**lane **A**rray (FPA) detector [5]. Today there are two types of infrared imaging systems available. Systems based on a Fourier transform spectrometer and systems operated with tunable lasers [5].

Infrared images are information rich. The possibility of obtaining infrared spectra spatially resolved provides information about the spatial distribution of chemical components. Further, physical information, such as size and shape, are obtained in high-quality visual images that can be obtained for the same regions for which infrared spectra are acquired [42] [39]. A great advantage of modern spectroscopy such as FTIR spectroscopy is that the technique is nondestructive, meaning that the sample is not fragmented or destroyed as it is necessary with conventional chemical analysis. Further, FTIR spectroscopy provides high spectral resolution, 2-10 cm^{-1}. The spatial resolution of infrared microscopy is restricted by the diffraction limit which is in the same magnitude as the wavelength of the incoming radiation. Mid-infrared radiation has wavelengths between 2.5 µm and 25 µm, which is also the order of sizes of biological cells. Thus, with state of the art infrared techniques, it is in general very hard to resolve subcellular components.

The IR absorbance spectrum is obtained by detecting the attenuation of the incoming IR light intensity. Mainly the attenuation is due to chemical absorption of the sample, however, it can also be due to scattering. When the attenuation is only due to chemical absorption, the absorbance spectrum is referred to as the pure absorbance spectrum. However, if the loss of incoming radiation is also due to scattering the absorbance spectrum is referred to as the apparent absorbance spectrum. Absorbance spectra from biological cells and tissue suffer from distinct physical effects such as Mie scattering, which was first identified in IR spectra in 2005 by Mohlenhoff et al. [30]. Hence, the absorbance spectra are not only showing chemical absorption features but also physical features. In order to correctly read the chemical contributions in the spectra it is necessary to preprocess the data. A frequently used model-based preprocessing technique is **E**xtended **M**ultiplicative **S**ignal **C**orrection (EMSC) [1]. The EMSC model is able to correct for physical features, such as shifts in the baseline, and to separate and quantify the chemical and physical variation in the spectra. A lot of effort has been made to successfully correct for Mie scattering, and with the state of the art algorithm, **M**ie **E**xtinction **E**xtended **M**ultiplicative **S**ignal **C**orrection (ME-EMSC), it is possible to recover the pure absorbance spectra of microscopic samples. However, the algorithm is computationally expensive and the development of faster algorithms has been achieved using deep learning by Magnussen et al. [23]. In recent years, this development has gone one step further and taken advantage of the physical scattering effects to resolve subcellular information. Both physical and chemical features are determined by the complex refractive index and the absorptivity of the sample. Thus, by solving the inverse scattering problem, it is possible to obtain 3D optical, structural and chemical information from the FTIR spectroscopic measurements of intact biological cells. By training a deep convolutional **D**e**S**cattering **A**uto**E**ncoder (DSAE) it is possible to predict the size of the cell wall and the cell interior of an intact cell [24]. In addition, the algorithm can retrieve the chemical composition of the cell wall and cell interior and predict the refractive index of the sample. Thus, by taking advantage of the physical effects it is possible to obtain information rich spectra at the diffraction limit. The approach is called 3D IR diffraction tomography.

To circumvent the diffraction limitation due to the infrared wavelength, new infrared spectroscopic instrumentation that probes the near field has been invented: **O**ptical-probed **P**hoto-**T**hermal induced **I**nfra**R**ed microspectroscopy (O-PTIR) and **A**tomic **F**orce **M**icroscopy-based **I**nfra**R**ed spectroscopy (AFM-IR). Both instrumentations take advantage of the Photothermal process that occurs when an infrared laser beam is focused onto the sample in a pulsed duration. As late as the 1980s and the 1990s the first experiments of PTIR were notified, hence the technology is quite new and still under development [13] [15]. O-PTIR uses both an IR QCL and a green visible laser which are made collinear and guided onto the sample. The IR laser is absorbed by the sample and causes a photothermal effect. As a result of the absorption of radiation from the IR laser, the reflectivity of the green laser changes. It can be shown that the change in reflectivity is proportional to the absorption of IR radiation. By using the green laser as the probing laser, the O-PTIR system is no longer limited to the diffraction limit of the IR radiation wavelength but of the green laser wavelength which is in the sub-micrometer range [33].

The AFM-IR based method was invented by A. Dazzi and it was patented in 2007. An infrared laser is focused onto the sample causing a temperature change which leads to a rapid expansion of the sample. Further, the expansion causes a deflection of the cantilever which starts to oscillate. It has been shown that the oscillation is proportional to the IR absorption [8]. The spatial resolution of the AFM-IR is limited by the tip radius of the cantilever, few tens of nanometers [32] [6]. Highly detailed topographical images at the nanometer scale can be achieved, which has opened the door to subcellular physical and chemical resolution.

Different sample preparation methods for O-PTIR and AFM may be used with the goal of resolving subcellular chemical information. One approach is by micro-sectioning cells to obtain infrared spectra of the cell wall and cell interior. Another approach that has been done in a study by Zhang et al. [44] is to probe in-depth measurements of intact cells with O-PTIR. The study claims that they successfully obtained depth-resolved chemical information of biological cells.

The thesis is a methodological study where the overall aim is to compare different micro-

and nano spectroscopic instrumentation and evaluate how well it is possible to obtain subcellular spatially resolved chemical information of biological cells. Three different infrared techniques are compared, i.e. O-PTIR, AFM-IR and FTIR with a FPA detector where the 3D infrared diffraction tomography approach by Magnussen et al. [24] is used. Further, two different preparation methods for biological cells are considered, namely sections of embedded cells and intact cells. Cells sections are measured with O-PTIR and AFM-IR with the goal of distinguishing the cell wall from the cell interior. The second preparation method used, is the use of intact cells which are measured by FTIR FPA and where the chemical pure absorbance spectra of the cell wall and cell interior are reconstructed by using a pre-trained algorithm from the approach of Magnussen et al. [24]. In addition, depth-resolved spectra of intact cells are measured with O-PTIR, similar to the method described in the study by Zhang et al. [44]. The biological samples considered are filamentous fungus *Mucor circinelloides*, two different yeasts strains (*Rhodotorula graminis*, *Phaffia rhodozyma*) and one type of algae (*Aurautiochyrium limaanium*).

Chapter 2

Theory

2.1 Infrared spectroscopy

We will compare O-PTIR, AFM-IR and 3D infrared diffraction tomography which are all different configurations using infrared spectroscopy, therefore we consider now the fundamentals of infrared spectroscopy.

Spectroscopy is the study of electromagnetic radiation and how it interacts with matter [37]. The interaction with light and matter includes absorption, transmission and scattering of radiation by the material. Electromagnetic radiation consists of a whole spectrum of different frequencies and wavelengths as shown in Fig. 2.1. This thesis is limited to the infrared (IR) part of the electromagnetic spectrum, which is in the region between microwaves and visible light. Further, IR radiation can be divided into three regions, i.e **F**ar-**I**nfra**R**ed (FIR), **mid**-**I**nfra**R**ed (mid-IR) and **N**ear-**I**nfra**R**ed (NIR), with wavelengths between 500-25 μm, 25-2.5 μm and 2.5-0.8 μm, respectively. A common practice in IR spectroscopy is to use wavenumber ($\tilde{\nu}$) instead of wavelength, which is defined as $\frac{1}{\lambda}$, where λ is the wavelength. Wavenumber has the unit cm^{-1}, and Mid-IR spectroscopy lies in the spectral range between 400- 4000 cm^{-1}, also illustrated in Fig. 2.1. More precisely this thesis focuses on the mid-IR region of the IR spectrum.

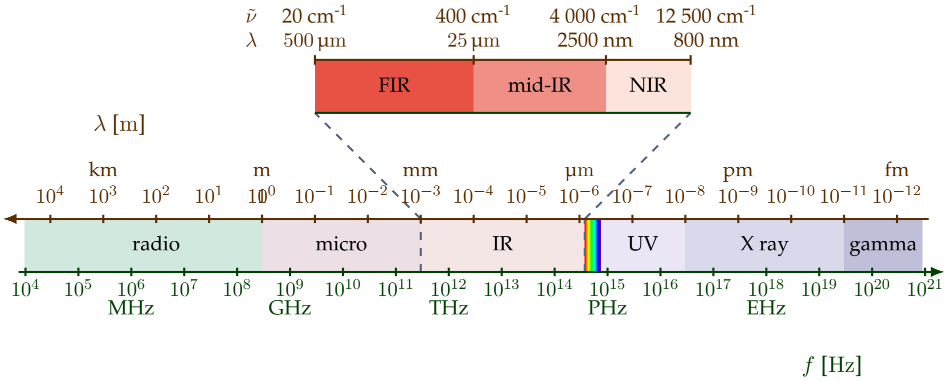

$\tilde{\nu}$ 20 cm^{-1}
λ 500 µm
400 cm^{-1}
25 µm
4 000 cm^{-1}
2500 nm
12 500 cm^{-1}
800 nm
FIR
mid-IR
NIR
λ [m]
km
m
mm
µm
pm
fm
10^4 10^3 10^2 10^1 10^0 10^{-1} 10^{-2} 10^{-3} 10^{-4} 10^{-5} 10^{-6} 10^{-7} 10^{-8} 10^{-9} 10^{-10} 10^{-11} 10^{-12}
radio
micro
IR
UV
X ray
gamma
10^4 10^5 10^6 10^7 10^8 10^9 10^{10} 10^{11} 10^{12} 10^{13} 10^{14} 10^{15} 10^{16} 10^{17} 10^{18} 10^{19} 10^{20} 10^{21}
MHz
GHz
THz
PHz
EHz
f [Hz]

FIGURE 2.1: The electromagnetic spectrum with the different regions in separate colors. The green and red arrows denote the frequency f [Hz] and wavelength λ [m], respectively. A zoomed illustration of the IR region showing the three regions, far-infrared, mid-infrared and near-infrared with the respective wavelengths and wavenumbers $\tilde{v} = \frac{1}{\lambda}$. By courtesy of Johanne Heitmann Solheim, Faculty of Science and Technology, NMBU.

Today, a broad range of IR spectroscopy techniques exists. Fourier Transformed Infrared (FTIR) Spectroscopy is the state of the art technique within IR instrumentation. FTIR uses the mid-IR and NIR parts of the electromagnetic spectrum. With this technique it is possible to calculate a spectrum in the frequency domain from an interferogram measured in the time domain, using the Fourier transform. Infrared spectra make it possible to identify chemical substances in intact sample materials and to obtain physical information about the sample. Recently new methods and instrumentations for measuring the IR absorbance spectra with nanometer resolution have been developed, such as O-PTIR and AFM-IR [8]. Common for spectroscopic techniques is that they all are non-destructive methods, which means that the IR radiation does not fragment the sample. Hence, for spectroscopy, in general, a destruction of the sample is not necessary and chemical structures can be analyzed in their native form, contrary to chemical and biochemical analysis techniques.

Absorption of radiation by a material is a quantum effect requiring that molecules have discrete energy levels and by absorption of electromagnetic radiation, the molecules change their state from lower energy levels to higher energy levels. Incident radiation does not have enough energy to excite electrons in atoms from lower to higher energy levels, but the energy of the radiation is sufficient to excite the intrinsic chemical vibrational modes of the molecule. The vibrational modes of complex samples result in spectral fingerprint that is characteristic for the samples. Vibrational modes can be classified into two different motions, i.e. bending and stretching. Further, bending vibration can be divided into rocking, twisting, scissoring and wagging. Stretching can also be divided into symmetric- and asymmetric stretching, depending on whether the motion is in phase or out of phase. The excited intrinsic vibrational modes result in characteristic peaks in the IR absorbance spectrum which can be interpreted as a fingerprint of a molecule. The fingerprint makes it possible to identify different molecules because every molecule has fairly different vibrational modes, which results in different absorbance spectra [35]. IR spectroscopy is used in a wide range of biological science and life science. Biological material, such as cells and tissue, consists of the main components proteins, lipids and carbohydrates. Thus, biological materials all have similar main components and only small features are distinguishing them apart. In section 3 and in Fig. 3.5 spectra of *Mucor circinelloides*, *rhodotorula graminis* and *phaffia rhodozyma* measured with **H**igh **T**hroughput **S**creening (HTS) FTIR are shown.

When absorbance spectra are recorded it is important to distinguish between chemical and physical effects. Figure 2.2 illustrates how absorbance spectra are recorded. First, as shown in Fig. 2.2A, a background spectrum is recorded, where the incident radiation coming from the source with intensity $\tilde{I}$ propagates through the optical system towards an empty sample holder. The transmitted radiation passing the empty sample holder is recorded at the detector. The background intensity is denoted by I_0. In Fig. 2.2B it is illustrated how the sample intensity is obtained. The sample is mounted on the sample holder and the intensity that transmits through the sample is measured at the detector. Due to the conservation of energy, the intensity will be either absorbed, transmitted and/or scattered. The following equation shows this relationship

$$\tilde{I}_0 = I_A + I_T + I_{sca}. \tag{2.1}$$

where I_A, I_T and I_{sca} are the absorbed, transmitted and scattered intensity, respectively. In ideal cases, scattering is not taken into account and the incident light is either absorbed or transmitted through the sample. The transmission of a sample can then be expressed as the ratio of the intensity reaching the detector and the incoming intensity as follows

$$T(\tilde{\nu}) = \frac{\tilde{I}(\tilde{\nu})}{I_0(\tilde{\nu})}. \tag{2.2}$$

A common practice is to represent the spectra according to absorbance and not transmission, this is because the absorbance is proportional to the concentration of constituents according to Beer-Lambert's law [1]. The relationship between transmission and absorbance $Z(\tilde{\nu})$ is as following

$$Z(\tilde{\nu}) = -\log T(\tilde{\nu}). \tag{2.3}$$

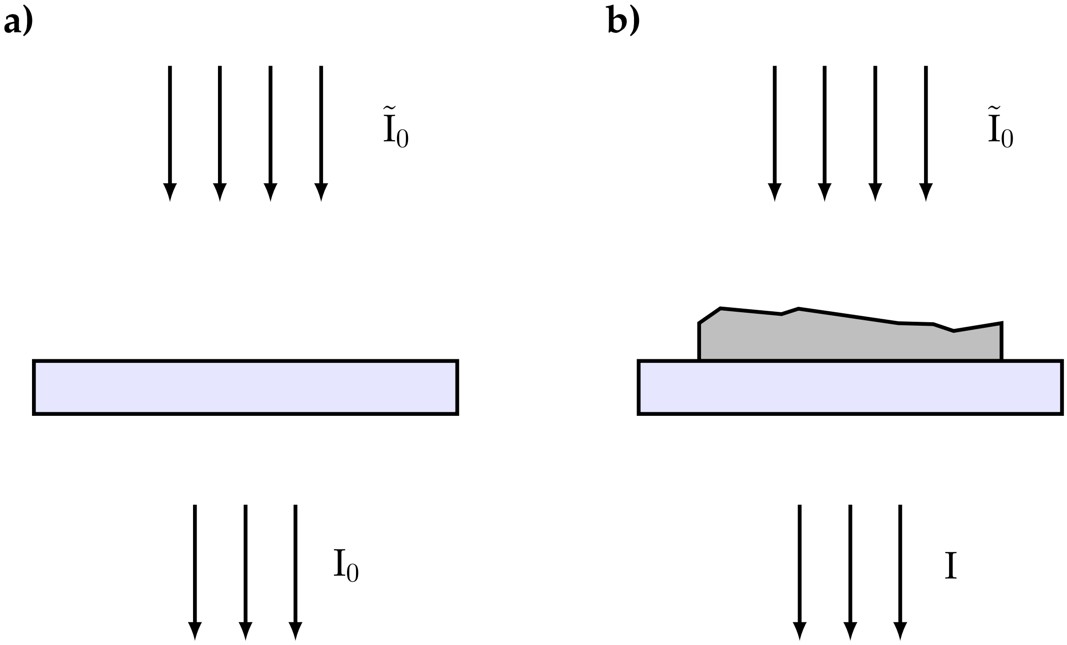
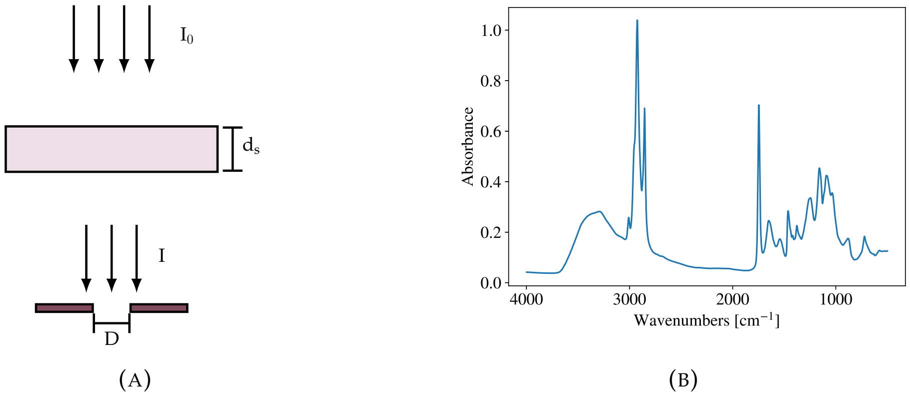


FIGURE 2.2: Simplified illustration of how IR absorbance spectra are recorded. **(A)** The reference spectrum is recorded when the incident radiation $\tilde{I}$ probes an empty sample holder and the transmitted radiation I reaches the detector. **(B)** The incident radiation is probing the sample which is mounted on the sample holder, and the transmitted radiation is recorded. By courtesy of Johanne Heitmann Solheim, Faculty of Science and Technology, NMBU.

2.2 The definition of the pure- and apparent absorbance spectrum

In electromagnetic theory absorption describes the phenomenon of light absorption by matter, causing for example the excitation of vibrational modes, when infrared radiation is used. Absorbance is a physical quantity of unit one defined by Eq. 2.3. The radiation coming from the IR source and attenuated by chemical absorption of molecules leads to chemical features in the spectra. In this case, the obtained absorbance may be considered as 'pure ' and denoted by the pure absorbance spectrum (Z_{pure}) [19]. However, in general some intensity may be lost due to scattering, causing unwanted physical features in the spectra. In this case, the obtained absorbance may be denoted as the apparent absorbance spectrum (Z_{app}) [39].

Pure absorbance spectrum is an ideal case whereas the measured spectrum is only affected by attenuation by chemical absorption. In Fig. 2.3a a visualization of an ideal case is shown, typically the sample is a thin film with a thickness d_s. In this case, the radiation reaching the detector *D* is only lost due to chemical absorbance. The pure absorbance spectrum as a function of wavenumber is illustrated in Fig. 2.3b. The characteristic of a pure absorbance spectrum is a constant flat baseline and all absorbance peaks correlate with the actual absorption of the vibrational modes. This means that high absorbance peaks correlate with high absorption, where absorbance equal to one means that 90% of the intensity has been absorbed by the sample at the corresponding wavenumber. Further, absorbance equal to zero means that all of the incident radiation transmits the sample, for instance, the flat region at 2000 cm^{-1} is typically named the silent region due to zero absorption. The spectrum illustrated is of the filamentous fungus *Mucor Circinelloides*.



(A) (B)

FIGURE 2.3: **(A)** Visualization of a transmission experiment. The incoming radiation intensity I_0 propagates towards the sample with a thickness d_s, and the transmitted intensity I reaches the detector D. Here, the loss in intensity is only due to chemical absorption by the sample. **(B)** Pure absorbance spectrum of the filamentous fungus *Mucor Circinelloides*, measured with an HTS FTIR instrument [11]. By courtesy of Johanne Heitmann Solheim, Faculty of Science and Technology, NMBU.

To get a better understanding of how the pure absorbance spectrum is obtained, a closer look at the intensity of the incident- and detected plane wave is necessary. The incident intensity can be expressed in terms of the incident electric field amplitude E_0 as following [19]

$$I_0 = \frac{c}{2}\epsilon_0|\vec{E_0}|^2, \tag{2.4}$$

where the constant c is the speed of light, ϵ_0 is the permittivity in vacuum. The electric field amplitude is changing depending on the depth position x in the sample, and can be expressed as

$$E(x) = E_0 e^{i\tilde{k}x}, \tag{2.5}$$

where $\tilde{k}$ is the complex angular wavenumber and can further be written as

$$\tilde{k} = \alpha + i \cdot \kappa = k_0 \cdot \tilde{n}, \tag{2.6}$$

where α and κ are the real and imaginary part of $\tilde{k}$, respectively. k_0 is the angular frequency in vacuum and $\tilde{n}$ is the complex refractive index. When the plane wave is detected by the detector it has transmitted through the whole sample, given $x = d$. The intensity recorded by the detector can now be written as

$$I = \frac{c}{2}\epsilon_0|\vec{E}|^2 = \frac{c}{2}\epsilon_0|\vec{E_0}|^2 e^{-2\kappa d}. \tag{2.7}$$

With Eq. 2.4 and Eq. 2.7 it is now possible to calculate the transmittance according to Eq. 2.2

$$T = e^{-2\kappa d} = e^{-4\pi n' d\tilde{v}}, \tag{2.8}$$

where n' is the imaginary part of the complex refractive index. The pure absorbance spectrum is derived in the following way

$$Z_{pure}(\tilde{v}) = -\log T(\tilde{\nu}) = \frac{4\pi n' d\tilde{v}}{\ln 10}. \tag{2.9}$$

$Z_{pure}(\tilde{v})$ is only explaining the chemical absorption by the matter, which means that $\frac{-4\pi n'\tilde{v}}{\ln 10}$ is a measure of the absorptivity of the sample, denoted as k. From Eq. 2.9 it is evident that the absorbance spectrum is wavelength dependent and proportional to the sample thickness d. From the same equation, Beer-Lambert's law can be derived as following

$$Z(\tilde{v}) = k_j(\tilde{v}) \cdot d \cdot c_j, \tag{2.10}$$

where k_j refers to the characteristic absorptivity for each component, d is the sample thickness and c_j is the concentration of each component$j$, if the sample is a mixture [19].

Only in ideal cases, the thin film is scatter free. In most practical cases this approximation is not applicable. Hence the term apparent absorbance is introduced. The apparent absorbance spectrum displays attenuation caused by both absorption and scattering features as illustrated in Fig. 2.4a. Due to the sample shape and size not all of the radiation will be transmitted or absorbed, but some of the incident radiation will also be lost due to diffuse scattering. As shown in Fig. 2.4b the scattering is prominent in the absorbance spectrum, displayed as a tilted baseline.

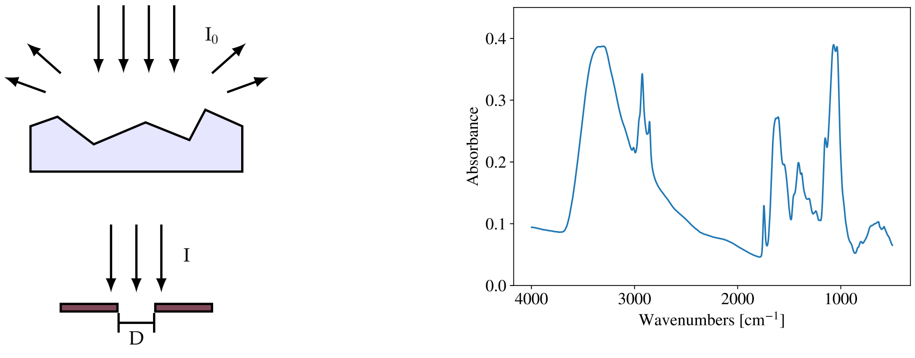
Assuming that the amount of radiation lost due to scattering is constant with respect to the wavenumber over the considered wavenumber range in the mid-infrared, the apparent absorbance spectrum can be expressed as

$$Z_{app}(\tilde{v}) = -\log \frac{I(\tilde{\nu})}{\alpha I_0(\tilde{\nu})} = -\log \frac{I(\tilde{\nu})}{\alpha I_0(\tilde{\nu})} + \log \alpha, \tag{2.11}$$

where α is a constant scaling effect for the intensity due to diffuse scattering. This scaling of the intensity turns into a constant baseline of the absorbance with the value $\log \alpha$. In addition to diffuse scattering, variations in the source intensity between the background and the sample intensity may lead to the same constant baseline variations. We see that if the source intensity is scaled with a factor of α compared to the background measurement a baseline shift occurs in the absorbance spectra. The scaling coefficient may vary with the wavenumber, which may cause a tilted baseline, as illustrated in the absorbance spectrum in Fig. 2.4b [39].

For single cell spectroscopy more sophisticated scatter effects occur. Biological samples like single cells and or cells in a tissue have spherical morphological structures. Since the wavelength of the incident infrared radiation is in the same order as the size of biological cells, strong scattering effects appear. For cells, these scattering effects have been interpreted as Mie-type scattering, which causes highly distorted infrared absorbance spectra [30] [38].

When absorbance spectra suffer from these distinct physical effects, the absorbance does not longer have Beer-Lamber behavior and estimating the chemical composition of the sample is not possible anymore. Modelling scatter effects is thereby an important activity in IR spectroscopy, and several methods have been proposed to correct the spectra and reconstruct the pure absorbance spectra to retrieve the chemical fingerprint [[18] [3] [4] [19] [20] [38] [23]].



(A) (B)

FIGURE 2.4: **(A)** Visualization of a transmission experiment when incoming radiation intensity I_0 is lost due to both chemical absorption and scattering. The attenuated radiation intensity I reaches the detector D. **(B)** Apparent absorbance spectrum recorded from the filamentous fungus *Mucor Circinelloides* with an HTS FTIR instrument [11] . By courtesy of Johanne Heitmann Solheim, Faculty of Science and Technology, NMBU.

