FORESTRY AND NATURAL SCIENCES

Lassi Rieppo

Infrared Spectroscopic Characterization of Articular Cartilage

Publications of the University of Eastern Finland
Dissertations in Forestry and Natural Sciences

UNIVERSITY OF EASTERN FINLAND

LASSI RIEPPO

Infrared Spectroscopic Characterization of Articular Cartilage

Publications of the University of Eastern Finland
Dissertations in Forestry and Natural Sciences
No 90

Academic Dissertation
To be presented by permission of the Faculty of Science and Forestry for public examination in the Auditorium ML3 in Medistudia Building at the University of Eastern Finland, Kuopio, on December, 12, 2012,
at 12 o'clock noon.

Department of Applied Physics

Kopijyvä Oy
Kuopio, 2012
Editors: Prof. Pertti Pasanen, Prof. Kai-Erik Peiponen, Prof. Matti Vornanen, Prof. Pekka Kilpeläinen

Distribution:
University of Eastern Finland Library / Sales of publications
P.O. Box 107, FI-80101 Joensuu, Finland
tel. +358-50-3058396
http://www.uef.fi/kirjasto

ISBN: 978-952-61-0963-3 (printed)
ISSNL: 1798-5668
ISSN: 1798-5668
ISBN: 978-952-61-0964-0 (PDF)
ISSN: 1798-5676 (PDF)

Author's address:
Department of Applied Physics
University of Eastern Finland
P.O.Box 1627
FI-70211 KUOPIO
FINLAND
email: lassi.rieppo@uef.fi

Supervisors:
Professor Jukka Jurvelin, Ph.D.
Department of Applied Physics
University of Eastern Finland
email: jukka.jurvelin@uef.fi

Adjunct Professor Simo Saarakkala, Ph.D.
Department of Medical Technology
Institute of Biomedicine
University of Oulu
email: simo.saarakkala@oulu.fi

Adjunct Professor Jarno Rieppo, M.D., Ph.D.
Institute of Biomedicine
University of Eastern Finland
email: jarno.rieppo@gmail.com

Reviewers:
Professor Nancy Pleshko, Ph.D.
Department of Mechanical Engineering
Temple University
Philadelphia, PA, USA
email: npleshko@temple.edu

Associate Professor Achim Kohler, Ph.D.
Department of Mathematical Sciences and Technology
Norwegian University of Life Sciences
Aas, Norway
email: achim.kohler@umb.no

Opponent:
Senior Investigator Richard Spencer, M.D., Ph.D.
Magnetic Resonance Imaging and Spectroscopy Section
National Institute on Aging
National Institutes of Health
Baltimore, MD, USA
email: spencer@helix.nih.gov

ABSTRACT

Articular cartilage (AC) is avascular and aneural tissue that covers the ends of long bones. The function of AC is to reduce the stresses exposed on the subchondral bone and to minimize the friction between the articulating bones during locomotion. Osteoarthritis (OA) is globally the most common joint disease. In OA, cartilage degeneration causes pain and leads to decreased joint mobility. AC is mainly composed of collagen, proteoglycans (PGs) and interstitial water. The early OA changes in the composition of AC occur before there are any clinical signs of the disease. Characterization of these changes is essential if one wishes to understand the disease. Fourier Transform Infrared (FTIR) spectroscopic imaging is a powerful tool for chemical analysis at the microscopic level. In AC, high specificity in the FTIR spectroscopic parameters for collagen and PGs is required. This thesis work compared the FTIR spectroscopic analysis methods for compositional analysis of AC. Another aim was to analyze the interrelationships between the spectroscopic data and biomechanical properties of AC. The results show that important information about the biochemical composition of AC can be extracted from the FTIR spectra. The biochemical specificity can be optimized with the use of multivariate regression methods and the results further improved with variable selection algorithms. The biomechanical properties can also be predicted from FTIR spectra with similar or better specificity than with the previous biochemical methods. Due to the complex structure of AC, the average composition cannot fully explain its biomechanical properties. Subsequently, the model may be further improved by inclusion of a layered tissue structure with a variable composition and collagen network orientation in the depth-wise direction.

Universal Decimal Classification: 543.42
National Library of Medicine Classification: QT 36, QU 55.3, WE 300, WE 348, WN 180
Medical Subject Headings: Cartilage, Articular; Microscopy; Spectro-

scopy, Fourier Transform Infrared; Collagen; Biomechanics; Multivariate Analysis; Regression Analysis; Osteoarthritis/diagnosis
Yleinen suomalainen asiasanasto: nivelrusto; mikroskopia; spektroskopia; kollageenit; biomekaniikka; monimuuttujamenetelmät; regressioanalyysi; nivelrikko - - diagnoosi

To Kia

Acknowledgements

This study was carried out during 2008-2012 in the Department of Applied Physics, University of Eastern Finland.

I am grateful to my principal supervisor, Professor Jukka Jurvelin, Ph.D, for giving me the opportunity to work in his research group and for sharing his expertise in cartilage research. I want to offer my sincere thanks to my second supervisor, Adjunct Professor Simo Saarakkala, Ph.D., for all the help with various issues which arose during this thesis project. I want to express my gratitude to my third supervisor, my brother, Adjunct Professor Jarno Rieppo, M.D., Ph.D., for introducing me into the world of scientific work and for the continuous supervision and support. I have learned so much through our discussions about research.

I want to thank Professor Heikki J. Helminen, M.D., Ph.D., for his fatherly support and encouragement at the beginning of my research career. I thank Tommi Närhi, M.D., for the help with sample preparation and, more importantly, for being a friend outside work. I also thank my other co-authors, Mikko Lammi, Ph.D., and Jaakko Holopainen, M.D., for their contributions to the manuscripts and for providing the purified cartilage constituents.

I am grateful to the official reviewers of this thesis, Professor Nancy Pleshko, Ph.D., and Associate Professor Achim Kohler, Ph.D., for their constructive criticism and suggestions to improve the thesis. I thank Ewen MacDonald, D.Pharm, for the linguistic review.

It has been a priviledge