2.3 Spatial resolution

Every optical system is limited by spatial resolution, and factors such as misalignment and imperfections in the lenses can affect the resolution in a negative way. However, there is one important and critical parameter in every optical system which limits the spatial resolution, namely the diffraction limit. The diffraction limit can be derived from the Rayleigh criterion

$$\Delta x \geq 0.61 \frac{\lambda}{n \sin \theta}, \tag{2.12}$$

where Δx is the spatial resolution, n is the refractive index of the medium and θ is the half-angle of the maximum angle of the aperture. $n \sin \theta$ refers to the numerical aperture denoted as NA [22]. According to the equation, it is clear that the diffraction limit is in the same order as the wavelength λ of the light source. Mid-IR light ranges from around 2.5 µm to 25 µm, which is in the same order as the sizes of many biological cells. Thus, due to the limitation it is not possible to achieve information on subcellular level or nano size level. As seen, spatial resolution sets boundaries of the size of objects that can be spatially resolved by optical methods.

2.4 Fourier transformed infrared spectroscopy

2.4.1 Michelson interferometer and Fourier transform

FTIR spectroscopy is based on the interferometer designed by Michelson more than 100 years ago, namely the Michelson interferometer [29]. The configuration consists of collimated radiation which propagates towards a beam splitter, where the radiation is split into two perpendicular beams, as shown in Fig. 2.5 [14]. Assuming an ideal beam splitter, 50% of the beam is transmitted and the other 50% is reflected. The beamsplitter introduces a phase difference between the two beams due to a phase change of 90° of the reflected beam. The transmitted beam strikes a movable mirror, while the reflected beam strikes a fixed mirror. Both beams are reflected by the mirrors, and propagate back to the beam splitter and recombine. Depending on the position of the movable mirror, the two beams are in phase or out of phase when they recombine, creating either constructive or destructive interference. The intensity of each beam striking the detector depends on the phase of the two different beams and how they interfere. If the beams are out of phase causing destructive interference the beam will propagate back to the source causing zero intensity at the detector.

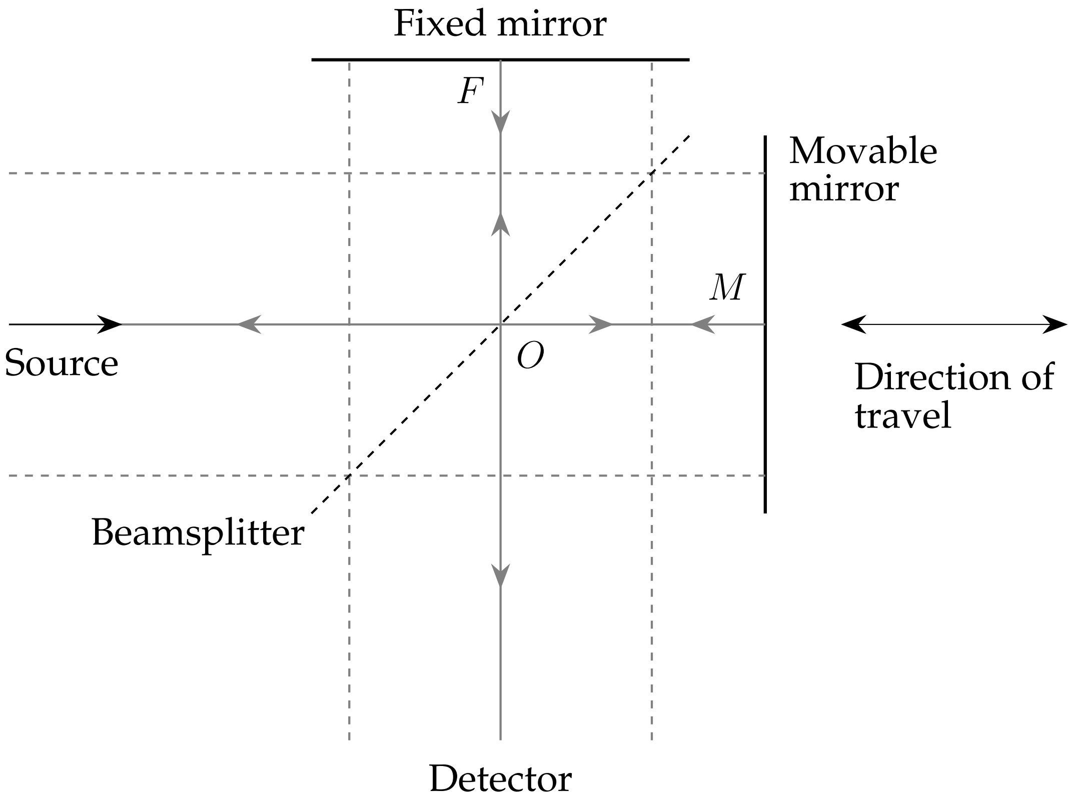


FIGURE 2.5: Visualization of the Michelson interferometer. The incident light propagates towards a beamsplitter, where it is either transmitted or reflected. Further, the light strikes either the fixed or the movable mirror, where it reflects and propagates back to the beamsplitter. The light recombines and reaches either the detector or is reflected back to the source depending on the phase of the beam. With inspiration from [14].

Depending on the movable mirror, the phase difference will vary and the mirror can either be a continuous-scan interferometer, where the mirror moves at a constant velocity, or it can be a step-scan interferometer, where the mirror moves between some fixed points, or it can be a rapid-scan interferometer, where the mirror moves continuously with high velocity, approximately 0.1 cm s^{-1} [35]. The **O**ptical **P**hase **D**ifference (OPD) of the two beams is called retardation and is denoted by the symbol δ. The spectral information achieved from the Michelson interferometer is the fluctuation of intensity as a function of the retardation and has the symbol $I'(\delta)$. If the mirror is moved with constant velocity, which is most common, the detected intensity varies sinusoidally given by

$$I'(\delta) = 0.5I(\tilde{\nu_0})(1 + \cos 2\pi\tilde{\nu_0}\delta). \tag{2.13}$$

The equation consists of a constant part and a sinusoidally part, however, it is only the non-constant part that is of interest and is known as the interferogram $I(\delta)$

$$I(\delta) = 0.5I(\tilde{\nu_0})\cos 2\pi\tilde{\nu_0}\delta, \tag{2.14}$$

where $I(\tilde{\nu_0})$ is the intensity of the source. It is practically impossible to have an ideal beamsplitter, thus $I(\tilde{\nu_0})$ is multiplied by a wavenumber dependent factor of $\cos 2\pi\tilde{\nu_0}\delta$, which corrects for non-ideal beamsplitter. The signal achieved by the interferogram is depending on the efficiency of the beamsplitter, the response at the detector and the characteristics of the amplifier, in addition to the intensity of the source. The source intensity is the only factor that varies for each measurement, the other parameters are fixed for each system configuration. Taking these factors into consideration, it is possible to modify Eq. 2.14 by a single wavenumber-dependent correction factor, $H(\tilde{V_0})$. Hence, the signal achieved by the interferometer is

$$S(\delta) = 0.5H(\tilde{\nu_0})G(\tilde{\nu_0})I(\tilde{\nu_0})\cos 2\pi\tilde{\nu_0}\delta, \tag{2.15}$$

where $G(\tilde{\nu_0})$ is the responsitivity by the detector and the amplifier, measured in $(V \cdot W^{-1})$. To simplify the equation let $B(\tilde{v})$ be equal to $0.5H(\tilde{\nu_0})G(\tilde{\nu_0})I(\tilde{\nu_0})$. The interferogram can now be expressed as

$$S(\delta) = B(\tilde{\nu_0})\cos 2\pi\tilde{\nu_0}\delta, \tag{2.16}$$

where $B(\tilde{\nu_0})$ is the intensity of the source at the specified wavenumber. The spectral information is finally achieved by computing the cosine Fourier transform of $S(\delta)$, and this is why it is called Fourier transform infrared spectroscopy. The Fourier transform turns the signal measured at the detector as a function of path difference $S(\delta)$ into the intensity as a function of the wavenumber $\tilde{v}$.

2.5 The photothermal effect based techniques

2.5.1 The photothermal effect

Optical-Photothermal IR spectroscopy and Atomic Force microscopy-based IR spectroscopy (AFM-IR) are two techniques based on the detection of the photothermal effect that occurs when IR radiation is absorbed by the sample. It results in a change of temperature, refractive index within the sample and induces thermal expansion and photomechanical pressure change [2]. Equation 2.9 shows that the power absorbed by the sample is proportional to the imaginary component of the refractive index, n'. When the IR radiation is focused onto the sample some of the IR power is absorbed by the material. The absorbed power P_{abs} can be expressed by [8]

$$P_{abs} = \int_V \frac{\pi c \epsilon_0}{\lambda} Im(\tilde{n}(\lambda)^2)|E_{loc}|^2 dV, \tag{2.17}$$

where V is the volume of the sample, c is the speed of light, λ is the wavelength of the radiation, $\tilde{n}(\lambda)^2 = n^2 + in\kappa - \kappa^2$, where only the imaginary part is considered, namely $in\kappa$. Further, $|E_{loc}|^2$ is the electric field inside the sample. The amount of energy which is absorbed by the sample is strongly depending on the electric field inside the sample and the refractive index. Normally, the electric field inside a sample is not known. In addition, the refractive index may vary within the sample and cannot be obtained as a function in space. Consequently, it is necessary to make assumptions. As we recall, mid-IR radiation is in the range of 2.5-25 µm and the sample size is in the size range of a few micrometers. Hence, the electric dipole approximation can be used, stating that small variations of the electric fields inside the sample can be neglected. Therefore, it can with good approximation assume that the electric field inside the sample is constant [21]. The second assumption is that the absorption corresponding to a molecular vibration is weak from which it follows that $n^2 >> n'^2$. By Eq. 2.9 the imaginary part of the complex refractive index n' can be calculated, by assuming that the absorbance $Z = 1$, the sample thickness is 10 µm, and $\lambda = 6$ µm. Then, solving the equation gives $n' = 0.1$ which is much smaller than n which is around 1.3. With these two assumptions the absorbed power P_{abs} in Eq. 2.17 can be expressed as [8]

$$P_{abs} = I_{inc}\alpha_{opt}V\tilde{v}n'(\tilde{v}), \tag{2.18}$$

I_{inc} is obtained by the relationship of $|E_0|^2$ from Eq. 2.4, and $\alpha_{opt} = 2\pi n$. From this equation it is evident that the power absorbed by the sample is proportional to the absorptivity of the sample explained in Eq. 2.9, since both the power absorbed and the pure absorbance spectrum are proportional to $\tilde{v}$ and n'. The IR radiation irradiates the sample in a pulsed duration time, corresponding to the rectangular box function. The concept of the photothermal expansion effect is to take advantage of the fact that IR radiation which is absorbed by the sample, causes a change in temperature. The increasing temperature in the sample induces thermal stress, which is relaxed by a mechanical expansion of the sample. With the Fourier heat equation, it is possible to mathematically express the relationship between the temperature change in time and space and the power absorbed by the sample.

$$\rho C_p \frac{\partial T}{\partial t} = \frac{\Pi(t_p) \cdot P_{abs}}{V} + k_{th}\Delta T, \tag{2.19}$$

where ρ is the density of the sample and C_p is the heat capacity. $\Pi(t_p)$ is the rectangular box function with length t_p. Assuming that the laser signal has the shape of a rectangular pulse with a duration of t_p. The quantity k_{th} represents the thermal conductivity and Δ is the Laplacian operator. $\Pi(t_p) \cdot P_{abs}$ may be denoted as the power of the heating source. Equation 2.19 gives a clear relationship between the laser duration, the power absorbed by the sample and the temperature change of the sample. When using an **O**ptic **P**arametric **O**scillator (OPO) laser, which is a short pulse laser with a duration time of one nanosecond, the laser duration is much smaller than the heat relaxing time of the sample. It can be shown that it is possible to solve the Fourier heat equation with respect to the temperature and obtain [8]

$$T = \frac{T_{max}}{t_p}t, \text{ when } 0 \leq t \leq t_p, \tag{2.20}$$

$$T = T_{max}e^{-\frac{(t-t_p)}{\tau_{relax}}}, \text{ when } t_p \leq t, \tag{2.21}$$

where the maximum temperature increase is expressed as $T_{max} = \frac{P_{abs}t_p}{V\rho C_p}$ and the relaxation time is expressed as $\tau_{relax} = \frac{\rho c_p}{k_{eff}}a^2$, where a is the radius of the sample and k_{eff} is the external heat conductivity of the environment surrounding the sample, i.e the air and the sample slide. By integrating the temperature with respect of time it is possible to visualize the characteristic of the samples temperature change as shown in Fig. 2.6. There is a linear relationship between time and temperature when the time is less or equal to the laser duration, reaching the maximum temperature when the time is equal to t_p. When the time is exceeding t_p the temperature is decreasing exponentially. The mechanical expansion caused by the increasing temperature can be expressed as

$$u(t) = aG\alpha_T\Delta T(t), \tag{2.22}$$

where a is the size of the heated region, G is a constant depending on the geometry of the sample, α_T is the thermal expansion coefficient and $\Delta T(t)$ is the rise in temperature due to the absorption of the IR laser.

It has now been shown that the power absorbed by the sample and the absorptivity of the sample are both proportional to the wavenumber and the imaginary part of the refractive index. Hence the power absorbed and the absorptivity of the sample are linearly dependent on each other. Further, It is shown by the Fourier heat equation that the change in temperature is proportional to the power absorbed. Lastly, the thermal expansion is shown to be proportional to the temperature change. Thus, the photothermal expansion of the sample is proportional to

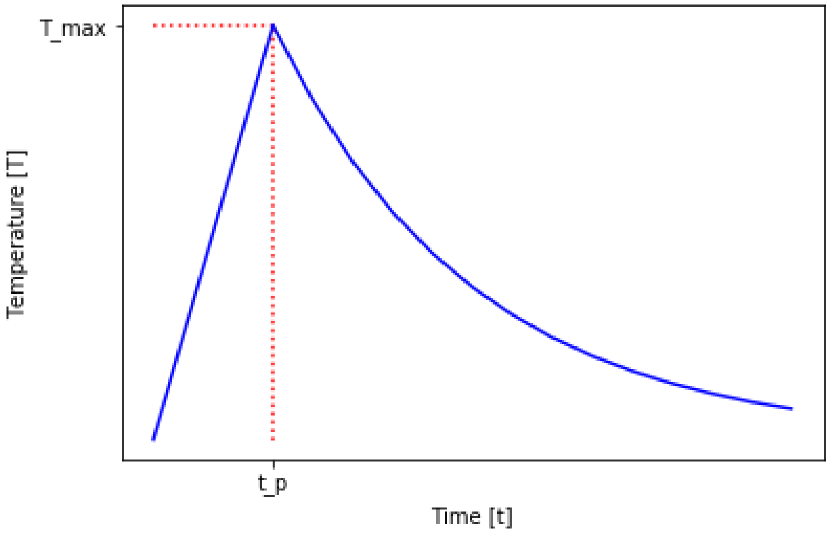


FIGURE 2.6: The blue graph shows the temperature change of the sample when an IR laser illuminates it. The characteristic temperature change shows that the laser interval (t_p) is less than the relaxation time of the sample, as obtained example with an OPO laser.

the power absorbed by the sample [9].

2.5.2 Optical-probed photothermal induced infrared microspectroscopy

Optical-probed **P**hoto**T**hermal induced **I**nfra**R**ed microspectroscopy (O-PTIR), is a newly emerging technique that uses a visible laser to probe the photothermal effect occurring inside the sample. The instrumentation setup is shown in Fig. 2.7 where the number 1 denotes a tunable IR laser beam and how it is guided to the sample surface. The number 2 denotes the meeting point of the IR laser and the green visible laser, where the two beams are made collinear and guided towards a Cassegrain microscope objective which is denoted with the number 3. The objective focuses the two beams onto the sample surface. The number 4 denotes the sample where the IR laser and the green laser are focused onto the sample. The IR laser illuminates a larger area of the sample and the absorption of the IR laser causes a photothermal expansion. As described in the previous chapter the absorption of IR radiation causes a change in temperature, refractive index and thermal expansion. The green laser probes a smaller area, and due to the changes in temperature and refractive index the reflection of the green laser changes. The reflected beam is guided back to the detector marked with the number 5 in the illustration [31]. Due to the green laser, also denoted as the probing laser, the instrument is wavelength independent in the IR spectral region. A priori, This means that the spatial resolution is not dependent on the IR laser, but on the green laser which is a continuous wave with a wavelength of 532 nm. Hence, submicron spatial resolution is achieved. Another advantage of O-PTIR is that the IR absorbance spectrum is not distorted by IR Mie scattering [16]. So far the theory behind this technique is not well explained in the literature, and there are no good explanations of how the IR O-PTIR signal is recorded. However, one way it is explained is that the change in reflectivity of the visible laser is caused by the thermal expansion causing a change in the intensity of the scattered light, also known as Rayleigh scattering. By deriving the Rayleigh scattering equation it is possible to calculate the photothermal response in the following way

$$\Delta P_{PR} \propto \frac{\sigma N}{\kappa C_p} \frac{\partial n}{\partial T} P_{pr} P_{IR}, \tag{2.23}$$

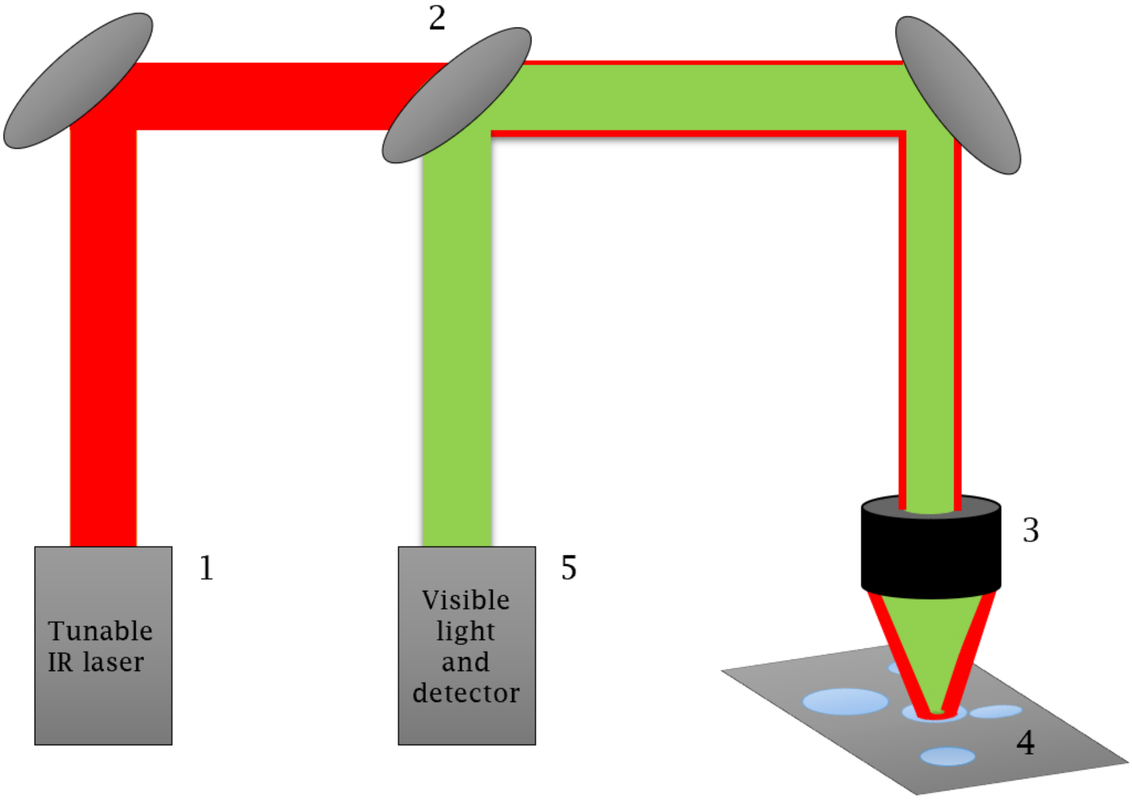


FIGURE 2.7: Schematic view of the O-PTIR system. **(1)** The laser path of the pulsed tunable IR laser is in red. **(2)** The IR laser and the green probe laser are made collinear. **(3)** A microscope objective focuses the two collinear beams onto the surface of the sample. **(4)** The sample absorbs the IR laser radiation, causing a photothermal expansion effect that causes a change in the reflection of the green probe laser. **(5)** The reflected light returns back to the detector, and the IR response is recorded.

where ΔP_{PR} is the probe power detected, σ is the absorbance cross section of the sample, N is the atomic number density, κ is the heat conductivity, C_p is the heat capacity, n is the refractive index, T is the temperature, P_{pr} is the power of the visible laser and P_{IR} is the power of the IR laser [44]. The probe power monitored by the detector is then processed and allows to indirectly measure the IR response of the sample, unlike traditional IR instruments where the direct absorbance of the sample is measured. Another explanation of how the O-PTIR signal is obtained is due to thermal diffusion, where the refractive index of the air around the sample changes. Thus the reflectivity of the green laser changes and is then recorded. However, this is not explained well in the literature.

2.5.3 Atomic force microscopy-based infrared spectroscopy

Atomic **F**orce **M**icroscopy-based **I**nfra**R**ed spectroscopy (AFM-IR) is a technique where atomic force microscopy and IR spectroscopy are combined [28]. The technique takes advantage of the linear proportionality between the photothermal expansion effect and the absorptivity of the sample to obtain an IR absorbance spectrum. The instrumental setup is shown in Fig. 2.8 and consists of an infrared laser that highlights the sample, a cantilever, a red diode and a photodiode. When measuring, the tip of the AFM cantilever is placed in contact mode with the sample. The IR laser is tuned to a specific wavelength in the mid-infrared, causing an almost instantaneous photothermal expansion due to an instant temperature increase in the absorbing area. Further, the photothermal expansion causes a force impulse on the cantilever through the tip, causing a deflection on the cantilever. Following a temperature rise, the tip of the cantilever reacts accordingly, and the cantilever starts to oscillate on its eigenmodes. The cantilever oscillation is recorded with the optical system of the AFM: a red diode and the photodiode, as seen in the yellow dashed line in the figure. The value of the maximum peak-to-peak amplitude of the oscillation may be plotted versus the IR radiation illuminating the sample to reconstruct the local IR absorbance spectrum [8]. AFM-IR provides spatial resolution of less than 50 nm allowing subwavelength infrared imaging. The resolution is only limited by the radius of the cantilever tip [9]. The high spatial resolution makes it possible to achieve spectral information far below the conventional FTIR technique. Thus, this makes AFM-IR measurements an attractive technique in the study of for example subcellular compartments of biological cells. The technique allows high spatially-resolved IR mapping of cells and provides highly resolved spectral data.

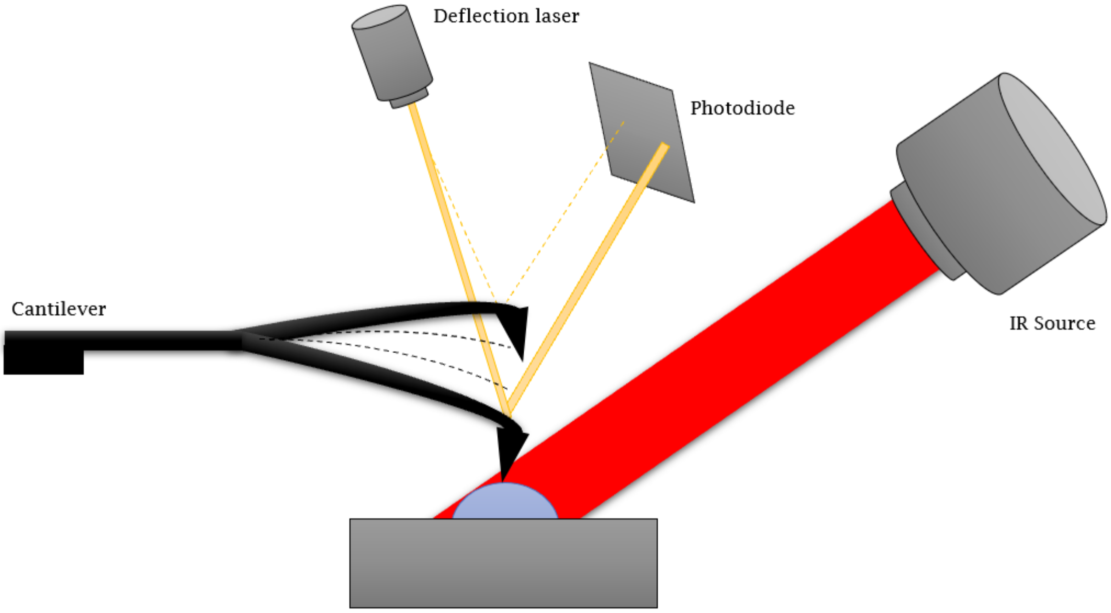


FIGURE 2.8: Schematic view of the AFM-IR setup with the infrared laser illuminating the sample. The cantilever is in contact mode and starts to oscillate when the sample absorbs the IR radiation. The oscillation changes the angle of the deflection laser, which is recorded by the photodiode. With inspiration from [28].

2.6 Spectral preprocessing and data analysis

2.6.1 The multiplicative signal correction

Multivariate **S**ignal **C**orrection (MSC) is a frequently used technique in IR spectroscopy to remove physical effects like scattering, scaling effects and baseline shift in apparent absorbance spectra. The aim of the prepossessing is to reconstruct the pure absorbance spectrum by removing the physical effects to obtain only the chemical contributions in the spectrum. MSC was first introduced in 1983 by Martens et al. [25]. The MSC model is based on the Beer-Lambert Law. MSC is the simplest version of the correction. In MSC the apparent absorbance spectrum $Z_{app}(\tilde{v})$ is modelled as

$$Z_{app}(\tilde{v}) = a + b \cdot Z_{ref}(\tilde{v}) + \epsilon(\tilde{v}), \tag{2.24}$$

where a describes the constant baseline shift, b is a scaling parameter, $Z_{ref}(\tilde{v})$ is the reference spectrum and $\epsilon(\tilde{v})$ takes care of the residuals which are not described by the model. The reference spectrum is ideally a good representative of the samples chemical composition. In many cases, the mean of the apparent absorbance spectra in the dataset is used as the reference spectrum. The corrected spectra are then standardized with respect to this average which represents for example the average thickness of the sample. Using a reference spectrum which is a good representative of the sample set considered leads to a stable correction model [39]. Least squares regression is performed in order to obtain the parameters a and b. The corrected absorbance spectrum is obtained by removing the physical effects from the apparent absorbance spectrum as follows

$$Z_{corr}(\tilde{v}) = \frac{Z_{app}(\tilde{v}) - a}{b}, \tag{2.25}$$

where Z_{corrr} is the corrected absorbance spectra. According to Eq. 2.24 $Z_{corr}(\tilde{v})$ can also be written as

$$Z_{corr}(\tilde{v}) = Z_{ref}(\tilde{v}) + \frac{\epsilon(\tilde{v})}{b}. \tag{2.26}$$

We recall that b is the scaling effect and $Z_{ref}(\tilde{v})$ is often the mean spectrum of $Z_{app}(\tilde{v})$. From Eq. 2.26 it can be seen that the interesting chemical features of each spectrum are taken care of by the residual spectrum $\epsilon(\tilde{v})$. An advantage of MSC is that it easily can be extended, which makes it a robust model when dealing with highly complex scatter effects.

2.6.2 The extended multiplicative signal correction

Extended **M**ultiplicative **S**ignal **C**orrection (EMSC) is the extended version of MSC introduced in 1991 by Martens and Stark [27]. It is a standard prepossessing method in IR spectroscopy because of its flexibility in choosing higher polynomial orders correcting for the complexity of the data set. Due to the fact that diffuse scattering is often wavenumber dependent, significant baseline distortions may occur which often can be approximated by polynomials. By adding wavenumber-dependent polynomials to the MSC model of Eq. 2.24 the model is extended as following [17]

$$Z_{app}(\tilde{v}) = a + b \cdot Z_{ref}(\tilde{v}) + c \cdot \tilde{v} + d \cdot \tilde{v}^2 + \epsilon(\tilde{v}). \tag{2.27}$$

While there are in principle no limitations on the number of polynomial orders, it is important to choose the order of the polynomial carefully. Adding higher polynomial orders is only meant to correct the physical artifacts, but if too high polynomial orders are used, important chemical features may be removed as well. EMSC can be extended to also include other chemical absorbance spectra so-called constituent model spectra [40]. The constituent model spectra can both include analytes and interferents, which have been previously termed good-spectra and bad-spectra respectively [26]. In cases when the sample is embedded in a medium whose spectral features afterward need to be subtracted from the spectra, an interferent spectrum representing the media may be used. Epoxy, paraffin and water are examples of different media with unwanted chemical features in the apparent absorbance spectra. When selecting an interferent spectrum one needs to be careful and avoid too high collinearity with important chemical features in the target pure absorbance spectrum that is to be retrieved. Analyte spectra can as well be used as model spectra in the EMSC model. An analyte spectrum is used as a stabilizing factor in the model, and the aim is not to remove the analyte spectrum in the correction process. Instead, it can be used to quantify the concentration of a known constituent as the parameter associated with a respective analyte is estimated in the EMSC modelling process for each measured spectrum [40]. In order to understand how the constituent spectra are included in the EMSC model we consider how the reference spectrum is achieved according to Beer-Lambert law. The measured spectrum $Z(\tilde{v})$ can be expressed as:

$$Z(\tilde{v}) = \sum_{j=1}^{J} c_j k_j(\tilde{v}), \tag{2.28}$$

where c_j is the concentration of compound j and k_j is the absorbance of compound j. When constituent spectra are included, the total chemical compounds can be written as

$$\sum_{j=1}^{J} c_j k_j(\tilde{v}) = \sum_{j=1}^{J} c_j(k_j - Z_{ref}(\tilde{v})) + Z_{ref}(\tilde{v}), \tag{2.29}$$

assuming the sum of concentrations is equal to 1. Rewriting $(k_j - Z_{ref}(\tilde{v}))$ to Δk_j, which is the chemical differences from the reference spectrum. The EMSC model can now be expressed by

$$Z_{app}(\tilde{v}) = a + b \cdot Z_{ref}(\tilde{v}) + c \cdot \tilde{v} + d \cdot \tilde{v}^2 + b \cdot \left\{ \sum_{j=1}^{J} c_j \Delta k_j(\tilde{v}) \right\} + \epsilon(\tilde{v}). \tag{2.30}$$

Further, the expression can be simplified further by expressing $b \cdot c_j$ as h_j, such that

$$Z_{app}(\tilde{v}) = a + b \cdot Z_{ref}(\tilde{v}) + c \cdot \tilde{v} + d \cdot \tilde{v}^2 + \sum_{j=1}^{J} h_j \cdot \Delta k_j(\tilde{v}) + \epsilon(\tilde{v}). \tag{2.31}$$

When both analyte and interferent spectra are included in the EMSC model, it can finally be expressed as

$$Z_{app}(\tilde{v}) = a + b \cdot Z_{ref}(\tilde{v}) + c \cdot \tilde{v} + d \cdot \tilde{v}^2 + f \cdot Z_{ana}(\tilde{v}) + g \cdot Z_{int}(\tilde{v}) + \epsilon(\tilde{v}). \tag{2.32}$$

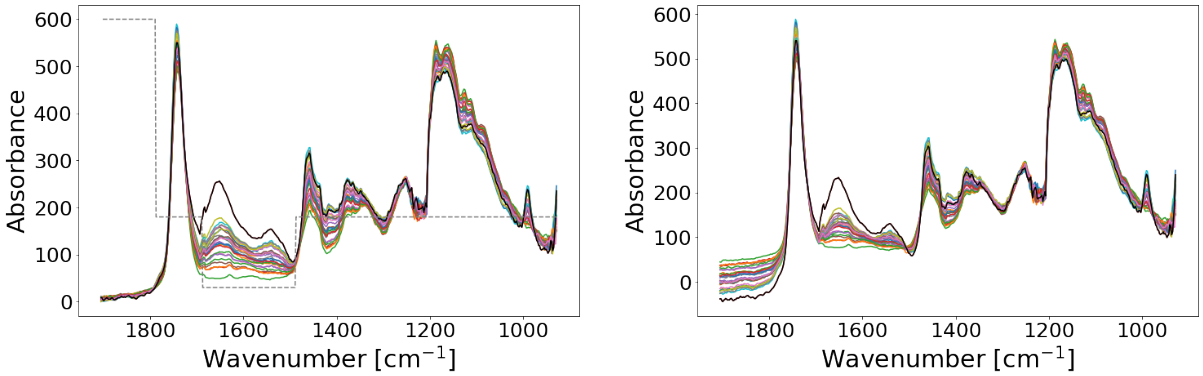
Z_{ana} and Z_{int} denotes the analyte spectrum and interferent spectrum, respectively. Finally, the corrected absorbance spectra can be expressed in the same way as Eq. 2.25 and Eq. 2.26, but including the extended version.

$$Z_{corr}(\tilde{v}) = \frac{Z_{app}(\tilde{v}) - a - c \cdot \tilde{v} - d \cdot \tilde{v}^2 - g \cdot Z_{int}(\tilde{v})}{b} \tag{2.33}$$

$$Z_{corr}(\tilde{v}) = Z_{ref}(\tilde{v}) + \frac{f \cdot Z_{ana}(\tilde{v}) + \epsilon(\tilde{v})}{b} \tag{2.34}$$

From Eq. 2.33 and Eq. 2.34 it is clear that the interferent and analyte spectrum has the opposite effects of each other.

Another useful method in EMSC is to apply weights on different regions in the spectra. For instance, regions with high chemical variations can be weighted down, and regions with high physical variations can be weighted up. The model focuses on correcting the variation that is weighted up instead of focusing on the whole spectrum. Figure 2.9 shows an example with and without weighted EMSC (see Fig. 2.9A and 2.9B, respectively). In the region above 1800 cm^{-1}, the weighted EMSC is able to correct for the scaling effect. Further, in the region between 1700 cm^{-1} and 1500 cm^{-1} the weighted model is not correcting for the chemical variability as desired in contrast to the non-weighted model. There are large chemical variations in the region between 1700 cm^{-1} and 1500 cm $^{-1}$, which we want to keep in the corrected spectra, thus this part is weighted down. There is no correct answer to how to apply weights, and in many cases, it is a process of trial and error. However, a good start is to weight physical variability around 1 and chemical variability around 0.1.



(A) (B)

FIGURE 2.9: Visualization of EMSC correction with and without weights. **(A)** Shows the correction with the weights 1, 0.05 and 0.3 shown by the grey dotted line. The weights are scaled by 600 in the figure for better visualization. **(B)** Shows the correction without any weighted parameters.

2.6.3 Principal component analysis

Principal **C**omponent **A**nalysis (PCA) is an often used tool in the data analysis of high dimensional and complex data sets. PCA is frequently applied for the multivariate data analysis of spectroscopic data. The aim of PCA is to decompose a data set into new latent variables named **P**rincipal **C**omponents (PCs). The principal components are linear combinations of the variables in the original data set and define a new coordinate system where the principal components are orthogonal with respect to each other [36]. The first principal component defines a new direction that maximizes the co-variance of the data set. Thereafter, the residual in the data set is considered and again the co-variance in the data is maximized, resulting in the second principal component etc. Since principal components are orthogonal to each other, the new variables are independent and uncorrelated. Usually, we are able to catch close to 100% of the explained variance by considering only the first few components in infrared data of biological samples. Therefore, PCA is an efficient tool for dimensionality reduction [14]. Mathematically PCA can be expressed as

$$X = \overline{X} + T_A P_A^T + E_A, \tag{2.35}$$

where X is the original matrix with dimension (N × K), where N is number of measurements and K is wavenumbers. Further, $\overline{X}$ is the mean of each row in X, and T_A is the score matrix with dimension (N × A) and P_A is the loading matrix with ($K \times A$) dimension. E_A denotes the residual matrix and represents the variance which is not expressed by the PC model $T_A P_A$. Figure 2.10 illustrates the dimension of the matrices. The score matrix is the new coordinates of the original data matrix returned by the PCA analysis, which can be plotted in a score plot where the PCs are the axes. The loading matrix explains how the original spectra are related to the new PCs, and may be visualized in a loading plot [36].

2.6.4 Multivariate curve resolution

Multivariate **C**urve **R**esolution (MCR) was developed by the field of chemometrics and can both be performed by a non-iterative and an iterative algorithm in the analysis of chemical

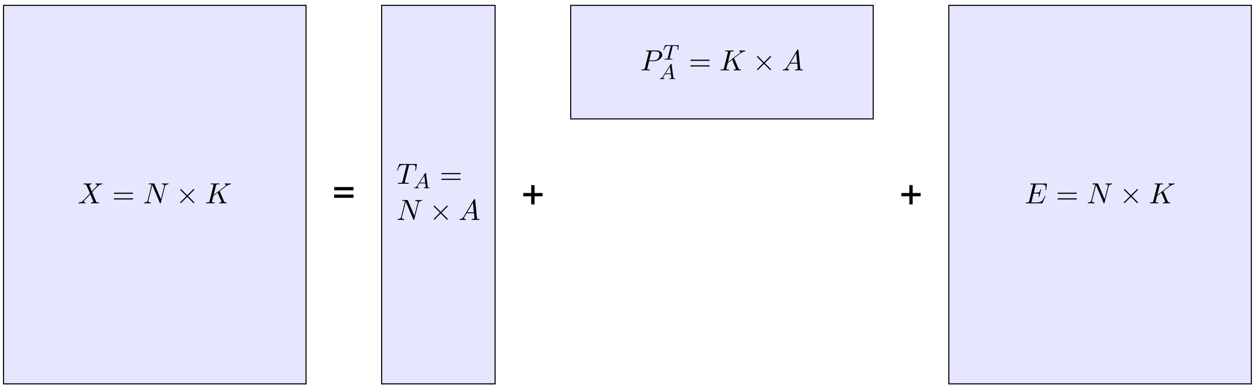


FIGURE 2.10: Visualization of the dimensions of the matrices in PCA, where X is the original data matrix, T_A is the score matrix, P_A^T is the loading matrix and E is the residual matrix.

mixture problems [34]. The iterative approach is the most common and the one considered in this thesis, more precisely **M**ultivariate **C**urve **R**esolution - **A**lternating **L**east **S**quare (MCR-ALS), is the most widely used algorithm. In the MCR of infrared spectroscopic data, a data matrix of absorbance spectra is decomposed by a linear model:

$$D = C \cdot S^T + E, \tag{2.36}$$

where D is the initial matrix with absorbance spectra, C and S are matrices representing the concentration and the respective pure component spectra. E is the residuals, which is not explained by the model. The dimensions of matrix D and E are (N $\times$ K), while matrix C is (N $\times$ A) and matrix S^T is (K $\times$ A). The matrix of concentrations C and the matrix of pure component spectra S have the same dimensions as the score and loading matrices in PCA, respectively (see Fig. 2.10). The difference is that in MCR certain conditions are imposed for the matrices C and S, such as for example non-negativity for the pure components S. The goal of MCR is to estimate the true values of C and S when only having information of the initial matrix D. The first step of the MCR-ALS algorithm is to determine the number of components, which can be the same as the number of principal components used in the PCA. The second step is to determine the initial estimates of C or S. In this thesis, the scores from the PCA have been used as initial estimates for C. The third step is the optimization of C and S by solving Eq. 2.36 with the iterative model ALS and by using a linear transformation matrix T. The model is updated as following [34]

$$D = C \cdot (T \cdot T^{-1}) \cdot S^T + E, \tag{2.37}$$

$$D = (C \cdot T) \cdot (T^{-1\cdot} S^T) + E. \tag{2.38}$$

In every iteration different constraints can be applied, whereas non-negativity is used in the data analysis in this thesis [41].

2.6.5 3D infrared diffraction tomography

IR spectroscopy aims to study chemical information in the examined material, however, due to loss of radiation caused by scattering, the chemical information in the spectra is often biased by scatter contributions. Biological samples have the same size as the magnitude of the wavelength of the radiation, thus highly scatter-distorted spectra, such as Mie-type scattering is prominent. Scattering is determined by the complex refractive index function of the sample. Further, the complex refractive index may vary in space since biological samples are not homogeneous, and the real and imaginary part of the refractive index are not constant over the sample volume [24]. Thus, both chemical and physical features are dependent on the refractive index and the absorptivity of the sample, causing the scattering features to be highly entangled with the chemical features in the absorbance spectra. The ME-EMSC algorithm has successfully managed to remove the Mie-scattering from the IR-absorbance spectra, by updating the estimate of the pure absorbance spectrum. The algorithm has become the state-of-art correction and is an open-source algorithm [38]. Deep learning models have later outperformed the ME-EMSC correction in terms of time and performance, as described by Magnussen et. al [23]. Further, by taking advantage of the fact that scattering and chemical absorption features in measured spectra for chemical cells are highly informative for the physical and chemical properties of cells, the inverse scattering problem of predicting chemical and physical properties of the cell wall and cell interior was solved by a deep convolutional neural network [24]. The algorithm predicts the physical properties such as size and refractive index of the cell wall and the cell interior, in addition, it predicts information rich pure absorbance spectra of the cell wall and cell interior. These predictions are made by an estimation of a highly scatted mean spectrum of the measured cell. The proposed algorithm has successfully managed to retrieve chemical information of biological cells and predict the size and chemical composition of the cell wall and the cell interior of an intact cell. This approach is termed 3D infrared diffraction tomography. In this thesis we use this algorithm, from the study made by Magnussen et al. [24], to predict the pure absorbance spectrum of cell wall and cell interior and the physical properties of the fungus *Mucor circinelloides* and the yeast cells *Phaffia rhodozyma* and *Rhodotorula graminis*. The approach will be compared with O-PTIR and AFM-IR and how well it is possible to obtain high quality chemical information of the cell wall and cell interior of both intact and sectioned biological cells.

Chapter 3

Method

3.1 Biological samples

The experimental part is divided into two different sections in regard to sample preparations, namely sectioned cells and intact cells. As for the sectioned cells, yeast *Rhodotorula graminis CCY 20-2-47* and algae *Aurautiochyrium limaanium* are used as biological samples. The size of the *Rhodotorula graminis* cell is approximately between 2.3 - 3.9 µm in diameter and the *Aurautiochyrium limaanium* has a cell size of approximately 5 - 10 µm in diameter. The cells were first treated with osmium and then sectioned with a cryo-microtome in 500 nm and 200 nm thin sections. To be able to make such fine sections, the samples were embedded in epoxy which is a type of resin. The epoxy is embedding the biological samples in liquid state. Samples were then dried for a period of three weeks, where they turned into a polymerized block. The sectioning was performed by the imaging center at the Norwegian University of Life Sciences. As intact cells, two different types of carotenogenic yeast were used, namely *Rhodotorula graminis CCY 20-2-47* and *Phaffia rhodozyma CCY 77-1-1*. Phaffia rhodozyma has a cell diamteter ranging from 6 µm to 10 µm. In addition, filamentous fungi *Mucor circinelloides FRR 5020* were studied. However, the exact cell size is difficult to define as it creates long hyphae which can vary in length and thickness. The cultivation of the intact cells was done by members of the BioSpec Norway group at Norwegian University of Life Sciences. Table 3.1 shows an overview of the different biological samples considered with the respective sample preparation and sample size for each sample.

TABLE 3.1: Overview of the instrumentation, sample preparation and sample size that were used for the different biological cells

	Rhodotorula graminis	*Phaffia rhodozyma*	*Aurautiochyrium limaanium*	*Mucor circinelloides*
Instrumentation	O-PTIR AFM-IR FTIR-FPA	O-PTIR FTIR-FPA	O-PTIR AFM-IR	O-PTIR FTIR-FPA
Sample preparation	Sectioned cells Intact cells	Intact cells	Sectioned cells	Intact cells
Sample size [µm]	2.3-3.9	6-10	5-10	Not defined

3.2 Instrumentation

All data measured with O-PTIR, AFM-IR and FTIR FPA hyperspectral measurements have been performed by the candidate. FTIR HTS measurements have been performed by members of the BioSpec Norway group at Norwegian University of Life Sciences, and **T**ransmission

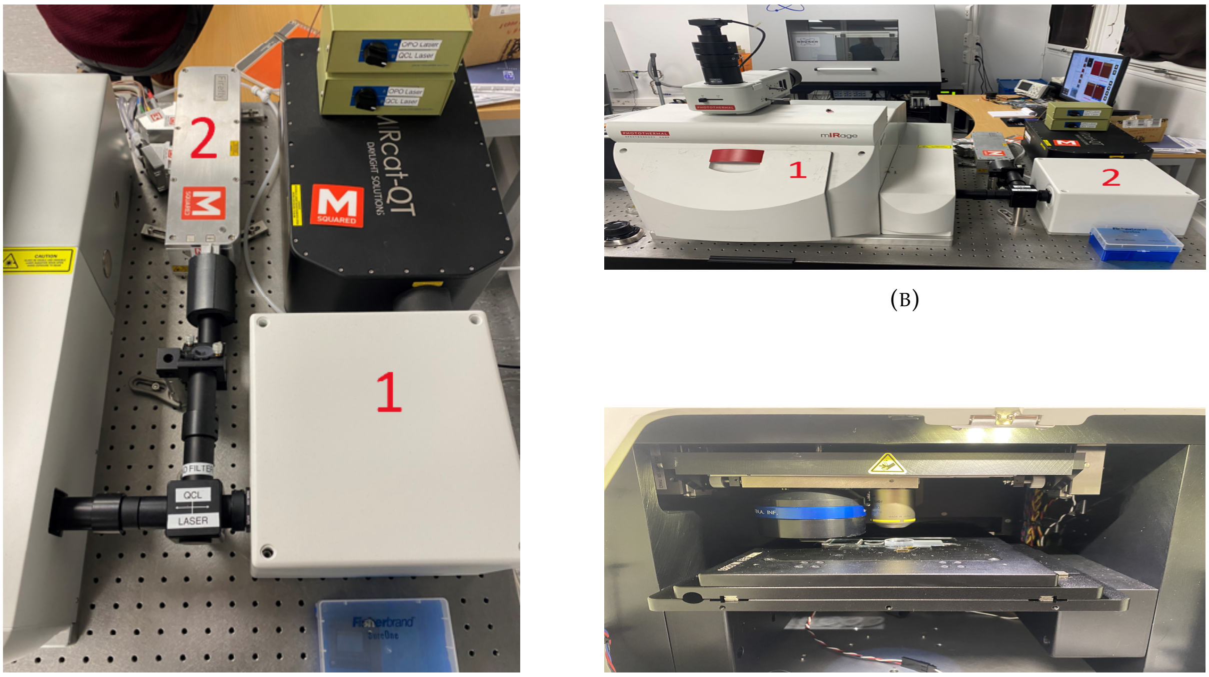
Electron **M**icroscopy (TEM) has been performed in an earlier study by the imaging center at the Norwegian University of Life Sciences.

O-PTIR measurements: All biological samples were measured with O-PTIR, in addition, measurement of pure epoxy was measured as reference. The O-PTIR instrument has four QCL lasers, which cover a spectral range of 910 cm^{-1} to 1903 cm^{-1}. Images of the O-PTIR instrument are shown in Fig. 3.1. Figure 3.1B shows the whole instrumental setup where the number **1** indicates the position of the microscope objective and number **2** shows the source of the IR laser beam, Further, Fig. 3.1A shows a closer image of IR laser setup. The setup depends on what laser type that is used. Number **1** indicates the path of the QCL laser and number **2** the path of the OPO laser. Figure 3.1C shows the stage where the sample is placed on. The stage is movable in x, y and z direction, where x and y are used to find the wanted position of the target sample, and changing the z position is used to get the right focus. Further, the image shows two different objectives, the one with a yellow thin stripe is a 10X objective with numerical aperture of 0.3 which gives a better visualization of larger sample areas. The objective with a blue stripe is a 40X objective with numerical aperture of 0.78. A background measurement was acquired every time a new sample was measured. The background measurement optimizes the position of the four lasers and it optimizes the power of the IR lasers which are shown in Fig. 3.2. Laser brakes occurs in the transitions between lasers and it needs to be corrected in the preprocessing. The laser breaks are approximately at 1693 cm^{-1}, 1433 cm^{-1} and 1205 cm^{-1}. An **A**valanche **P**hoto**D**iode (APD) detector, where used which is highly sensitive. Before measuring the samples the optical parameters needed to be optimized. This is done by looking at the response of the sample when changing the power of the IR laser and the green laser. By increasing the laser powers too much it is possible to burn the sample. In addition to tuning the two laser powers, the pulse rate and the pulse width may be optimized to obtain higher response of the sample. The optimal signal was achieved at 42% and 24% for the IR laser and 0.22% for the green probe power. The pulse rate and pulse width were mostly set to 100 kHz and 100 ns, respectively, however for some cases they were set to 200 kHz and 200 ns. These parameter yields for all samples except the algae because no optimal parameters were obtained due to bad quality of the measured spectra. An average of four spectra was used for all acquired spectra. The O-PTIR data were measured at Paris-Saclay university, institute of chemical physics in Paris, France.

Tentative depth-resolved O-PTIR measurements were performed on intact cells. Figure 3.3 shows a simplified version the optical setup for depth-resolved measurements. The green laser probes the sample at the focal point, and the sample stage can be moved up and down in the Z-direction to probe at different depths within the sample.

AFM-IR measurements: The 500 nm thin sectioned cells were as well measured with the AFM-IR. For the algae, it was not possible to obtain optimal parameters, because of bad quality of the spectra. With *Rhodotorula graminis*, the IR laser power was set to 9.79% with a pulse rate of approximately 400 kHz and a pulse width of 160 ns. Spectra were obtained as an average of two spectra with a spectral range from 900 cm^{-1} to 1900 cm^{-1}. The height images varied in width size between 1 µm and 4 µm and height size between 4µm and 5µm. The AFM-IR measurements were performed at Paris-Saclay university, institute of chemical physics in Paris, France.

FTIR FPA hyperspectral measurements: FTIR microspectroscopic imaging were measured of intact cells with an FTIR Hyperion 3000 from Bruker, with a FPA detector. The measurements were performed in transmission mode, with a resolution of 8 cm^{-1}. The number of



(A)

(B)

(C)

FIGURE 3.1: Images of the O-PTIR instrumentation. **(A)** shows the IR laser setup, where **1** and **2** illustrates the beam paths of the QCL laser and the OPO laser, respectively. **(B)** Number **1** indicates the location of the microscope and number **2** labels the IR lasers. **(C)** Image of the stage and the two different objectives.

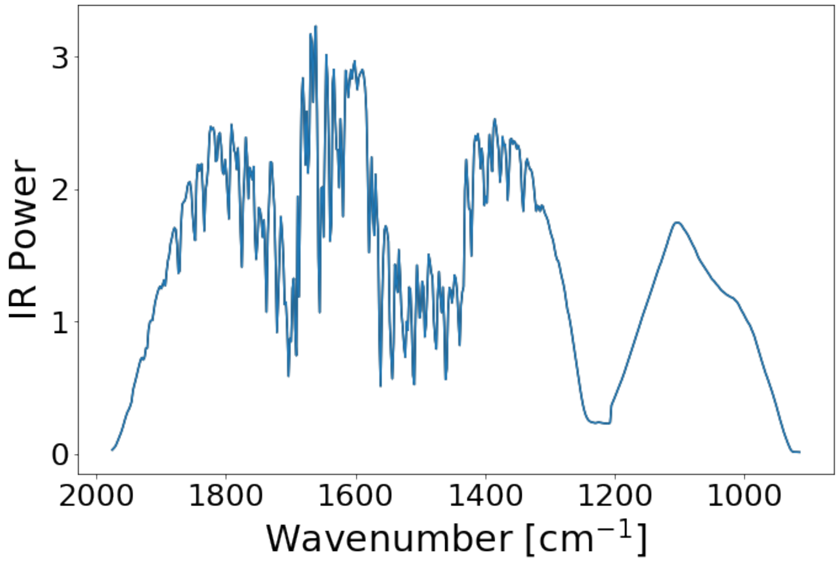


FIGURE 3.2: Optimized power of the four infrared QCL lasers. Data is obtained from an arbitrary background measurement.

scans varied between 32 and 64 scans. A background spectrum was acquired before every measurement.

In table 3.1 the three techniques O-PTIR, AFM-IR and FTIR FPA are shown with the corresponding measured biological samples.

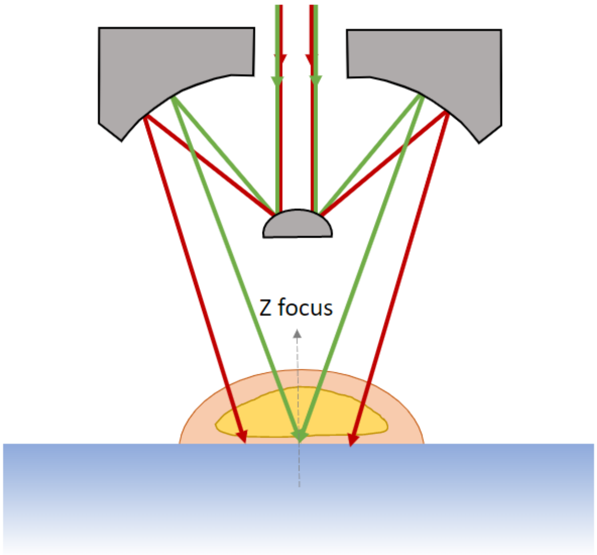


FIGURE 3.3: Schematic view of the depth-resolved O-PTIR measurement. The green laser probes the sample at different focus points within the sample by moving the stage in the Z-direction. The IR laser is illuminating a larger part of the sample. With inspiration from A. Dazzi.

FTIR HTS measurements: **H**igh **T**hroughput **S**creening (HTS) measuremnts of *Mucor circinelloides*, *rhodotorula graminis* and *phaffia rhodozyma* was performed by BioSpec Norway group at Norwegian University of Life Sciences. The measurements were measured with spectral resolution of 6 cm^{-1} and number of scans was 64.

Transmission electron microscopy: TEM image of sectioned *Rhodotorula graminis* is shown in Fig. 3.4, were subcellular structures are seen. The cell wall is the edge around the whole cell and the cell interior is considered to be everything within the cell wall. The image has been performed by the imaging center at the Norwegian University of Life Sciences.

3.3 Spectral preprocessing and data analysis

O-PTIR and AFM-IR: All spectra obtained from O-PTIR and AFM-IR have been laser corrected using a Python script made by Paris-Saclay University, institute of chemical physics in Paris, France. In general, the laser breaks occur approximately at 1693 cm^{-1}, 1433 cm^{-1} and 1205 cm^{-1}, however, in some cases the laser breaks may be shifted a few wavenumbers. It is therefore important to look at the acquired spectra to be sure where the laser breaks occur.

FTIR FPA hyperspectral images: The FPA hyperspectral images of intact cells was applied to the 3D infrared diffraction tomography approach which is a pre-trained deep convolutional neural network algorithm published by Magnussen et al. [24].

To analyze the data the software Quasar version 1.5.0 and the programming language Python was used.

Three different terms are used to label the sampling of spectral data, namely point scan, line scan, and spectral image. Point scan refers to a measurement of a single spectrum at one position of the sample, line scan denotes sampling of spectra along a line, and with spectral image we refer to sampling of spectra in an image plane.

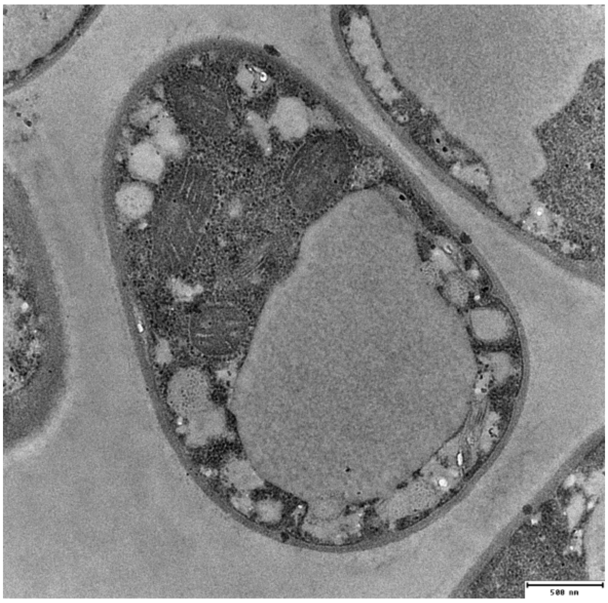
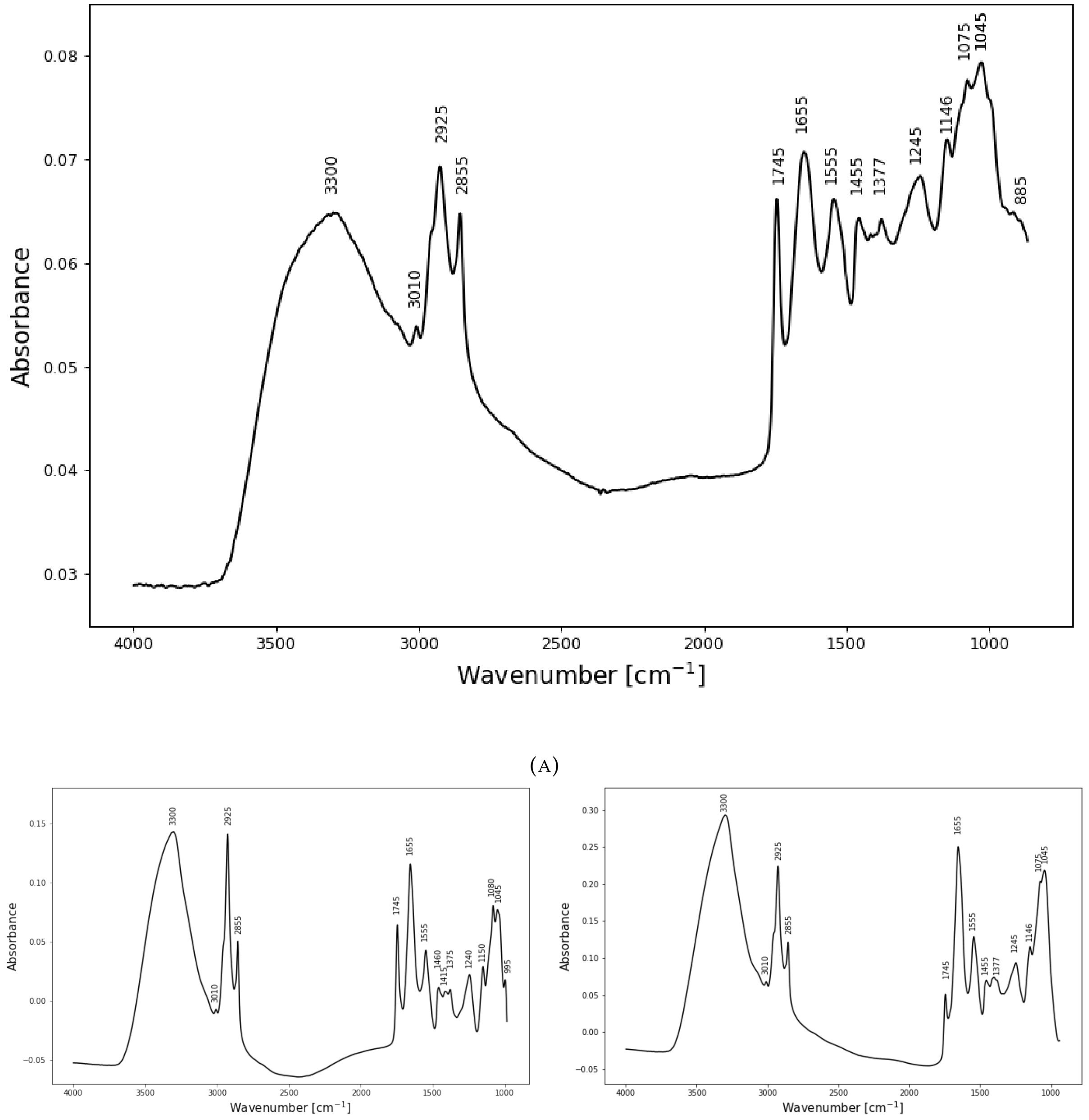


FIGURE 3.4: TEM image of *Rhodotorula graminis* showing subcellular structures.

FTIR HTS spectra: Figure 3.5 A, B and C show HTS coupled to FTIR measurements of *Mucor circinelloids*, *Rhodotorula graminis* and *Phaffia rhodozyma*, respectively, which can be considered as relatively pure absorbance spectra. The most important peaks are annotated with the corresponding wavenumbers, and table 3.2 shows the respective peak assignments. In general, the cell wall consists of a higher concentration of carbohydrates, which are associated with peaks in the spectral region between 1200 cm^{-1} and 1000 cm^{-1}, relative to the cell interior which has a higher concentration of lipids and proteins. The peaks at 2925 cm^{-1}, 2855 cm^{-1}, 1745 cm^{-1} and 1146 cm^{-1} refer to lipids, as shown in table 3.2. Further, the two peaks at 1655 cm^{-1} and 1455 cm^{-1} are assigned to proteins, which are denoted as Amide I and Amide II. The figures show that the fungi and the two yeasts share the same vibrational bands. However, the heights of the absorbance peaks vary a lot. The peak heights are associated with the amount of molecules the different chemical compounds have. Thus, by evaluating the ratio of the area of different absorbance bands, it is possible to obtain quantitative information about the chemical composition of biological samples.



(A)

(B)

(C)

FIGURE 3.5: FTIR-HTS spectra of **(A)** *Mucor circinelloides* **(B)** *Rhodotorula graminis* **(C)** *Phaffia rhodozyma*. The most important peaks are annotated with the corresponding wavenumber.

TABLE 3.2: Peak assignments with the corresponding wavenumbers of biological cells [12].

Wavenumber (cm^{-1})	**Peak assignment**
3500-3200	O-H stretching (carbohydrates)
3275	N-H stretching (chitin/chitosan)
3010	= C-H stretching (lipids)
2925	>CH2 of acyl chain (lipids)
2855	C-H (CH2) stretching (lipids)
1745	-C = O stretching in esters (lipids)
1655	-C = O stretching, Amide I (proteins, chitin)
1555	C-N-H deformation, Amide II (proteins, chitin)
1455	-C-H (CH2, CH3) bending (lipids)
1377	-C-H (CH3) bending (chitin)
1245	P = O stretching (polyphosphates)
1146	C-O-C stretching in esters (lipids)
1075	C-O stretching (charbohydrates)
1045	C-O-C stretching (carbohydrates)
885	P-O-P stretching (polyphosphates)

Chapter 4

Results and discussion

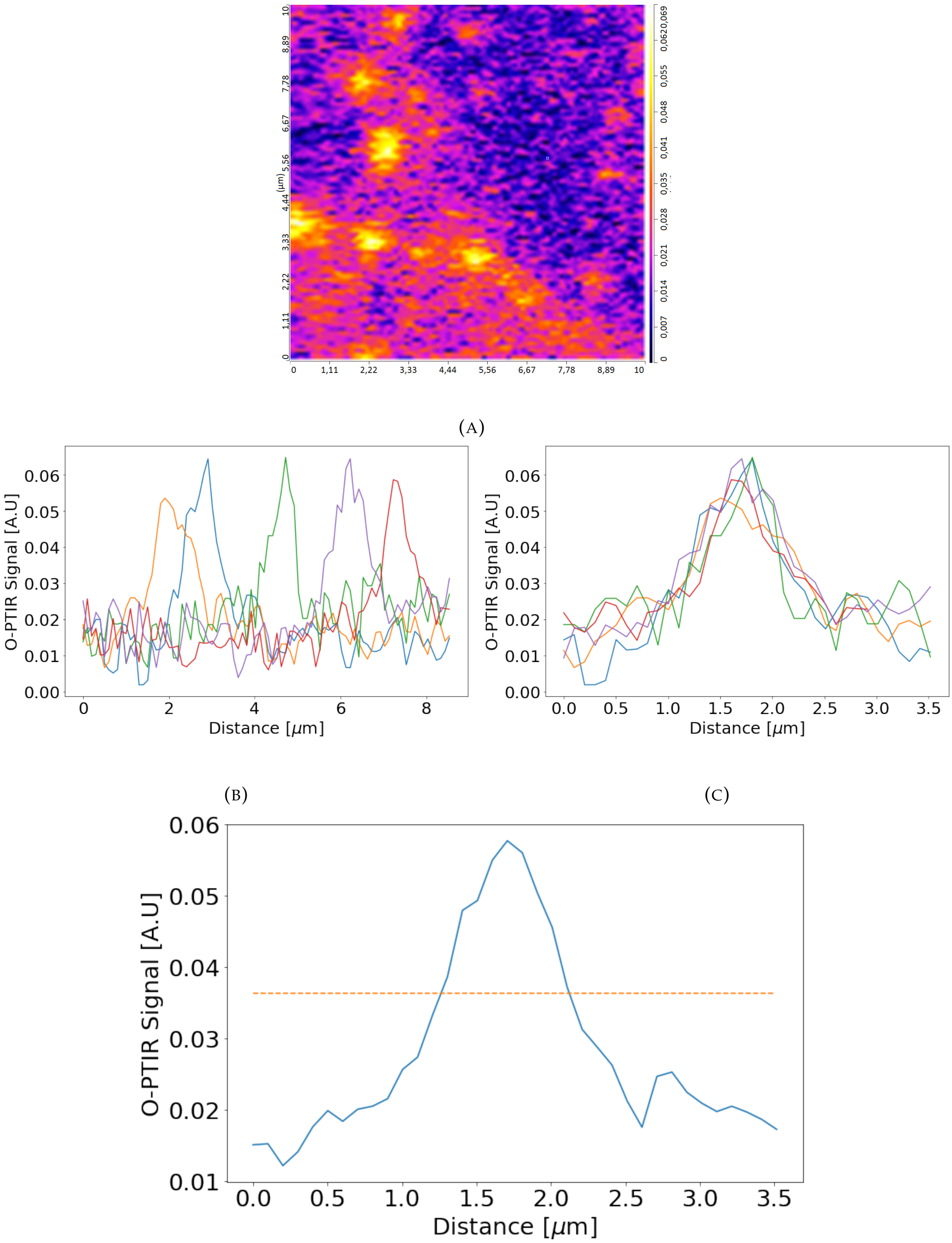
In the following, the three techniques O-PTIR, AFM-IR and 3D IR diffraction tomography will be compared with respect to their ability to obtain infrared spectral data from subcellular components of cells. In addition, sample preparation procedures will be compared for the three techniques. Sectioned cells of *Rhodotorula graminis* and *Aurautiochyrium limaanium* and, intact cells of *Rhodotorula graminis*, *Phaffia rhodozyma* and *Mucor circinelloides* will be compared and the preparation procedures will be evaluated. Further, we will discuss how well it is possible to obtain subcellular chemical information of depth-resolved measurements using O-PTIR. Lastly, FPA hyperspectral images of intact cells will be evaluated with the 3D IR diffraction tomography algorithm from Magnussen et al. [24].

4.1 Spatial resolution of O-PTIR

As mentioned in section 2.5.2 the lack of a theoretical explanation of the instrument makes it difficult to know precisely what the spatial resolution is. Thus, a measurement testing the spatial resolution has been performed. Figure 4.1A shows a mid-infrared photothermal image where the high-intensity points are **P**oly**S**tyrene (PS) beads with a diameter range of 500 nm. Further, Fig. 4.1B shows five line scans through five of the different beads. The offset value is estimated to be 0.015 which has been corrected in Fig. 4.1C. The scans have been aligned such that the peaks are overlapping. Figure 4.1D shows the average of the five line scans. The **F**ull **W**idth **H**alft **M**aximum (FWHM) illustrated with the orange dashed line was calculated giving a distance of 0.86 µm. This result indicates that it is not possible to spatially resolve particles of size 500 nm. A similar experiment using a PMMA bead of 500 nm to calculate the FWHM has been done in a study by Zhang et al. [44] resulting in 0.61 µm.

4.2 Sectioned cells measured with O-PTIR

Figure 4.2 shows three separate O-PTIR measurements of sectioned *Rhodotorula graminis* where the spectra are corrected by weighted EMSC using a constant offset. As reference spectrum the average spectrum of *Rhodotorula graminis* is used. Figure 4.2A, B shows spectra of 500 nm sections of *Rhodotorula graminis* cells measured at different positions probing different cells, and Fig. 4.2C shows spectra of 200 nm section of *Rhodotorula graminis* cells. Different thicknesses of the sections were used to evaluate if the sample thickness has an effect on the quality of the obtained spectra. By comparing Fig. 4.2A, B with Fig. 4.2C it shows that the thickness of the sectioned cells does not effect the quality of the spectra. Further, in Fig. 4.2A the most significant absorbance peaks are annotated with a dashed line with the corresponding wavenumber. Due to the epoxy, the spectra do not only show chemical absorbance peaks due to absorption by the biological cells as described in table 3.2, but also chemical absorbance peaks due to absorption by the epoxy. Figure 4.2D shows the optical image corresponding to the spectra measured in Fig. 4.2C, where the blue line (see the red arrow) is the acquired line scan. The



(A)

(B)

(C)

(D)

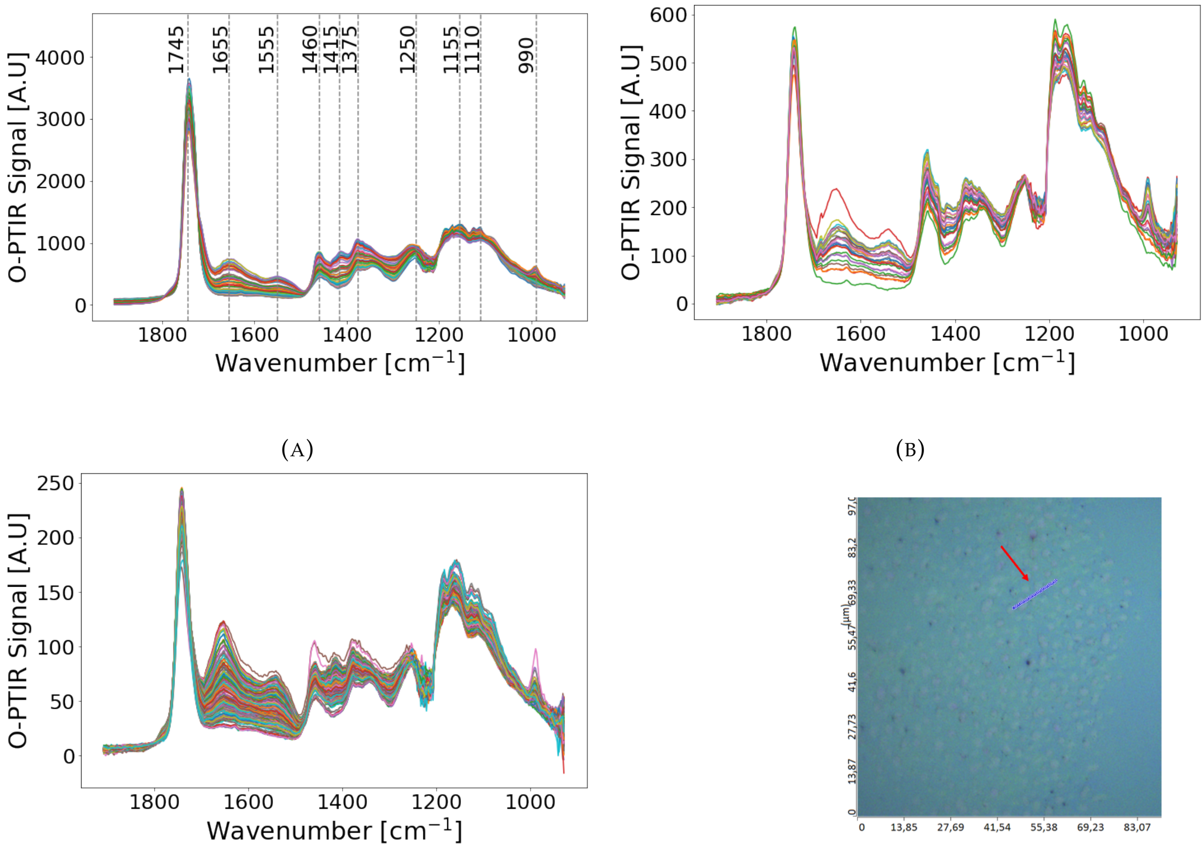
FIGURE 4.1: **(A)** Mid-infrared photothermal image of 500 nm polystyrene beads. **(B)** Line scan through five different beads. **(C)** The line scans are aligned according to the peak maximum. **(D)** Average line scan where the orange dashed line marks the FWHM = 0.86 µm

cells are visualized as light points, while the black points are artifacts. The image shows that it is not possible to get a good visualization of cells with a diameter less than 4 µm by looking at the optical image. However, the infrared response is good resulting in high quality absorbance spectra. The ratio of the absorbance band at 1745 cm^{-1} and the spectral range between 1200 cm^{-1} and 1000 cm^{-1} varies a lot when comparing the three different measurements in Fig. 4.2. This variation may be explained by the power of the IR laser. By increasing the IR laser power, the intensity of the bands at 1200-1000 cm^{-1} decreases as seen in Fig. 4.2A, and by decreasing the IR laser power the intensity of the same bands increases in intensity as seen in Fig. 4.2B. Thus, the IR laser power does affect the reflectivity of the green laser. However, it is not fully understood how it is affected. The background measurement can also be a reason for the observed variation. Generally, in conventional IR spectroscopy, the background spectrum is acquired before every measurement. However, in O-PTIR measurements this is not the practice. The instrumental setup has predefined instrumental parameters for the background, and commonly the background is only acquired once for every sample. Thus, when changing parameters such as the IR laser power or the probe power, the same background measurement is used, which may cause shifts in the absorbance spectra. A separate study should evaluate the relationship between the two lasers and the spectral response, and how the background measurements should be done.

The sectioned cells were embedded in epoxy which was necessary in order to obtain thin sections (200 - 500 nm) using a cryo-microtome. The cells are too soft and by embedding the cells with epoxy it turns into a polymerized block which makes it possible to make thinner sections. The presence of epoxy around and in the cell is expected to lead to contamination of the absorbance spectra with chemical signatures from epoxy. Therefore, it is important to know what the chemical absorbance spectra of pure epoxy look like in order to be able to identify epoxy signals in IR spectra and to be able to assess the degree of overlap of epoxy and cell signals in the spectra. Figure 4.3 shows EMSC corrected spectra of only epoxy where absorbance peaks are annotated with the respective wavenumbers. The spectra are corrected with a constant offset and the mean spectrum is used as the reference spectrum. There are many absorbance peaks of epoxy that overlap with the FTIR HTS spectra of intact *Rhodotorula graminis* cells from Fig. 3.5B, such as the peak at 1745 cm^{-1}, the two peaks at 1460 cm^{-1} and 1375 and the three peaks at 1250cm^{-1}, 1155 cm^{-1} and 1110 cm^{-1}. Thus, epoxy and *Rhodotorula graminis* share many of the same absorption bands, and the absorbance peaks in Fig. 4.2 show contributions from both *Rhodotorula graminis* and epoxy. However, there are some peaks in the spectra of sectioned *Rhodotorula graminis* in Fig. 4.2 that are not present in the epoxy spectra, meaning that the signals only come from chemical absorption of yeast and not epoxy. These are the amide I and amide II peaks at 1655 cm^{-1} and 1550 cm^{-1}, and the peaks at 1415 cm^{-1} and 990 cm^{-1}. The same peaks are present in the FTIR HTS measurements of intact *Rhodotorula graminis* in fig 3.5B, verifying that chemical absorption origins from the yeast cells. In order to obtain subcellular information of the cell wall and cell interior it is necessary to subtract the chemical signal origin from epoxy from the sectioned cells in fig 4.2. For the correction of spectra of the nano sections with respect to epoxy, it is important to evaluate if the pure epoxy spectra show variations. This may affect the choice of 'bad spectra' (i.e. interferent spectra) in the EMSC correction model. From Fig. 4.3 we can conclude that the epoxy spectra do not show significant chemical variation.

4.2.1 Removing epoxy from the sectioned cells

Figure 4.4A shows spectra of sectioned *Rhodotorula graminis* where the spectra are corrected with weighted EMSC using a polynomial order zero and the mean spectrum as the reference spectrum. In addition, the spectra were normalized with respect to the peak at 1745 cm^{-1}. Further, the black epoxy spectrum in Fig. 4.4A was subtracted from all spectra in Fig. 4.4A,



(A) (B)

(C) (D)

FIGURE 4.2: O-PTIR measurement of sectioned *Rhodotorula graminis*. The spectra are corrected with weighted EMSC with polynomial order zero. The mean spectrum is used as the reference. **(A)** Absorbance spectra of a 500 nm thin section, where the most prominent absorbance peaks are illustrated with dashed lines and the respective wavenumber. **(B)** Absorbance spectra of a 500 nm thin section. **(C)** Absorbance spectra of a 200 nm thin section. **(D)** Optical image of the line scan measured in C.

where the result is displayed in fig 4.4B. Negative peaks are observed in 1730 cm^{-1} and in the region between 1200 cm^{-1} and 1000 cm^{-1}, This indicates that there are chemical absorption bands of yeast in these regions, consequently to much signal has been subtracted. At the wavenumbers 1780 cm^{-1} and 1750 cm^{-1} two small peaks appears. They may partially be related to absorption by yeast cells. It turned out that it was not possible to obtain pure chemical absorbance signal of *Rhodotorula graminis* by subtracting the epoxy spectrum directly. Further data treatment was necessary.

A weighted EMSC model, with polynomial order zero, was performed on the same data as in Fig. 4.4. Instead of using the mean spectrum as a reference, the mean spectrum of intact *Rhodotorula graminis* cells measured with FTIR microspectroscopy imaging was used. An epoxy spectrum showed in Fig. 4.4A colored in black, was included as an interferent spectrum in the EMSC model. As we recall from section 2.6.2, an interferent spectrum contains chemical information that is unwanted in the spectra. Figure 4.5A shows the result of the EMSC correction, where some of the epoxy seems to be successfully subtracted. The peak at 1745 cm^{-1} is reduced and the height of the peak looks like the actual lipid peak that is present in yeast cells. We checked the Pearson correlation coefficient between the mean spectrum and

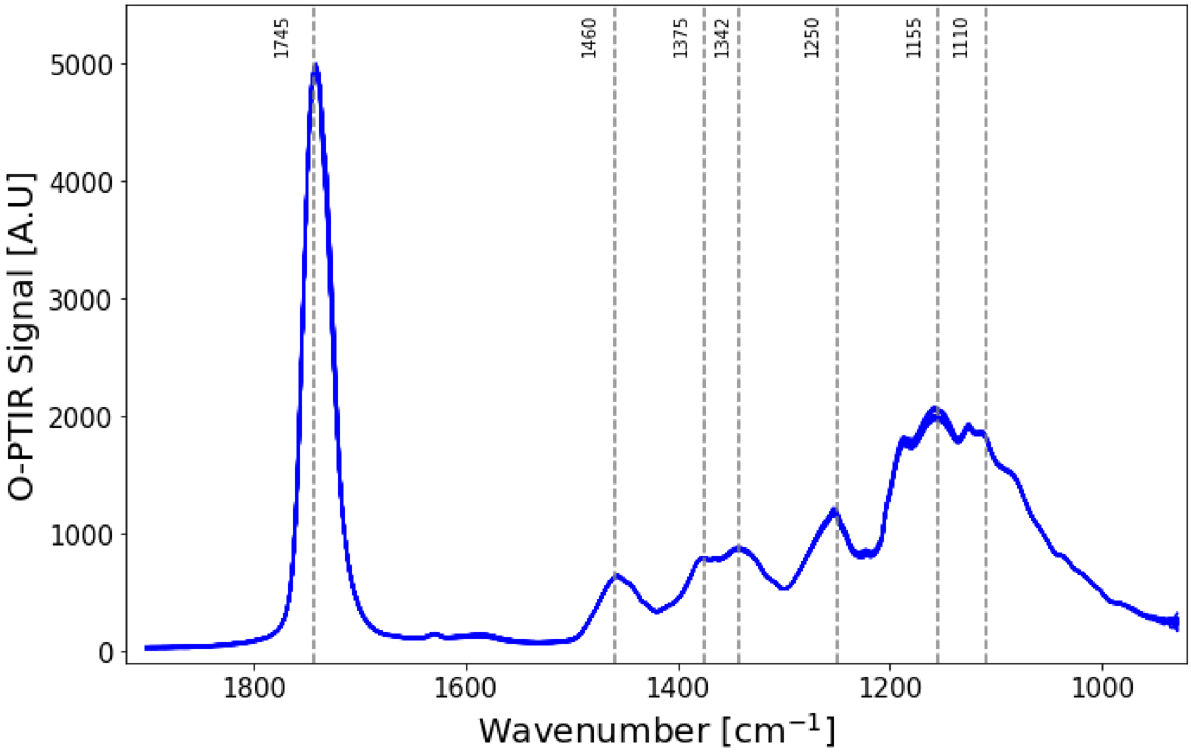
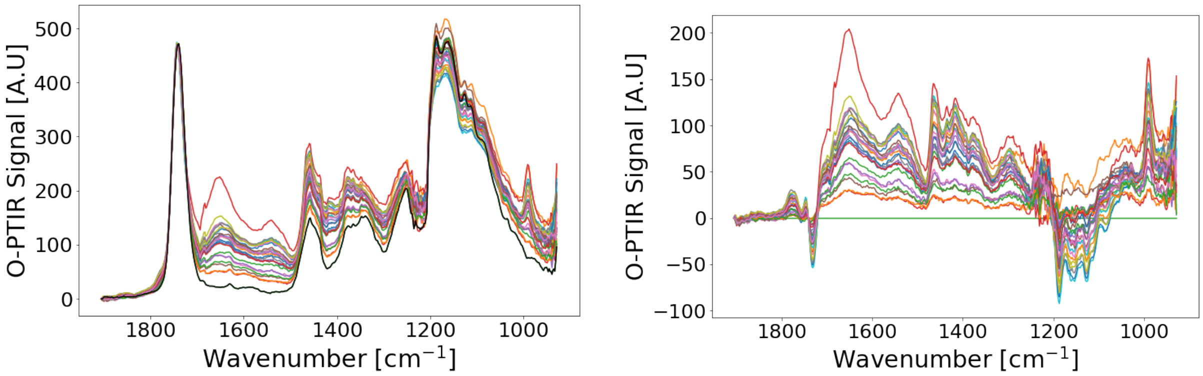


FIGURE 4.3: O-PTIR measurement of pure epoxy. The spectra are EMSC corrected with polynomial order zero and the mean spectrum is used as the reference. The spectra are obtained from different positions on a 500 nm sectioned film. Significant peaks are illustrated with dashed lines with the respective wavenumber.



(A) (B)

FIGURE 4.4: O-PTIR spectra of a 500 nm sectioned *Rhodotorula graminis*. **(A)** Weighted EMSC corrected spectra with polynomial order zero and the mean spectrum as reference. The spectra are in addition normalized with respect of the 1745 cm^{-1} peak. The black spectrum shows only signals of epoxy. **(B)** The black epoxy spectrum is is subtracted from all spectra.

the reference spectrum. It turned out that the Pearson correlation is 0.36. Therefore, we do not expect that there is any instability in the estimation of the multiplicative effect related to the reference spectrum and the parameter related to the interferent spectrum, i.e. epoxy spectrum.

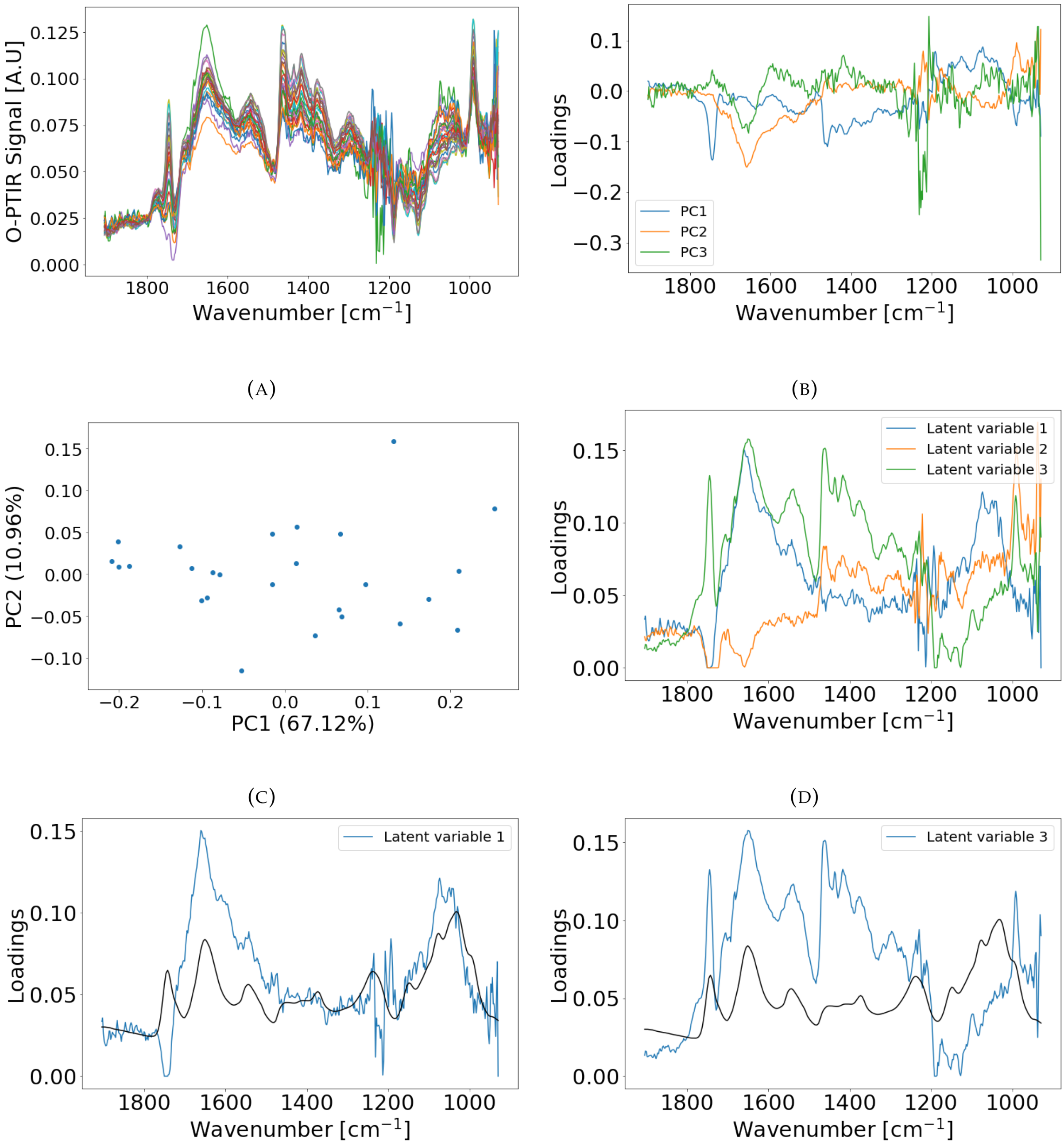
Further, a PCA analysis was performed using the EMSC corrected data. The total variance explained by the three first PCs are 83.25 %, where 67.12 % is explained by the first PC, 10.96 % is explained by the second PC, and 5.18 % is explained by the third PC. The loadings of the three first PCs are displayed in Fig. 4.5B. The blue spectrum visualizes the loading of PC1. It has a significant peak at 1745 cm^{-1} where it shows strong signals both from epoxy and lipids from yeast. There are two peaks at 1655 cm^{-1} and 1555 cm^{-1}, which do not contain any signals from epoxy, only signals from proteins. Further PC2, shows strong signals in the protein region and it has a significant peak at 990 cm^{-1}. Figure 4.5C shows the score plot for the first and second PCs along the x-axis and the y-axis, respectively. It is not possible to draw any conclusions from the score plot, which might be because the spatial resolution most likely is not good enough to get signals from pure cell wall and cell interior of samples with a size of 2-3 µm such as *Rhodotorula graminis*.

Further, an MCR analysis is performed with non-negative constraints. For the MCR analysis, the corrected data from the EMSC correction presented in Fig. 4.5A and the scores from the PCA analysis in Fig. 4.5C were used for initializing the MCR algorithm. The resulting three latent variables of the MCR analysis are shown in Fig. 4.5D, where the blue, orange and green spectra are the first, second and third latent variables of the MCR analysis, respectively. In Fig. 4.5E and F the first and third latent variables are shown in blue together with the mean spectrum of the intact yeast cell which was used as the reference spectrum in the EMSC model. Considering the spectral range below 1400 cm^{-1} for the first latent variable and the range above 1400 cm^{-1} for the third latent variable, they share a lot of the same signals as the mean spectrum. Looking closer at the peak at 1745 cm^{-1} and the two peaks at 1650 cm^{-1} and 1550 cm^{-1}, the ratio between the peaks looks accurate in comparison to the mean spectrum of intact *Rhodotorula graminis*. It seems like the epoxy has been successfully extracted from this region. However, the epoxy could not be removed successfully for all spectral data of the same sample following this preprocessing approach. In most of the cases, the contribution from epoxy to the spectra was so strong that it has not been possible to remove the contribution from epoxy from the spectra of sectioned yeast cells. The epoxy signal has a strong influence on the spectra, making the data analysis hard. Even though we were able to remove some of the epoxy signals, it was not possible to distinguish clearly the cell wall and the cell interior apart from each other for the O-PTIR data.

O-PTIR measurements of sectioned algae are displayed in Fig. 4.6. It was not possible to obtain good spectra for the algae and most of the measurements resulted in noisy spectral data that was difficult to interpret (not displayed). The algae samples have high roughness and thick cell wall, which may have caused issues when cryo-microtome the cells. The algae were also measured with AFM-IR, with the same poor result. It is not clear why it was not possible to obtain high quality spectra such as of yeast. Due to these results, no further investigation regarding algae has been done.

4.2.2 Sectioned cells measured with AFM-IR

A section of *Rhodotorola Graminis* was measured by AFM-IR. The result is displayed in Fig. 4.7. Figure 4.7A shows the absorbance spectrum of only epoxy in orange and an arbitrary absorbance spectrum of the sectioned cell. The spectra are corrected by weighted EMSC with polynomial order zero and with the mean spectrum as the reference spectrum. The two spectra share the same absorbance peaks at 1745 cm^{-1}, 1460 cm^{-1}. However, the intensities of the peaks are different. For instance at 1745 cm^{-1}, the epoxy spectrum has an intensity of approximately 24 mV, and the cell spectrum has an intensity of approximately 12 mV. A topographical images of the measured cell is shown in Fig. 4.7C. The color bar shows the height of the cell and it is a clear contrast of the cell wall and the cell interior. Figure 4.7B shows spectra of the cell with the same EMSC correction as 4.7A. The spectra do not show considerable chemical absorption in the carbohydrate region 1200-1000 cm^{-1}, even though the *Rhodotorola Graminis*



(A) (B) (C) (D) (E) (F)

FIGURE 4.5: O-PTIR spectra of 500 nm sectioned *Rhodotorula graminis*. **(A)** Weighted EMSC corrected with polynomial order zero (see the main text for information on the reference and the interference model spectra). **(B)** The loading plot from PCA analysis. The three first principal components are displayed. **(C)** Score plot from PCA analysis with the first and second principal component along the x-axis and y-axis, respectively. **(D)** The three first latent variables of the MCR model, which are obtained by MCR of corrected EMSC spectra and the scores from the PCA analysis. **(E)** and **(F)** show the first and third latent variables from the MCR, respectively. The black spectrum is the mean spectrum of *Rhodotorula graminis* measured with FTIR microspectroscopic imaging.

cell is expected to show chemical absorption in this area, as expected from the HTS FTIR spectra in Fig. 3.5B. The reason why there is no signal from yeast visible in this region is most likely

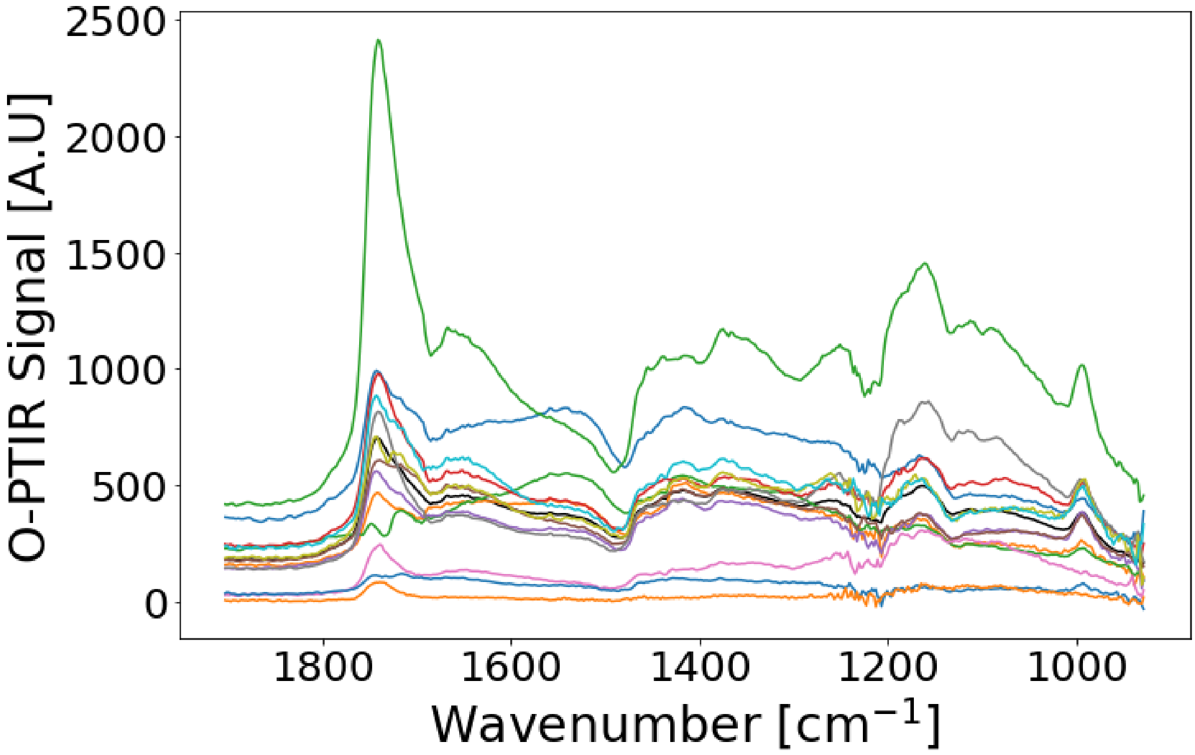


FIGURE 4.6: Raw spectra of 500 nm sectioned algae *Aurautiochyrium limaanium*, measured with O-PTIR.

because the epoxy signal is strong in this region. Thus, important signals of carbohydrates in the cell wall are thereby lost because of the dominance of epoxy. Therefore, nano sections that are obtained by embedding the cells in epoxy are not preferable if the interest is to obtain absorption signals in the range below 1200 cm^{-1}. The AFM-IR signal in this range is in addition very low, because of the power of the laser.

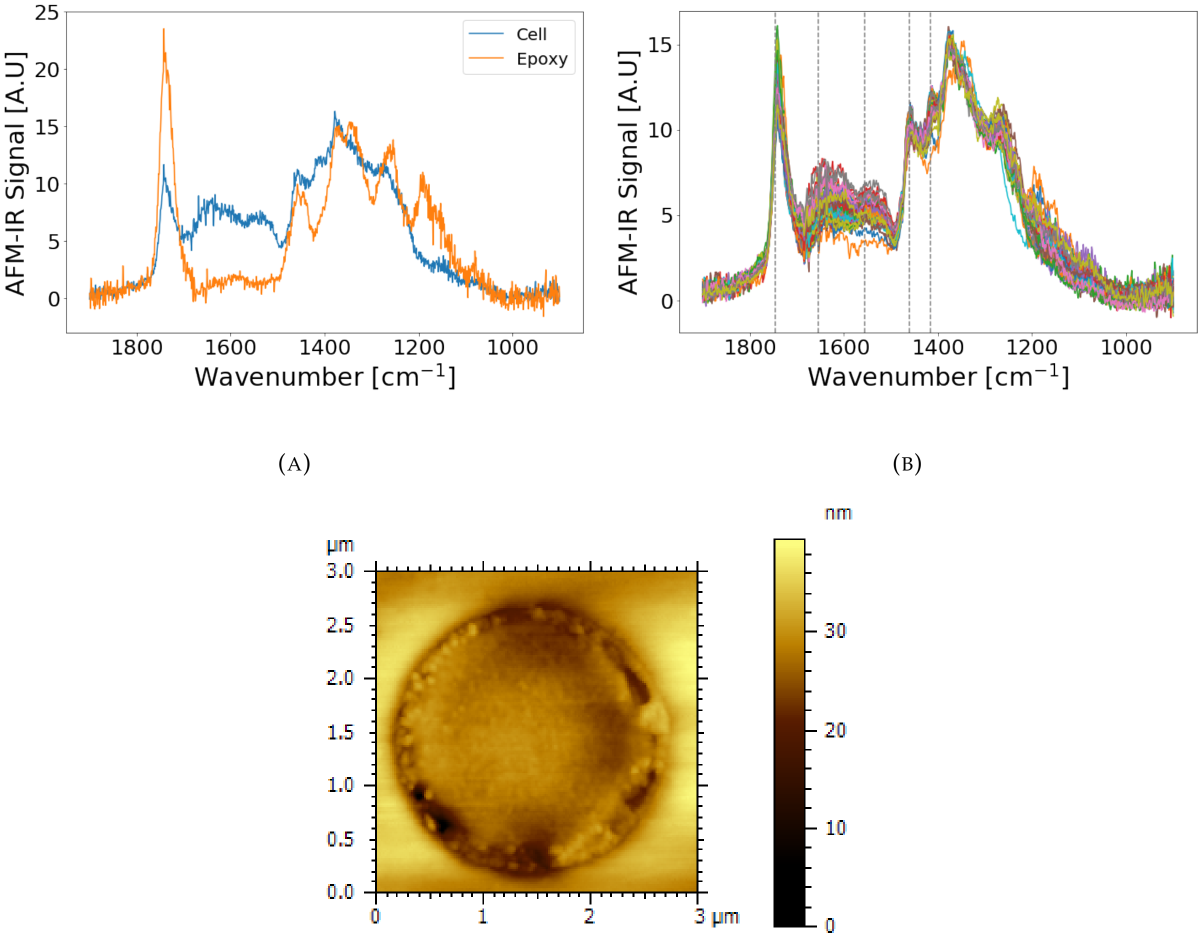
Five significant peaks are presented in Fig. 4.8 as heat maps and the five peaks are shown as dashed lines in Fig. 4.7C at the wavenumbers 1745 cm^{-1}, 1655 cm^{-1}, 1555 cm^{-1} 1460 cm^{-1} and 1415 cm^{-1}, respectively. Further, the heat maps in Fig. 4.8 show the integrated area across these five peaks which are as follows: 1770-1690 cm^{-1}, 1575-1685 cm^{-1}, 1495-1580 cm^{-1},1470-1450 cm^{-1} and 1400-1425 cm^{-1}. Each pixel in the heat map represents the integration of the respective peak area for each spectrum in the pixel. The color illustrates the intensity of the peaks, with yellow as the maximum intensity and dark blue as the minimum intensity. In Fig. 4.8A the spectra are smoothed with a Savitzky Golay filter using a window size of three. In Fig. 4.8B the spectra were in addition corrected with a weighted EMSC with polynomial order zero and using the mean spectrum as the reference spectrum. The reason why we correct the spectra with a weighted EMSC is to evaluate if the contrast of the intensity between the cell wall and cell interior increases. The integral across the 1745 cm^{-1} peak has a maximum value of 500 and a minimum value of 200 for both Fig. 4.8A and 4.8B. The heat map does not show a significant difference in the intensity of the cell interior and the cell wall, even though we expect more lipids in the cell interior. As seen in Fig. 4.7A, the epoxy signal at 1745 cm^{-1} is strong and interferes with the lipid signal, which makes it difficult to separate the cell wall and the cell interior with respect to the lipid peak at 1745 cm^{-1}. However, the heat map without EMSC correction shows a slightly higher intensity in the cell interior and it seems that the lipid content varies, anyways it is not possible to separate the cell wall and cell interior by looking at the 1745 cm^{-1} peak. The integration across the 1655 cm^{-1} peak has a maximum value of

220 and a minimum value of 40 for both Fig.4.8A and 4.8B, and it shows that there is a higher content of protein (amide I) in the cell wall than in the cell interior. The contrasts are seen in both Fig. 4.8A and 4.8B. However, the signal is stronger in the EMSC corrected heat map. The three last heat maps which show integrated peak areas across the peaks 1555 cm^{-1}, 1460 cm^{-1} and 1415 cm^{-1}, have maximum and minimum values equal to 90 and 10, 18 and 6, and 28 and 2, respectively for both Fig. 4.8A and Fig. 4.8B. All three heat maps show strong signals in the cell interior and low signals in the cell wall, whereas the EMSC corrected heat maps show slightly improved contrasts. The peak at 1555 cm^{-1} is assigned to protein (amide II) and 1460 cm^{-1} to lipids. The assignment of the peak at 1415 cm^{-1} is not completely clear. It may be due to the contribution of several chemical compounds, such as lipids and chitin. By integrating the area across significant peaks it makes it possible to successfully distinguish the cell wall and the cell interior. Heat maps obtained from spectra with and without EMSC correction show this trend. The contrasts in intensity are slightly improved when the spectra are EMSC corrected, making it easier to distinguish the cell wall from the cell interior. The sectioned cell is uniform in thickness and the AFM-IR is probing the same depth into the cell which means that we do not expect a large variance in scaling. Thus, the EMSC correction will only slightly change the spectra. In conclusion, integrating the peak area and presenting it as a heat map is a valid approach to analyzing and visualizing the data of sectioned cells embedded with epoxy, making it possible to distinguish the cell wall and the cell interior. Hence, with this approach, we may potentially see different chemical compounds in the cell wall and cell interior which provide valuable insight into cell structure and chemistry.

The same approach using EMSC, PCA and MCR as illustrated for the O-PTIR data in Fig. 4.5 was applied to the data from AFM-IR. However, it was not possible to successfully distinguish the cell wall and the cell interior with this approach. The main problem with spectra obtained from sections prepared from epoxy embedded cells is the strong absorption of epoxy in the same spectral regions that show strong chemical absorption of fungal cells. It is difficult to successfully subtract the epoxy from the apparent absorbance spectra with EMSC. However, scientists have successfully subtracted samples embedded with paraffin and water from the spectra using EMSC [40], for samples whose characteristic absorbance signals do not overlap considerably with the interferent signals. It is therefore a reasonable hypothesis that it is possible to remove the epoxy with EMSC correction and other preprocessing models. However, more time and resources are necessary to follow up on this hypothesis.

4.3 depth-resolved O-PTIR measurements of intact cells

Figure 4.9A shows depth-resolved apparent absorbance spectra of intact *Phaffia rhodozyma* cell measured at the same point with different Z-focus (point scan measurement). The sample stage is first placed such that the focus point is first located at the top of the cell before it is gradually raised with a step size of approximately 0.5 µm. The spectra are laser corrected and in Fig. 4.9B, as well as being baseline corrected where the mean offset value of the spectra in the region above 1800 cm^{-1} is subtracted from the spectra. Figure 4.10A visualizes the development of the baseline corrected spectra with 30 different Z-foci. The first spectrum is marked as number one, indicating that the focus is on top of the cell, and spectrum 30 is probing below the cell. Further, Fig. 4.10A shows that the signal gradually increases in magnitude, and reaching maximum intensity at spectrum 12, the intensity is then gradually decreasing until zero intensity. Spectrum one and the spectra from 26 to 30 do not contain any signal, meaning that the cell is not probed. The main variation is seen in the two protein peaks, amide I and II, and in the carbohydrate region between 1200 cm^{-1} and 1000 cm^{-1}. Figure 4.10B shows the integration across the whole spectral range (1903-929 cm^{-1}), which gives the total intensity as a function of the Z-focus. The total intensity is low for the first Z-focus and gradually increases



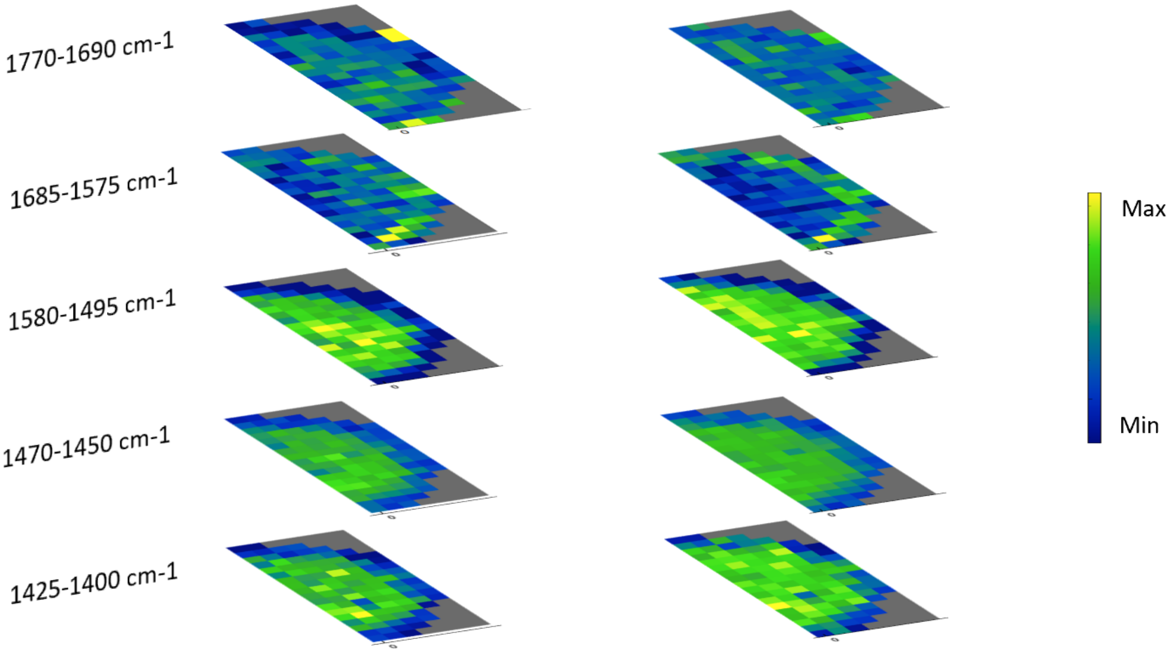
(A)

(B)

(C)

FIGURE 4.7: Measurement of a 500 nm sectioned yeast cell *Rhodotorola Graminis* recorded with AFM-IR. The spectra are corrected by weighted EMSC with the mean spectrum as a reference and polynomial order zero. **(A)** The orange spectrum illustrates only epoxy and the blue spectrum is a spectrum obtained from a cell embedded with epoxy. **(B)** Spectra of a cell embedded with epoxy, where important signals from yeast are marked with dashed lines. **(C)** A visual image of the recorded cell, where the color bar gives the height of the cell.

until it reaches a maximum at spectrum 12, where the intensity starts decreasing as the Z-focus increases further. Assuming that the maximum total intensity is in the middle of the cell, i.e. spectrum 12, and knowing that the cell has a diameter between 6-10 µm, the cell is between spectrum number 2 and number 22. This corresponds well with Fig.4.10A as we see that the main signal is in this region. In Fig. 4.10C, the integral across the lipid band at 1745 cm^{-1} (1767-1731 cm^{-1}) is shown for each Z-focus. It shares a similar characteristic as the total intensity displayed in Fig. 4.10B, only the maximum intensity reaches approximately 800 compared to 150000 for the total intensity. It is expected that the shapes are similar since the lipid signal is normally higher in the cell interior and lower in the cell wall. Figure 4.10D shows the ratio of the integrated area across the peak at 1745 cm^{-1} (1767-1731cm^{-1}) and the peak at 1655 cm^{-1} (1706-1589 cm^{-1}). The ratio is approximately between 0.06 and 0.04 for the Z-focus 2 to 5, probing the cell wall. Further from Z-focus between 6 and 12, the ratio decreases between 0.02 and 0.4 where the protein content is increasing compared to the lipid content. In this region,

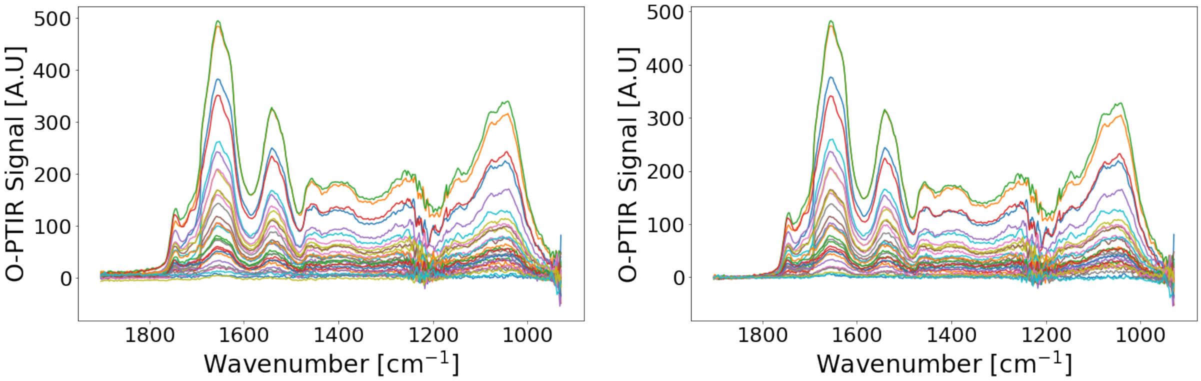


(A) (B)

FIGURE 4.8: Heat maps of the integration of five different peak bands from the same measurement as Fig.4.7. The color bar for each peak band shares the same axis for both **A** and **B** and the max and min values are: 1745 cm^{-1}: 500-200, 1655 cm^{-1}: 220-40, 1555 cm^{-1}: 90-10, 1460 cm^{-1}: 60-6, 1415 cm^{-1}: 28-2. **(A)** Spectra are preprocessed With a Savitzky Golay filter, window 3. **(B)** Spectra are corrected by weighted EMSC with polynomial order zero and the mean spectrum as a reference. In addition, the spectra are preprocessed with the Savitzky Golay filter, window 3.

the laser is probing the cell interior. Moreover, the ratio increases as the Z-focus increases to 25, meaning that the lipid content increases compared to the protein. This result shows that the lipid content is at its minimum when reaching the maximum intensity of the cell, compared to the amide I peak. Furthermore, as the laser probes through the sample and comes closer to the cell wall, the lipid content is higher compared to the amide I peak. In general, it is expected that the ratio should be at its highest at maximum total intensity, which is in the middle of the cell. However, it is not known exactly how the morphology of the cell is when placed on the slide, and the lipid body might lay closer to the cell wall along the sample slide and not exactly in the middle of the cell. In addition, it is not certain what the laser probes when the focus point is deep inside the cell. The signal will consist of contributions from the whole laser path and not only from the focus point.

Figure 4.11A shows depth-resolved spectra of *Mucor circinelloides*, where the point scan is of a single cell. Further, in Fig. 4.11B the spectra are laser corrected and the same baseline correction as previously described was performed. The same spectra are visualized in Fig. 4.12A, where the first spectrum is obtained with a Z-focus above the cell and where the Z-focus gradually decreases in depth with approximately 0.5 µm steps. In contrast to the measurements obtained for the Phaffia cell, we observe a large variation of the intensity regarding the lipid peak at 1745 cm $^{-1}$, in addition to the two protein peaks amide I and II. Some of the spectra do not show strong absorption signals even though the focus point is inside the cell, such as for example spectrum 11 and spectrum 21. Further, some of the amide I peaks are noisy which may occur due to saturation. However, in this case, we expect that the lipid peak at 1745 cm



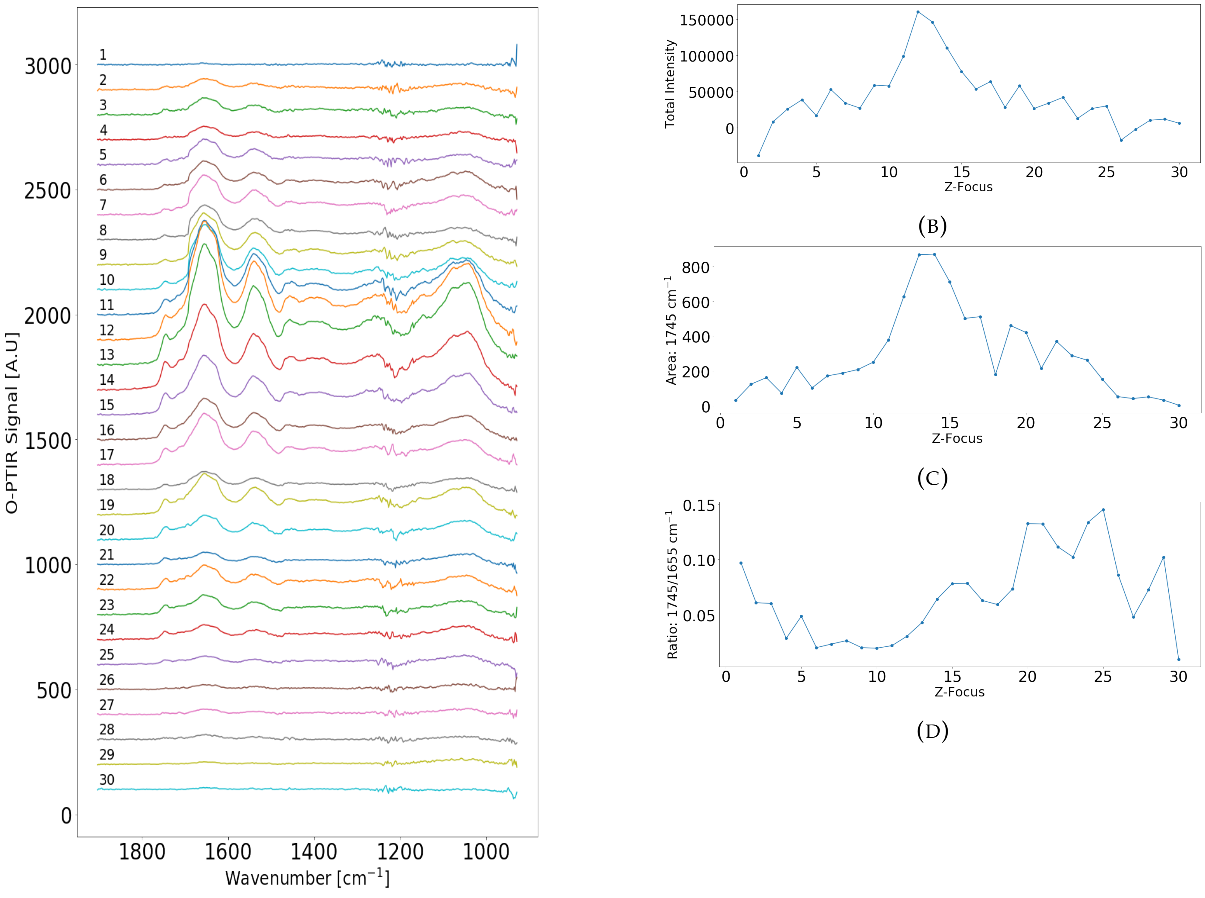
(A)

(B)

FIGURE 4.9: O-PTIR depth-resolved measurement of intact *Phaffia rhodozyma* with a step size of approximately 0.5 µm of the Z-focus. **(A)** Raw spectra. **(B)** Baseline corrected spectra.

$^{-1}$ shows saturation as well, which is not the case. These spectra may therefore be considered outliers and the discussed noise as artifacts. Figure 4.12B shows the total intensity as a function of the Z-focus. As expected, the intensity increases as the focal point reaches the cell interior and decreases as the focus reaches the lower cell wall. The maximum intensity is obtained for spectrum 23 with a total intensity exceeding 400000. The high total intensity value may indicate that the measured cell is thick, compared to the cell of Phaffia which showed less total intensity. In addition, the dependency of the intensity on the Z-focus is an indication of how big the cell is. Fig. 4.12C shows the integration across the 1745 cm^{-1} lipid area (1775-1718 cm^{-1}) which shows increasing signals in the cell interior and decreasing as the focus reaches the cell wall. However, Z-focus 15 to 20 shows decreasing amount of lipids, which should not be the case. Looking at the spectra with regard to the Z-focus, we see that for all spectra the amide peak is saturated or noisy. Fig. 4.12D shows the ratio between the 1745 cm^{-1} lipid peak and amide I band (1719-1603 cm^{-1}), where the ratio is more or less constant, except for a few points which can be considered as outliers. Lastly, Fig. 4.12E shows the ratio of the lipid peak at 1745 cm $^{-1}$ and the total intensity. However, the dependency of this ratio on the Z-focus is not conclusive. Fig. 4.13A shows the same as Fig. 4.12A, but the outliers are removed. The missing samples identified as outliers appear as empty space in Fig. 4.12A and as missing points in Fig. 4.13B-E. As for Fig. 4.12, the dependency of the peak areas and the ratios from the Z-focus is not conclusive.

Figure 4.14A shows depth-resolved spectra of *Mucor circinelloides*, where the point scan is focused on a hypha. The spectra are laser corrected, while a significant laser jump at 1687 cm^{-1} can still be seen. Different wavenumbers close to the laser break were tested in case the laser break was shifted, without succeeding in correcting the laser shift. In addition, the laser correction on the O-PTIR software was tested in an attempt to achieve better correction. However, a better correction could not be achieved. Since an improvement of the laser correction procedure was out of the scope of this thesis, no further attempt to improve the laser correction was made. An in-depth understanding of how the laser correction works should be urgently considered in further research on O-PTIR data. Fig. 4.14B shows the baseline corrected spectra, and Fig. 4.15A visualizes the spectra as the Z-focus penetrates through the hyphae. In total, the 18 different Z-scans indicate a thin sample. The chemical signals mainly occur in the two protein bands amide I and amide II, and in the carbohydrate region. In addition signals from the 1250 cm^{-1} peak were also prominent. The main difference between the hyphae and the cell is that the hyphae do not show any signals from lipids. The total intensity for each Z-focus is



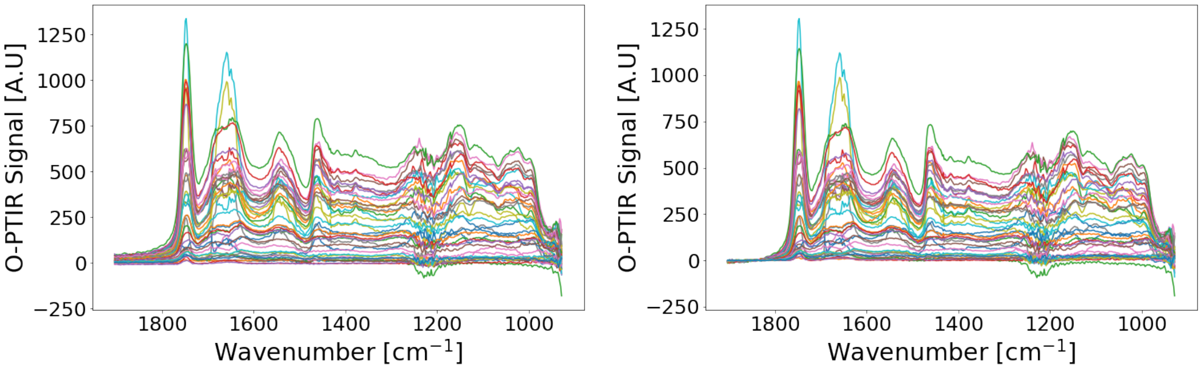
(A)

(B)

(C)

(D)

Figure 4.10: O-PTIR depth-resolved measurement of intact yeast *Phaffia rhodozyma*. **(A)** Baseline corrected spectra where spectrum 1 has the focus on the top of the cell and the spectra gradually decrease in depth with approximately 0.5 µm until spectrum 30 which is probing below the cell. **(B)** Integrated across 1903-929 cm^{-1} showing the total intensity along the y-axis and each Z-focus along the x-axis. **(C)** Integrated across 1767-1731 cm^{-1}. **(D)** The ratio of the integrated area 1767-1731 cm^{-1} and 1706-1589 cm1^{-1}.



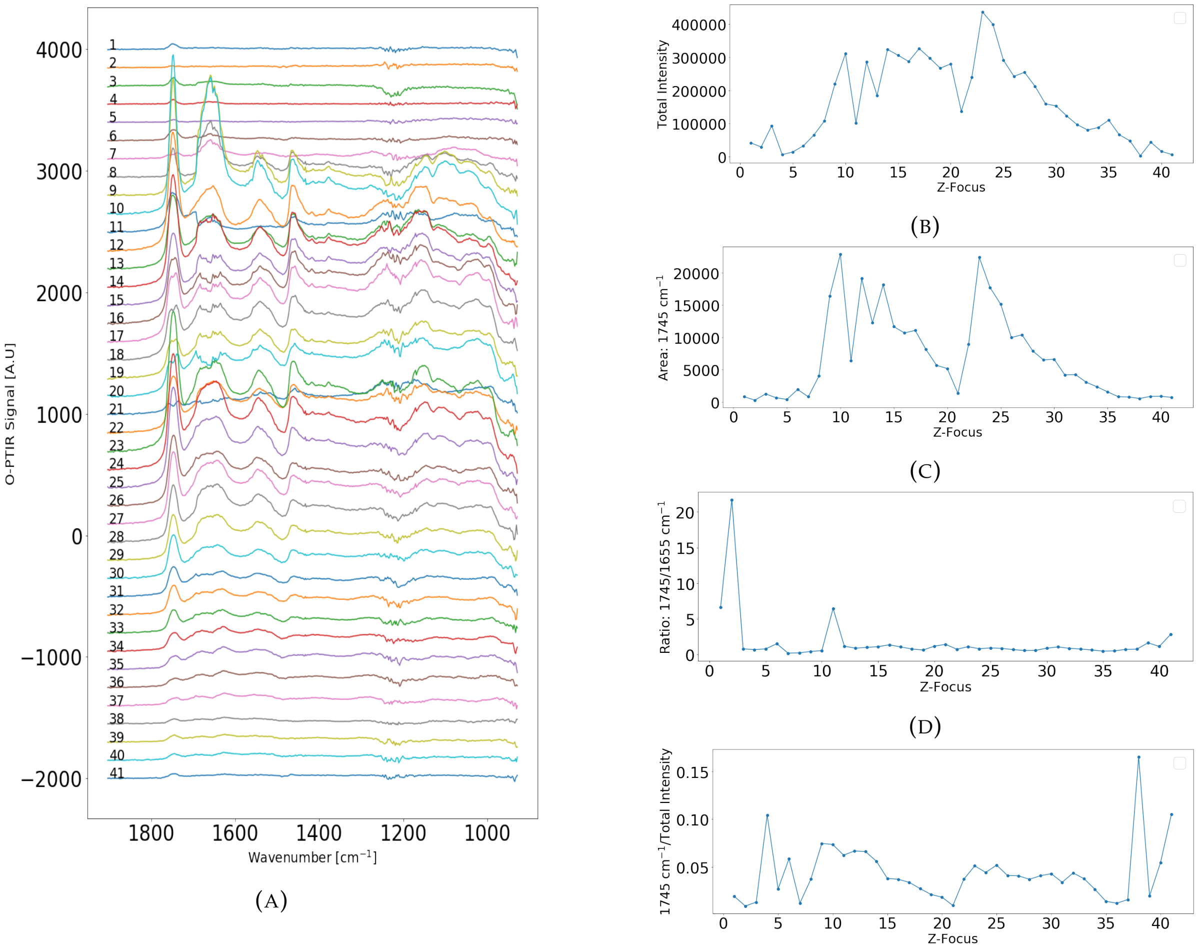
(A)

(B)

FIGURE 4.11: O-PTIR depth-resolved measurement of intact fungi *Mucor circinelloides* focused on a cell, with a step size of approximately 0.5 m of the Z-focus. **(A)** Raw spectra. **(B)** Baseline corrected spectra.

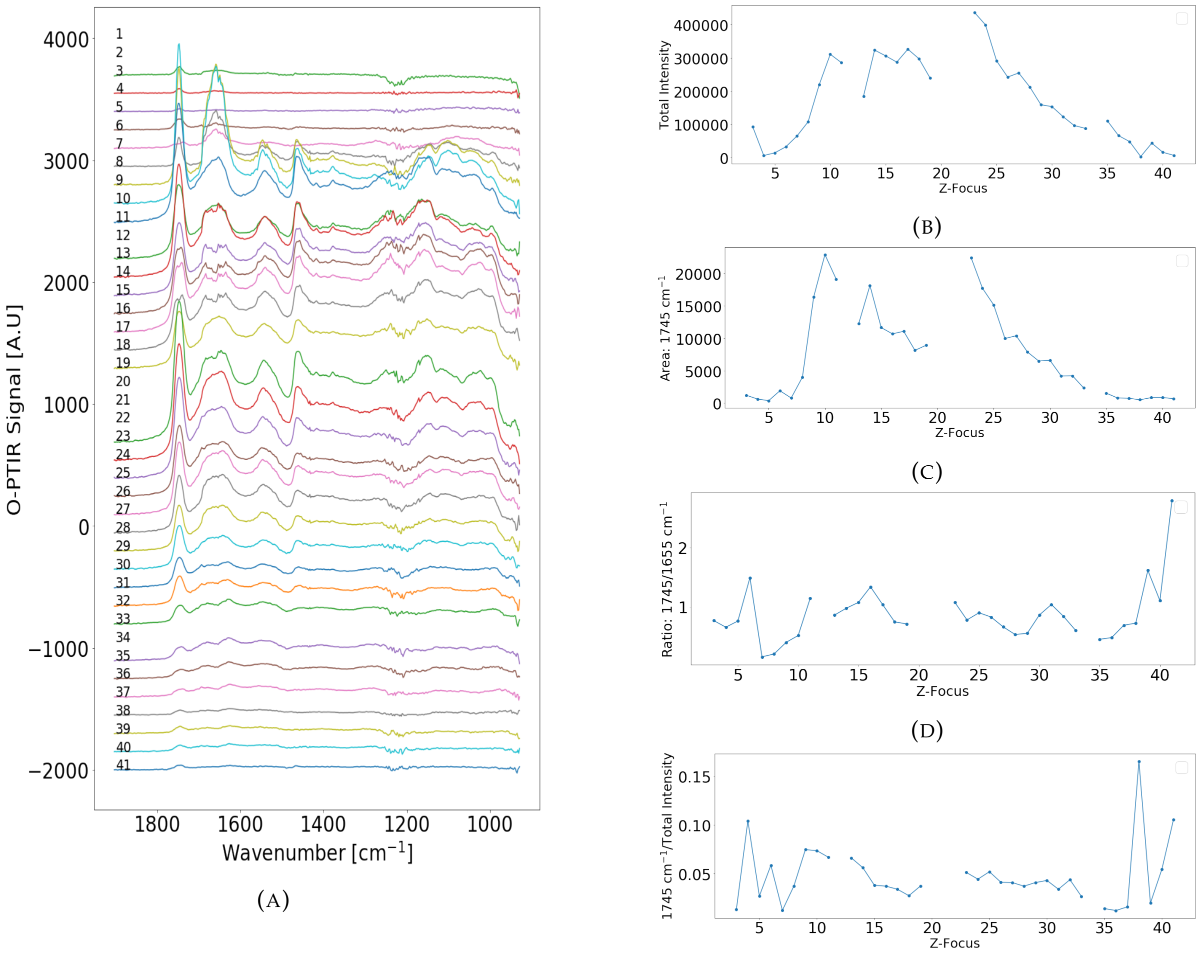
displayed in Fig. 4.15B, where it has been integrated across the whole spectral range (1903-929 cm^{-1}). In contrast to the depth-resolved spectra of Phaffia and Mucor as shown in Fig. 4.10 and Fig. 4.13, the total signal does not show a clear development. One reason why it may be more difficult to achieve signals of the cell wall and the cell interior is that the hyphae are thin, making it difficult to probe the cell wall and the cell interior separately.

Summarizing the results for the depth-resolved point scans of *Phaffia rhodozyma* and *Mucor circinelloides* (both a cell and hyphae), we can conclude that it is possible to obtain signals from different depths within the cell. The signal is gradually increasing as the focus reaches the cell interior and then gradually decreases as the focus reaches the bottom of the cell. For Phaffia and Mucor cells it was possible to see higher total intensity and lipid content in the middle of the cell, but for the hyphae it was not possible to obtain conclusive results when looking at the total intensity. The obtained results may indicate that in-depth measurements of O-PTIR do not allow to obtain highly sensitive in-depth-resolved signals. We may rather conclude that the signal could have contributions of the surrounding area of the focus point. For instant when probing with different Z-focus, the green laser penetrates different depths in the cell meaning it is not possible to only get information of the focus area but we most likely also get contributions along the whole laser path. As described in the section 2.5.2, it is the change in reflectivity of the green laser which is detected and translated into IR absorbance spectra. It is not understood how the reflectivity is affected by changing the Z-focus for the green laser from the surface of the sample to within the sample. With today's knowledge of the optical system and the physics behind it, it is not possible to know exactly how much of the signal is obtained from the actual focal point and what comes from the surrounding. In a study from Zhang et al. [44] depth-resolved images are performed, where they show depth-resolved signals of a lipid droplet and protein. The measurement from this paper is done in transmission mode of a PC-3 cell, which is a cancer cell. The cell is considerably bigger than the cells measured in this thesis, however, the thickness of the sample described by Zhang et al. [44] has a maximum thickness of 10 µm. Zhang et al. [44] are in total measuring 29 µm in depth, meaning that they are not only probing the cell but also far outside. Based on the result of this thesis, we conclude that when measuring in transmission mode as done by Zhang et al. [44], the laser collects signals from the whole cell and we expect contributions to the acquired signal along the whole laser path. We conclude that with current setups it is not possible to obtain clearly spatially resolved spectra with depth-resolved spectra with O-PTIR of subcellular chemical information of the cell. Further knowledge of the optical system and the physics of it is necessary.



(A) (B) (C) (D) (E)

FIGURE 4.12: O-PTIR depth-resolved measurement of intact *Mucor circinelloides* focused on a cell. **(A)** Baseline corrected spectra where spectrum 1 has the focus on the top of the cell and the spectra gradually decrease in depth with approximately 0.5 m until spectrum 41 which is probing below the cell. **(B)** Integrated across 1903-929 cm^{-1} showing the total intensity along the y-axis and each Z-focus along the x-axis. **(C)** Integrated across 1775-1718 cm^{-1}. **(D)** The ratio of the integrated 1775-1718 cm^{-1} and 1719-1603 cm^{-1}. **(E)** The ratio of the integrated 1775-1718 cm^{-1} peak and the total intensity.



(A)

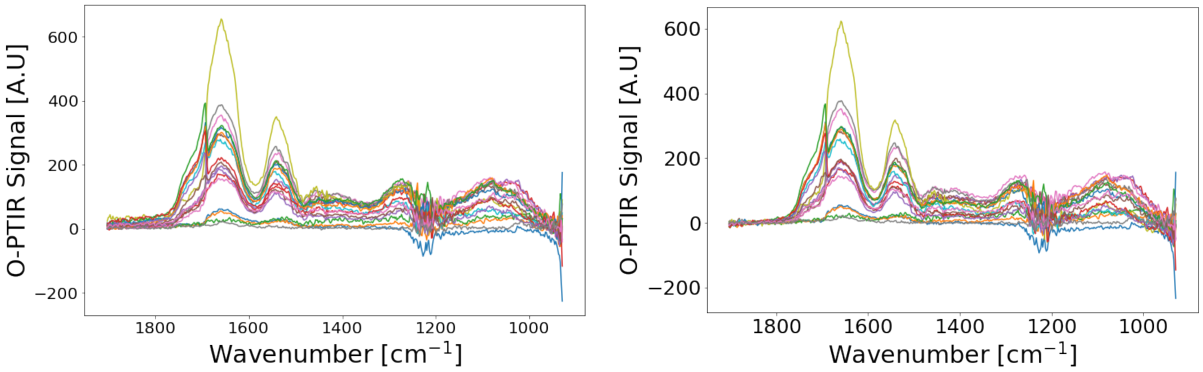
(B)

(C)

(D)

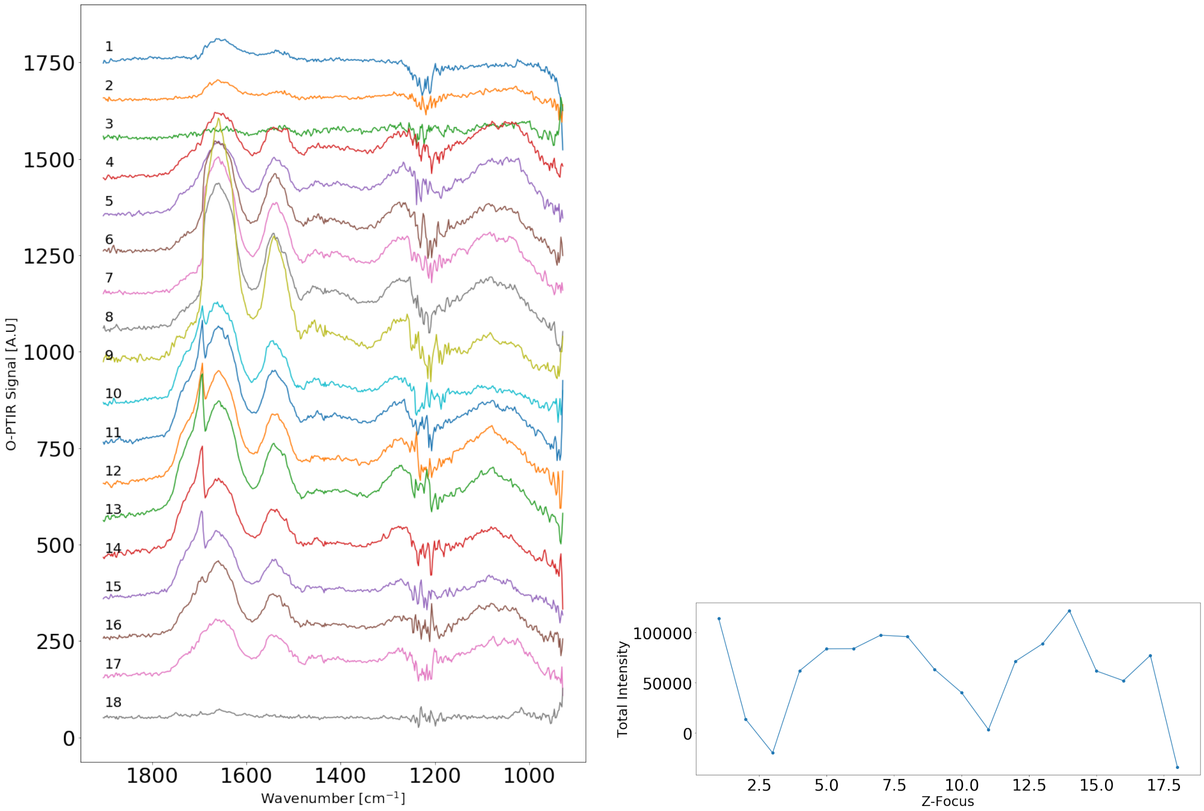
(E)

FIGURE 4.13: O-PTIR depth-resolved measurement of intact fungi *Mucor circinelloides* focused on a cell, where the outliers are removed and visualized with an empty spot. **(A)** Baseline corrected spectra where spectrum 1 has the focus on the top of the cell and the spectra gradually decrease in depth with approximately 0.5 m until spectrum 41 which is probing below the cell. **(B)** Integrated across 1903-929 cm^{-1} showing the total intensity along the y-axis and each Z-focus along the x-axis. **(C)** Integrated across 1775-1718 cm^{-1}. **(D)** The ratio of the integrated 1775-1718 cm^{-1} and 1719-1603 cm^{-1}. **(E)** The ratio of the integrated 1775-1718 cm^{-1} peak and the total intensity.



(A) (B)

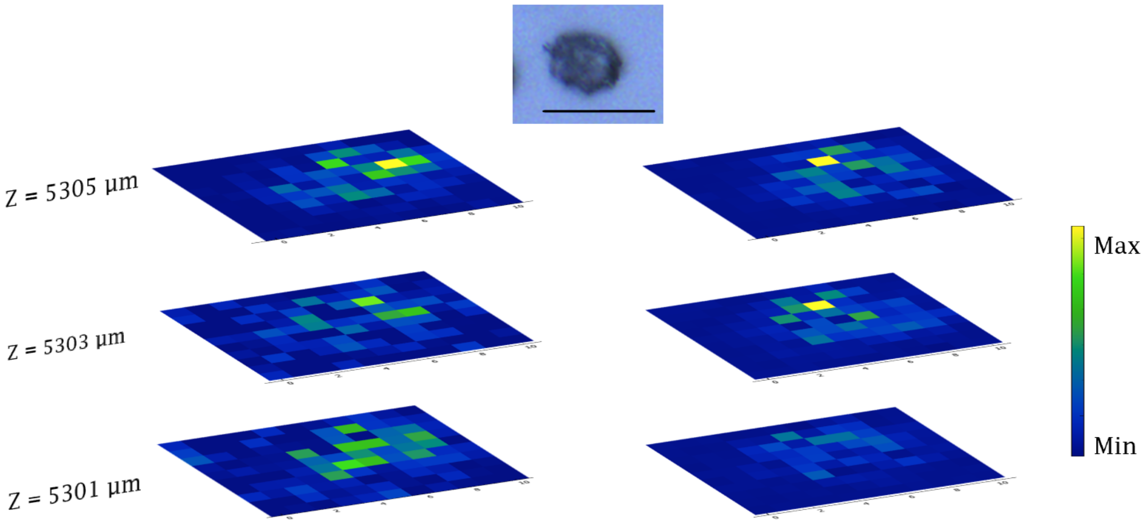
FIGURE 4.14: O-PTIR depth-resolved measurement of intact *Mucor circinelloides* focused on a hypha with a step size of approximately 0.5 µm of the Z-focus. **(A)** Raw spectra. **(B)** Baseline corrected spectra.



(A) (B)

FIGURE 4.15: O-PTIR depth-resolved measurement of intact *Mucor circinelloides* focused on a hyphae. **(A)** Baseline corrected spectra where spectrum 1 has the focus on the top of the cell and the spectra gradually decrease in depth with approximately 0.5 µm until spectrum 18 which probes below the cell. **(B)** Integrated across the whole spectrum (1903-929 cm^{-1}) showing the total intensity along the y-axis and each Z-focus along the x-axis.

In addition to the depth-resolved point scan, depth-resolved spectral images of *Phaffia rhodozyma* were measured with the O-PTIR, with three different Z-foci. In Fig. 4.16, the first focus was obtained at 5305 µm and then decreased with 2 µm, and the second and third spectral image was obtained at 5303 µm and 5301 µm, respectively. The images were both laser-corrected and baseline-corrected as described for the depth-resolved point scans. Figure 4.16 shows the heat maps of the spectra, where Fig. 4.16A illustrates the integration across the whole spectrum (1903-929 cm^{-1}), showing the total intensity, with a maximum value of 80000 and a minimum value at 0. Further Fig. 4.16B shows the integration across the 1745 cm^{-1} peak band, with a maximum value at 3000 and a minimum value equal to 0. A visual image of the cell is shown in the middle at the top of Fig. 4.16, where the scale bar corresponds to 13.85 µm. From the maps, it is not possible to see a clear separation between the cell wall and the cell interior.



(A)
Total intensiy

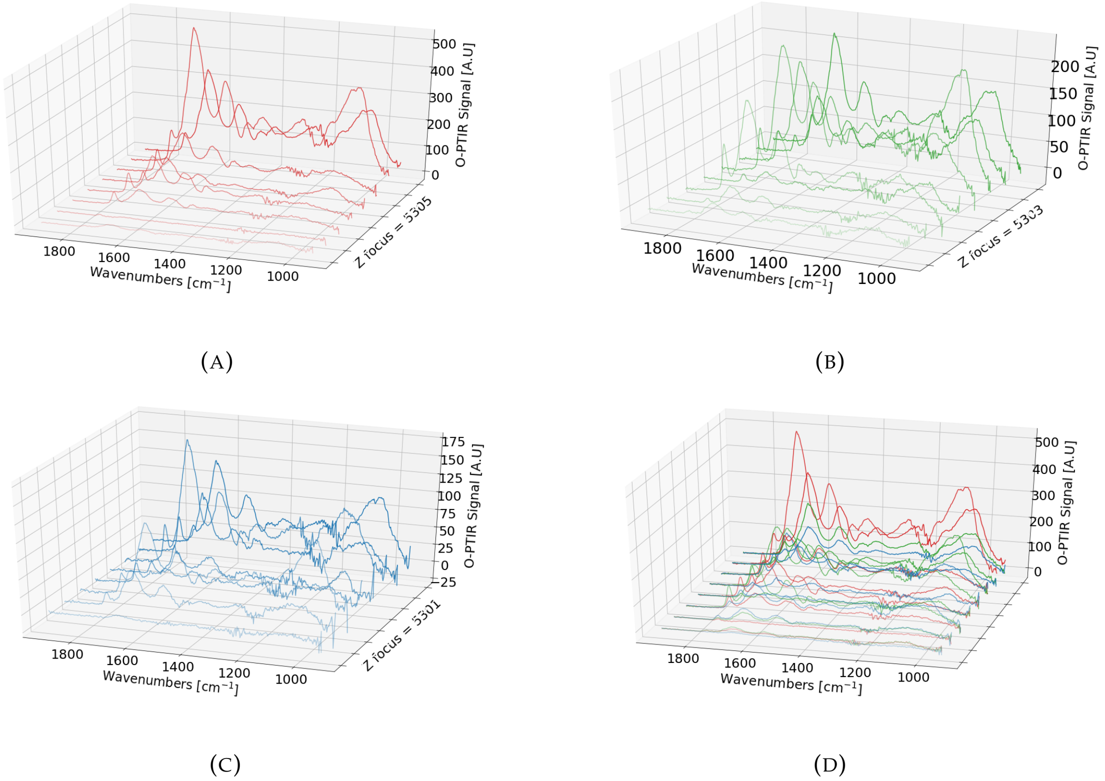
(B)
Integrated across the lipid band

FIGURE 4.16: Heat maps of depth-resolved spectral images of an intact *Phaffia rhodozyma* cell, where the Z-foci are 5305 µm, 5303 µm and 5301 µm. The spectra are baseline corrected. **(A)** Integrated across 1903-929 cm^{-1} showing the total intensity, where maximum intensity of the color bar is 80000 and minimum is 0. **(B)** integrated across the 1745 cm^{-1} lipid peak, where maximum intensity is 3000 and minimum is 0. Scale bar of the optical image: 13.85 µm

Figure 4.17A-C shows a line scan across the cell shown in Fig. 4.16, where Fig. 4.17A-C show the spectra with a constant focus on 5305 µm, 5303 µm and 5301 µm respectively. Figure 4.17D shows all three line scans in one figure using the same colors to mark the different Z-focuses as in Fig. 4.17A-C. For all Z-focus, the first three measurements show very weak signals, compared to the last spectra. It is not possible to obtain any information on the different depths in the cell.

4.4 3D infrared diffraction tomography

By using the pre-trained algorithm of Magnussen et al. [24] reconstructed spectra of the cell wall and cell interior were obtained for intact cells of *Mucor circinelloides*, *Phaffia rhodozyma* and *Rhodotorula graminis*. Figure 4.18A shows the raw spectra of an FTIR microspectroscopic



(A)

(B)

(C)

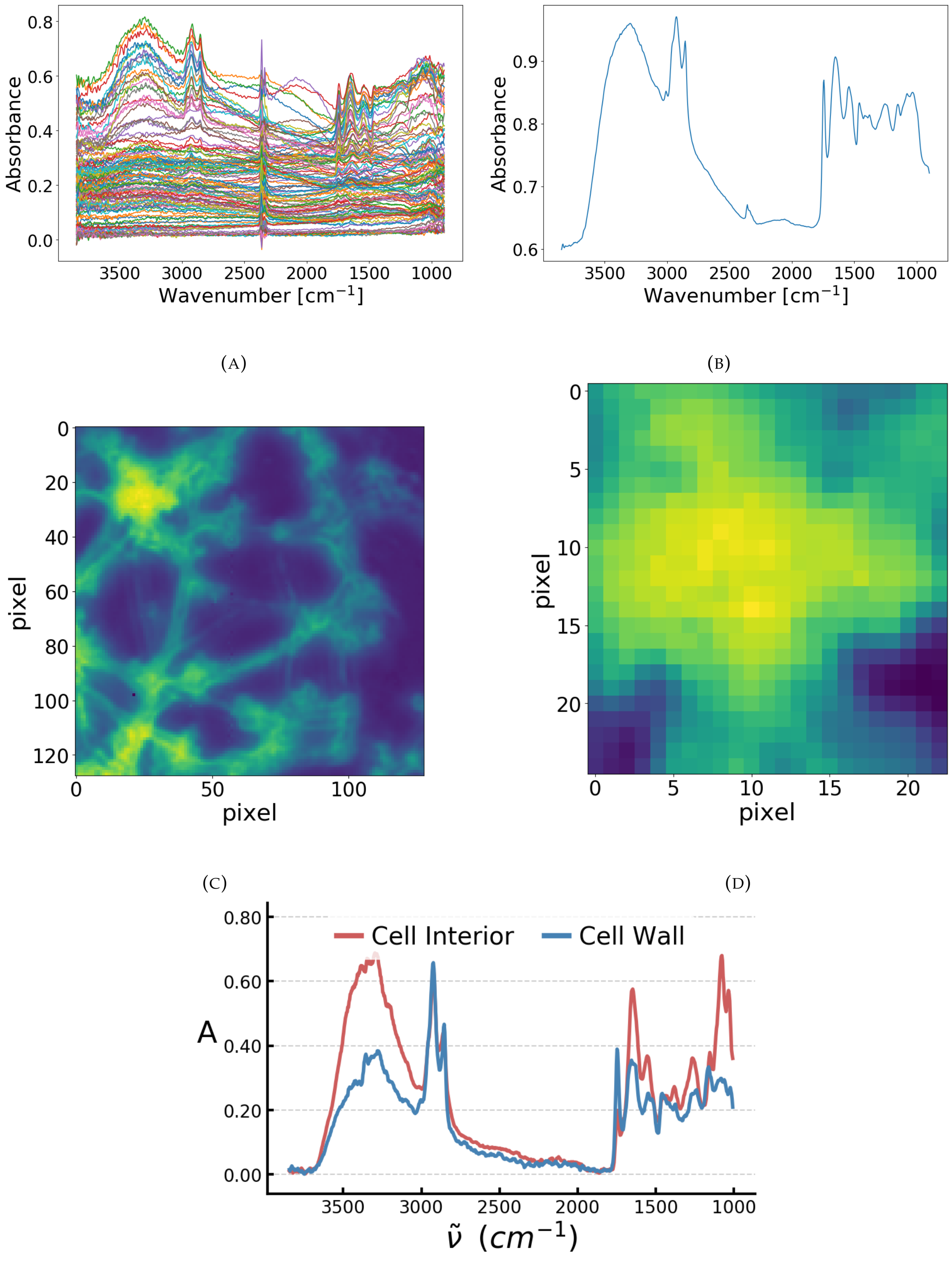
(D)

FIGURE 4.17: Depth-resolved O-PTIR measurements of an intact yeast *Phaffia rhodozyma* cell, where the spectra are baseline corrected. **(A)** Line scan with constant Z-focus equal to 5305 µm. **(B)** Line scan with Z-focus equal to 5303 µm. **(C)** Line scan with constant Z-focus equal to 5301 µm. **(D)** All three line scans are visualized in one figure with the same colors.

image of *Mucor circinelloides*. Figure 4.18C shows a corresponding chemical image, where the total absorbance is displayed as a heat map. Figure 4.18D shows the area of interest of the cell defining the area of which spectra where used for the 3D diffraction tomographic reconstruction of the cell wall and cell interior according to [24]. In Fig. 4.18B the average spectrum of the cell area of Fig. 4.18D is shown. Lastly, the reconstructed spectra of cell wall and cell interior, obtained with the reconstruction algorithm of Magnussen et. al.[24], are shown in Fig. 4.18E. The predicted inner and outer radius of the cell is 7.5 µm and 9.4 µm, respectively. The predicted refractive index of the cell core and cell wall is 1.33 and 1.35, respectively. We expect the refractive index of biological cells to be approximately equal to the refractive index of water which is 1.33. Thus, the predicted values of the refractive index is in a valid range. The red spectrum in Fig. 4.18E refers to the cell interior and does not show a strong signal at the lipid band at 1745 cm^{-1}, which is expected to be prominent in the cell interior. The blue spectrum of the cell wall shows a significant signal of the lipid peaks at 1745 cm^{-1}, 2855 cm^{-1} and 2925 cm^{-1}, which are not expected to be that prominent in the cell wall. In the paper [24], the authors were able to predict the chemical absorbance of the cell wall and the cell interior, which corresponded very well with the expected composition known from the literature [10]. In both this thesis and in the Magnussen study, FPA hyperspectral images are considered, while different fungal strains were investigated in [24] and this thesis. Figure 4.18C shows the results for different fungal cells, while no significant improvement were achieved when testing other cells. A reason why we are not able to reconstruct a meaningful spectrum of the cell wall and the cell interior might be because the mean spectrum obtained for the area of the cell according to Fig. 4.18D does not show significant scattering signals. The algorithm extracts information from the scattering features to solve the inverse scattering problem. Therefore, it is important that the average spectrum contains some scatter information in order to get a good prediction. In the spectral image in Fig. 4.18A, the averaging process eliminates the scattering. In future work, it needs to be evaluated how the scatter information in the spectra can be used in a better way for the 3D reconstruction.

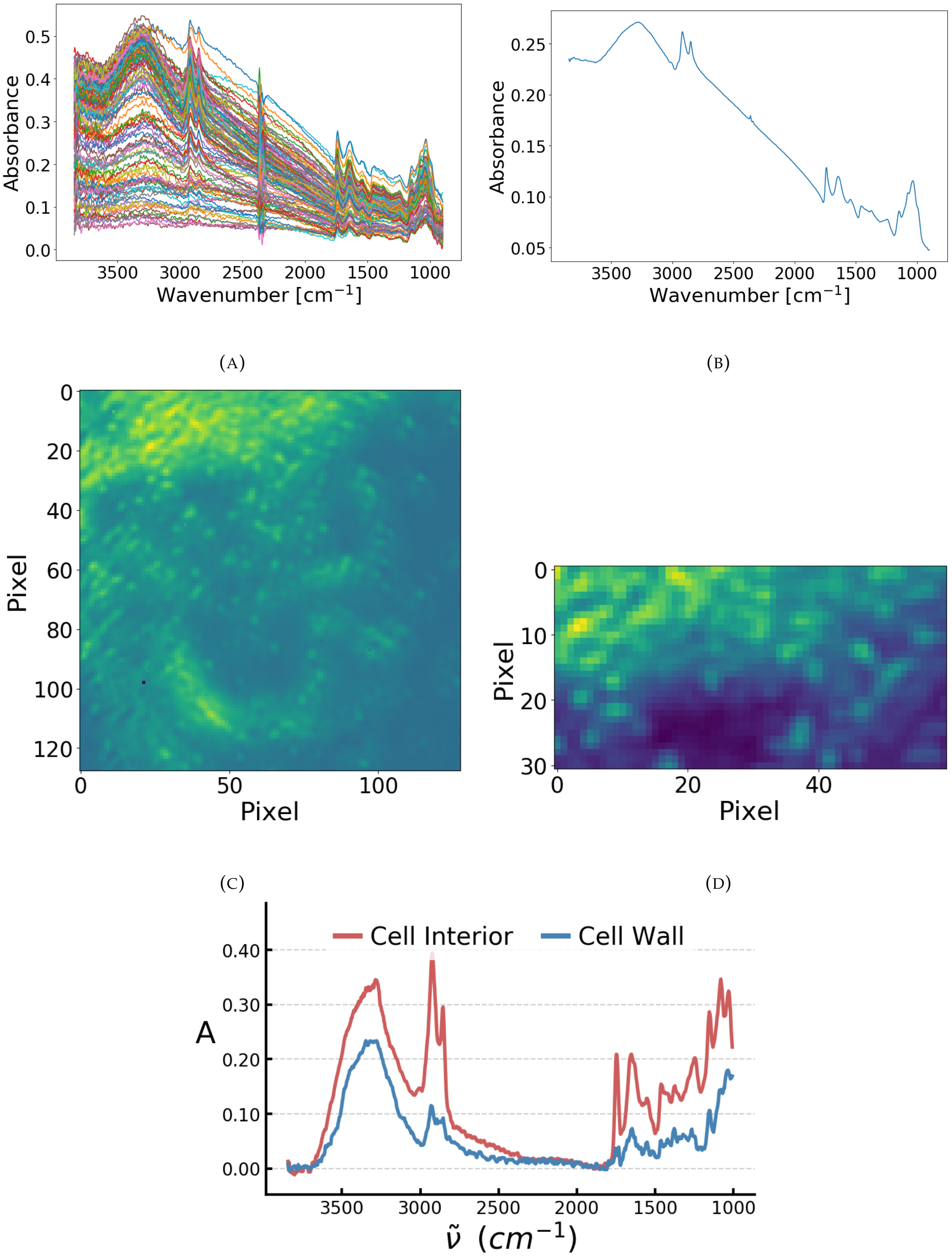
Figure 4.19A shows an FTIR microspectroscopic image of *Phaffia rhodozyma*, where significant scattering features are present. In Fig. 4.19B the respective hyperspectral image is visualized, and Fig 4.19C shows the hyperspectral image of the average spectrum displayed in 4.19D. The average spectrum has distinct physical effects from scattering. Fig4.19E shows the predicted cell interior and cell wall where the predicted inner and outer radius were 2.3 µm and 3.3 µm, respectively. The predicted size corresponds well with the actual size which is between 6 µm and 10 µm in diameter. The refractive index in the core and in the cell wall are predicted to be 1.33 and 1.36 respectively, which lie in a valid range for the refractive index. Clearly, we see molecular absorption assigned to lipids for the cell interior, with pronounced peaks at 2925 cm^{-1}, 2855 cm^{-1} and 1745 cm^{-1}. The results are promising showing information rich spectra of both the cell interior and the cell wall. The average spectrum does not originate from one cell but from many cells, as seen in Fig. 4.19C. Some cells are even overlapping. Since the deep-learning empowered 3D diffraction tomographic inverse model is based on spherical isolated samples, the optimal would be cells which are well separated and not overlapping. However when picking single cells the spectra had low quality and the prediction of cell wall and interior spectra did not acieve good results. Due to the limited spectral resolution of the FTIR microscopy, the resolution is limited. Satisfactory spectra where only obtained when the cells were clustered together. In this case, the predicted cell and cell interior spectra may however be not valid as the deep-learning empowered 3D diffraction tomographic inverse model is based on single cells.

Figure 4.20A shows spectra obtained from *Rhodotorula graminis*. The size of the sample is only a few micrometers, which is very close to the limit of the spatial resolution of the FTIR microscopic system. Thus, binning is used to be able to get chemical signals of the cell and not only noise. Therefore it is not possible to see the cells on the hyperspectral image in Fig. 4.20C. The average spectrum of the whole hyperspectral image is visualized in Fig. 4.20B. Finally, the predicted spectra of the cell wall and cell interior are visualized in Fig. 4.20D. Further, the predicted radius of the cell core is 8.0 µm and the radius of the cell wall is predicted to be 12.0 µm. The algorithm is thus not able to predict the correct size of the sample. The spectral resolution of the optical system is not good enough to get any signal of single cells which makes it impossible for the algorithm to predict the correct values. The predicted refractive index is 1.32 and 1.42 for the core and the cell wall, respectively, which is the result we expect. The spectrum of the cell wall contains a lot of noise and it is not possible to rely on the chemical information in the cell interior. Further, it would be interesting to measure the same cell with synchrotron radiation or with a QCL laser system instead of a conventional IR source. The advantage of using synchrotron IR radiation is the high-brightness feature. In comparison, the brightness of a synchrotron source is more than two orders larger than the brightness of a conventional IR source [43].



(A) (B) (C) (D) (E)

FIGURE 4.18: **(A)** Raw spectra of an FTIR microspectroscopic image of *Mucor circinelloides*. **(B)** Average spectrum of the region visualized in **D**. **(C)** Hyperspectral image with the total absorbance intensity visualized as a heat map. **(D)** Selected area of interest. **(E)** 3D reconstructed spectra of cell interior and cell wall, obtained by the pre-trained 3D diffraction tomographic reconstruction algorithm developed by Magnussen et at. [24].



(A)

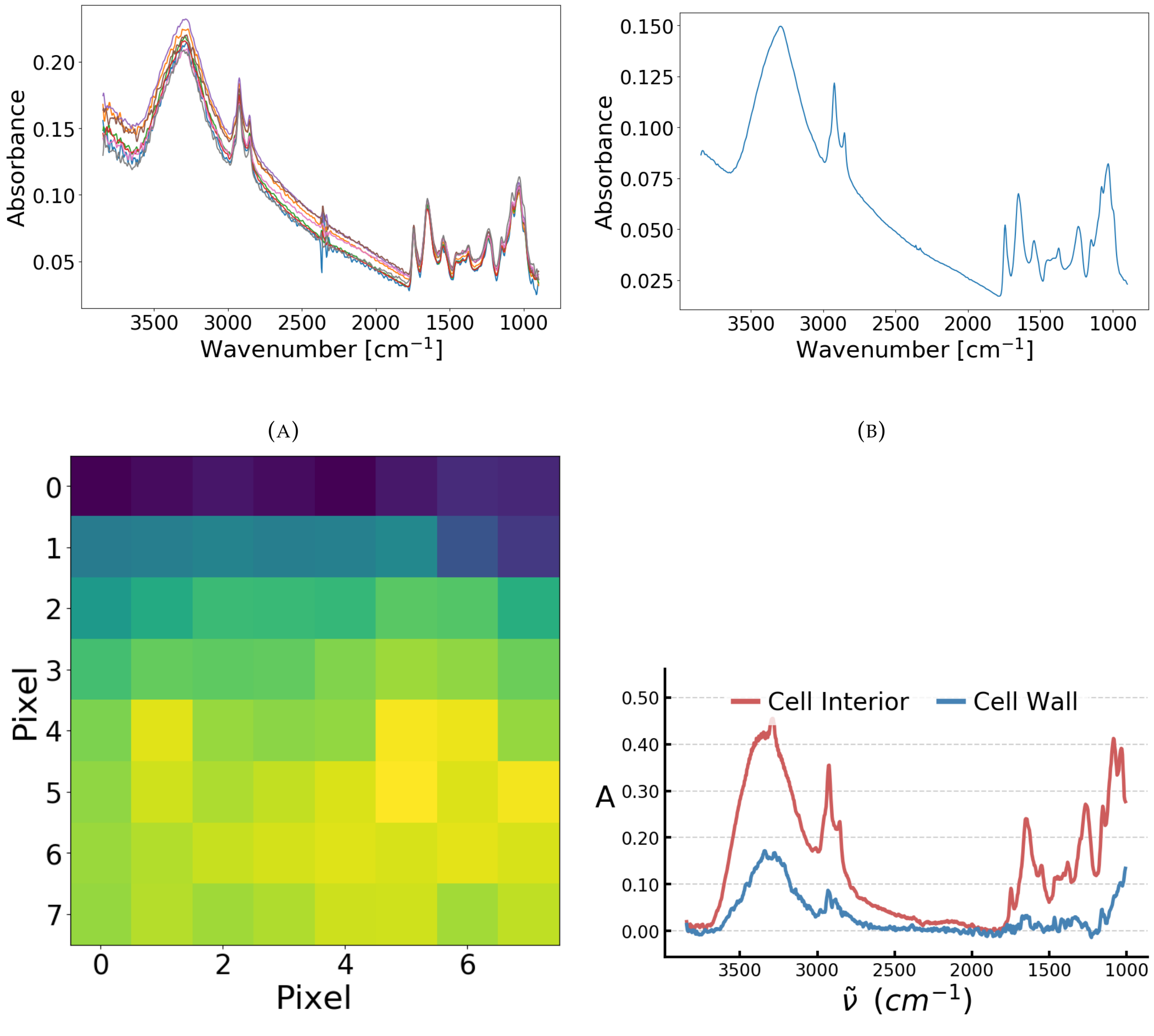
(B)

(C)

(D)

(E)

FIGURE 4.19: **(A)** Raw spectra of an FTIR microspectroscopic image of *Phaffia rhodozyma*, **(B)** Average spectrum of the region visualized in **D**. **(C)** Hyperspectral image with the total absorbance intensity visualized as a heat map. **(D)** Selected area of interest. **(E)** 3D reconstructed spectra of cell interior and cell wall, obtained by the pre-trained 3D diffraction tomographic reconstruction algorithm developed by Magnussen et at. [24].



(A)

(B)

(C)

(D)

FIGURE 4.20: **(A)** Raw spectra of an FTIR microspectroscopic image of *Rhodotorula graminis*. **(B)** Average spectrum of the region visualized in **C**. **(C)** Hyperspectral image with the total absorbance intensity visualized as a heat map. **(E)** 3D reconstructed spectra of cell interior and cell wall, obtained by the pre-trained 3D diffraction tomographic reconstruction algorithm developed by Magnussen et at. [24].

Chapter 5

Conclusions and Outlook

In this thesis, biological samples have been analyzed with three different micro- and nano-IR spectroscopic instrumentation to evaluate how well subcellular chemical information is obtained. Two different preparation methods have been considered, namely sectioned cells and intact cells.

Sectioned cells of *Rhodotorula graminis* and *Aurautiochyrium limaanium* were measured with the nano-IR spectroscopic instrumentation O-PTIR and AFM-IR. The optical image of O-PTIR does not have a high enough resolution to be able to get a good visualization of sectioned cells of *Rhodotorula graminis*. Thus, it is not possible to know exactly which cell compartment is probed. The absorbance spectra show clear spectral signatures, however, the spectra are highly contaminated with epoxy. Subtraction of epoxy was tried by correcting the spectra by weighted EMSC with epoxy as an interferent spectrum. Further, PCA and MCR were performed without succeeding to remove the epoxy and obtain subcellular chemical information. A problem with epoxy is that the molecules absorb in the same spectral range as the molecules of biological cells. In further studies, other mediums should be evaluated which do not share the same absorption bands as biological cells. The same sectioned cells were also measured with AFM-IR. The optical visualization has a much higher resolution than O-PTIR making it possible to obtain highly detailed information of cells on subcellular level. Further, the AFM-IR intensity spectra face the same issue as the absorbance spectra of O-PTIR, due to the epoxy. However, by integrating across significant bands such as the 1745 cm^{-1}, 1655 cm^{-1}, 1555 cm^{-1}, 1460 cm^{-1} and 1415 cm^{-1} and presenting them as heat maps, we successfully managed to distinguish the cell wall and the cell interior. The absorbances at the wavenumbers 1555 cm^{-1}, 1460 cm^{-1} and 1415 cm^{-1} show significantly higher signal in the cell interior compared to the cell wall, and the absorbance band at 1655 cm^{-1} shows higher intensity in the cell wall. According to the literature [10] this corresponds well with the chemical absorption peaks that are expected in the cell wall and the cell interior.

Sectioned cells of *Aurautiochyrium limaanium* were also measured with O-PTIR and AFM-IR, however, it was not possible to acquire consistent spectra, without noise or other features disturbing the sample. Thus, we were not able to obtain high quality spectral information of *Aurautiochyrium limaanium*.

Depth-resolved spectra were acquired with O-PTIR with different Z-focus. Intact cells of *Phaffia rhodozyma* and *Mucor circinelloides* were measured with this technique. The aim was to obtain chemical information on the different layers in the cells. The spectra show that the total intensity increases as the focus reaches the center of the cell, and then decreases as the Z-focus progresses further in depth of the cell. In addition to the total intensity, we see a similar shape of the area across the 1745cm $^{-1}$ lipid band. With respect to the total intensity and the lipid signatures, it seems like it is possible to obtain significant signals from the different layers. However, the recorded absorbance spectra do not only contain signals from the focus point but also signals from the surrounding area. Thus, the chemical variation in the spectra is not only coming from the respective layer of the cells. The idea of in-depth measurements with O-PTIR comes from the study by Zhang et al. [44]. The authors perform transmission measurements and not reflection measurements which has been done in this thesis. However, the same issues of sampling signal from other areas than the z-focus are valid. O-PTIR is a new technique and the physical background of the measurement principle when probing real samples is not fully understood. Measurement depth and resolution need to be investigated further in order to understand from which cell compartments chemical information is collected. To summarize, at the current stage, it is not possible to conclude that the difference in the absorbance spectra is due to chemical variation in the different layers of the biological cell or if it is due to an increase in the responsivity of the sample. The absorbance spectra show contributions of chemical absorption from the whole laser path of the probing laser and not only from the focus point.

Because of the diffraction limitation, due to the magnitude of the mid-IR radiation, it is not possible for FTIR-spectroscopy to obtain subcellular spatially resolved chemical information of biological cells on a few micrometers. However, by taking advantage of highly scattered distorted spectra, named Mie-scattering, which occurs due to the spherical morphology of biological cells, Magnussen et al. [24] have shown that it is possible to obtain 3D chemical information of the cell wall and cell interior. Earlier, scattering features have been seen as a disturbing effect that obscures the absorbance spectra and needs to be removed by preprocessing techniques such as ME-EMSC. By taking advantage of the relationship of scattering, refractive index and the absorptivity of the sample, the inverse scattering problem can be solved by using a 3D infrared diffraction tomographic approach which has successfully obtained subcellular chemical information of the cell wall and the cell interior. In this thesis, a pre-trained deep learning algorithm from Magnussen et al. [24] has been used on three different biological samples, namely *Rhodotorula graminis*, *Phaffia rhodozyma* and *Mucor circinelloides* measured with FTIR microspectroscopic imaging. *Rhodotorula graminis* is the smallest cell with a diameter range of 2.3-3.9 µm. This is in the limit of the spectral resolution of the instrument which is between 2-10 cm^{-1}. Thus measuring intact single cells caused considerable noise in the spectra and poor quality spectra. Therefore, binning was used to obtain chemical information about the cells. However, the 3D infrared diffraction tomography approach was not able to correctly predict the size of the cells and the chemical absorbance in the cell wall and cell interior. The main reason is that the spatial resolution of the instrument is too low and the binning makes it impossible to average across cells because the hyperspectral image is lost. A suggestion for further work is to measure the same sample with synchrotron IR radiation which has higher brightness of the source and which might lead to higher spatial resolution. Further, the algorithm successfully managed to reconstruct the chemical information of the cell wall and cell interior of intact *Phaffia rhodozyma* cells. In addition, the algorithm correctly predicted the size of the cell. The cells have a range between 6-10 µm in diameter, which is big enough to get chemical information about the cells without binning. Lastly, intact cells of *Mucor circinelloides* were measured with FTIR microspectroscopic imaging. The average spectrum did not contain significant scattering features, thus the algorithm did not manage to make a good prediction of the cell wall and cell interior. In addition, a poor prediction of the size of the cell was obtained. In further work, an evaluation of how we can extract the scattering features in a better way is necessary to get better predictions.

We have seen that it is difficult to obtain subcellular spatially resolved chemical information of biological cells. However, with AFM-IR of sectioned cells and the 3D infrared diffraction tomography approach of intact cells measured with FPA hyperspectral image we have been able to obtain chemical information and physical information about the cell wall and the cell interior. In table 5.1 a summary of the main results of the O-PTIR, AFM-IR and FTIR with the 3D IR diffraction tomography approach of both sectioned cells and intact cells are shown. The text colored in green shows were the methods was successful and the text colored in blue corresponds to situations where problems with the respective approach were observed. The grey empty columns are not considered and are outside the scope of this thesis.

As for O-PTIR we have not with certainty been able to get chemical information on subcellular level. However, this instrumentation is very new and there is a big lack of understanding the optical system and the physics of the configuration. Further research should therefore focus on the understanding of the system and then a reevaluation of how it is possible to use sectioned cells and in-depth measurements should be made.

We can see that the deep-learning empowered 3D diffraction tomographic infrared spectroscopy works well for larger cells. It has the advantage that no sectioning and embedding is needed. But the techniques is very limited in resolution.

AFM-IR has by far the best spatial resolution and it delivers premium spectra of subcellular components at nano-scale. However, further work with the development of sectioning protocols is needed to achieve spectra that are less contaminated by embedding materials.

TABLE 5.1: Summary of the main results of O-PTIR, AMF-IR, and FTIR-FPA using the 3D diffraction tomography approach. Overview of both sectioned and intact cells, in addition to the resolution. The green color indicates successful performance of the respective method, while the blue color indicates issues with the approach. The empty columns are not considered in this thesis.

	AFM-IR	O-PTIR	FTIR - 3D IR Diffraction Tomography
Sectioned cells	Possible to distinguish cell wall from cell interior. Contamination of epoxy especially between 1200-1000 cm^{-1}.	May work with bigger sample (>10µm). Dominated by epoxy signals which is impossible to remove.	
Intact cells		Good quality signal from both surface and within the cell. Uncertain on where the signal comes from when probing in-depth.	Works on samples > 5µm. High quality pure absorbance spectra of cell wall and cell interior. Not working on samples < 5µm. Need to be highly scattered spectra.
Resolution	> 50 nm	>532 nm	

Bibliography

[1] Nils Kristian Afseth and Achim Kohler. "Extended multiplicative signal correction in vibrational spectroscopy, a tutorial". In: *Chemometrics and Intelligent Laboratory Systems* 117 (2012), pp. 92–99.

[2] Yeran Bai, Jiaze Yin, and Ji-Xin Cheng. "Bond-selective imaging by optically sensing the mid-infrared photothermal effect". In: *Science Advances* 7.20 (2021), eabg1559.

[3] Paul Bassan et al. "Resonant Mie scattering in infrared spectroscopy of biological materials–understanding the 'dispersion artefact'". In: *Analyst* 134.8 (2009), pp. 1586–1593.

[4] Paul Bassan et al. "RMieS-EMSC correction for infrared spectra of biological cells: Extension using full Mie theory and GPU computing". In: *Journal of biophotonics* 3.8-9 (2010), pp. 609–620.

[5] Rohit Bhargava. "Infrared spectroscopic imaging: the next generation". In: *Applied spectroscopy* 66.10 (2012), pp. 1091–1120.

[6] Amy L Bondy et al. "Atomic force microscopy-infrared spectroscopy of individual atmospheric aerosol particles: subdiffraction limit vibrational spectroscopy and morphological analysis". In: *Analytical chemistry* 89.17 (2017), pp. 8594–8598.

[7] James W Cooley and John W Tukey. "An algorithm for the machine calculation of complex Fourier series". In: *Mathematics of computation* 19.90 (1965), pp. 297–301.

[8] Alexandre Dazzi, F Glotin, and R Carminati. "Theory of infrared nanospectroscopy by photothermal induced resonance". In: *Journal of Applied Physics* 107.12 (2010), p. 124519.

[9] Alexandre Dazzi and Craig B Prater. "AFM-IR: Technology and applications in nanoscale infrared spectroscopy and chemical imaging". In: *Chemical reviews* 117.7 (2017), pp. 5146–5173.

[10] Simona Dzurendova. "Sustainable fungal biorefineries: optimizing production of valuable metabolites in oleaginous Mucoromycota". In: (2021).

[11] Simona Dzurendova et al. "Metal and phosphate ions show remarkable influence on the biomass production and lipid accumulation in oleaginous Mucor circinelloides". In: *Journal of Fungi* 6.4 (2020), p. 260.

[12] Simona Dzurendova et al. "Microcultivation and FTIR spectroscopy-based screening revealed a nutrient-induced co-production of high-value metabolites in oleaginous Mucoromycota fungi". In: *PloS one* 15.6 (2020), e0234870.

[13] D Fournier, F Lepoutre, and AC Boccara. "Tomographic approach for photothermal imaging using the mirage effect". In: *Le Journal de Physique Colloques* 44.C6 (1983), pp. C6–479.

[14] Peter R Griffiths. "Fourier transform infrared spectrometry". In: *Science* 222.4621 (1983), pp. 297–302.

[15] Masaaki Harada et al. "Photothermal microscopy with excitation and probe beams coaxial under the microscope and its application to microparticle analysis". In: *Analytical Chemistry* 65.20 (1993), pp. 2938–2940.

[16] Oxana Klementieva et al. “Super-resolution infrared imaging of polymorphic amyloid aggregates directly in neurons”. In: *Advanced Science* 7.6 (2020), p. 1903004.

[17] A Kohler et al. “Extended multiplicative signal correction as a tool for separation and characterization of physical and chemical information in Fourier transform infrared microscopy images of cryo-sections of beef loin”. In: *Applied spectroscopy* 59.6 (2005), pp. 707–716.

[18] Achim Kohler et al. “Interpreting several types of measurements in bioscience”. In: *Biomedical vibrational spectroscopy* (2008), pp. 333–356.

[19] Tatiana Konevskikh et al. “Fringes in FTIR spectroscopy revisited: understanding and modelling fringes in infrared spectroscopy of thin films”. In: *Analyst* 140.12 (2015), pp. 3969–3980.

[20] Tatiana Konevskikh et al. “Mie scatter corrections in single cell infrared microspectroscopy”. In: *Faraday discussions* 187 (2016), pp. 235–257.

[21] Miriam Kosik et al. “Interaction of atomic systems with quantum vacuum beyond electric dipole approximation”. In: *Scientific reports* 10.1 (2020), p. 5879.

[22] Peter Lasch and Dieter Naumann. “Spatial resolution in infrared microspectroscopic imaging of tissues”. In: *Biochimica et Biophysica Acta (BBA)-Biomembranes* 1758.7 (2006), pp. 814–829.

[23] Eirik Almklov Magnussen et al. “Deep convolutional neural network recovers pure absorbance spectra from highly scatter-distorted spectra of cells”. In: *Journal of biophotonics* 13.12 (2020), e202000204.

[24] Eirik Almklov Magnussen et al. “Deep learning-enabled Inference of 3D molecular absorption distribution of biological cells from IR spectra”. In: *Communications Chemistry* 5.1 (2022), p. 175.

[25] Harald Martens, SA Jensen, and P Geladi. “Multivariate linearity transformation for near-infrared reflectance spectrometry”. In: *Proceedings of the Nordic symposium on applied statistics*. Stokkand Forlag Publishers Stavanger, Norway. 1983, pp. 205–234.

[26] Harald Martens, Jesper Pram Nielsen, and Søren Balling Engelsen. “Light scattering and light absorbance separated by extended multiplicative signal correction. Application to near-infrared transmission analysis of powder mixtures”. In: *Analytical Chemistry* 75.3 (2003), pp. 394–404.

[27] Harald Martens and Edward Stark. “Extended multiplicative signal correction and spectral interference subtraction: new preprocessing methods for near infrared spectroscopy”. In: *Journal of pharmaceutical and biomedical analysis* 9.8 (1991), pp. 625–635.

[28] Jérémie Mathurin, Ariane Deniset-Besseau, and Alexandre Dazzi. “Advanced Infrared Nanospectroscopy Using Photothermal Induced Resonance Technique, AFMIR: New Approach Using Tapping Mode”. In: *Acta Physica Polonica A* 137.1 (2020), pp. 29–32.

[29] Albert A Michelson. “XXVIII. Visibility of interference-fringes in the focus of a telescope”. In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 31.190 (1891), pp. 256–259.

[30] Brian Mohlenhoff et al. “Mie-type scattering and non-Beer-Lambert absorption behavior of human cells in infrared microspectroscopy”. In: *Biophysical journal* 88.5 (2005), pp. 3635–3640.

[31] *O-PTIR: non-contact, submicron, visible probe photothermal infrared spectroscopy*. 2023. URL: `https://www.photothermal.com/home/o-ptir/`.

[32] Victor W Or et al. “Lab on a tip: atomic force microscopy–photothermal infrared spectroscopy of atmospherically relevant organic/inorganic aerosol particles in the nanometer to micrometer size range”. In: *Analyst* 143.12 (2018), pp. 2765–2774.

[33] John Reffner. “Advances in infrared microspectroscopy and mapping molecular chemical composition at submicrometer spatial resolution”. In: *Spectroscopy* 33.9 (2018), pp. 12–17.

[34] Cyril Ruckebusch. *Resolving spectral mixtures: with applications from ultrafast time-resolved spectroscopy to super-resolution imaging*. Elsevier, 2016.

[35] Reiner Salzer. “Peter R. Griffiths, James A. de Haseth: Fourier transform infrared spectrometry (2nd edn.)” In: *Analytical and Bioanalytical Chemistry* 391.7 (2008), pp. 2379–2380.

[36] Jonathon Shlens. “A tutorial on principal component analysis”. In: *arXiv preprint arXiv: 1404.1100* (2014).

[37] Brian C Smith. *Fundamentals of Fourier transform infrared spectroscopy*. CRC press, 2011.

[38] Johanne H Solheim et al. “An open-source code for Mie extinction extended multiplicative signal correction for infrared microscopy spectra of cells and tissues”. In: *Journal of biophotonics* 12.8 (2019), e201800415.

[39] Johanne Heitmann Solheim. “Modelling scattering and absorption in the vibrational spectroscopy of cells and tissues”. In: (2023).

[40] Johanne Heitmann Solheim et al. “The Use of Constituent Spectra and Weighting in Extended Multiplicative Signal Correction in Infrared Spectroscopy”. In: *Molecules* 27.6 (2022), p. 1900.

[41] Roma Tauler. “Multivariate curve resolution applied to second order data”. In: *Chemometrics and intelligent laboratory systems* 30.1 (1995), pp. 133–146.

[42] David L Wetzel and Steven M LeVine. “Imaging molecular chemistry with infrared microscopy”. In: *Science* 285.5431 (1999), pp. 1224–1225.

[43] P Yu. “Application of advanced synchrotron radiation-based Fourier transform infrared (SR-FTIR) microspectroscopy to animal nutrition and feed science: a novel approach”. In: *British journal of nutrition* 92.6 (2004), pp. 869–885.

[44] Delong Zhang et al. “Depth-resolved mid-infrared photothermal imaging of living cells and organisms with submicrometer spatial resolution”. In: *Science advances* 2.9 (2016), e1600521.

Norges miljø- og biovitenskapelige universitet
Noregs miljø- og biovitskapelege universitet
Norwegian University of Life Sciences

Postboks 5003
NO-1432 Ås
Norway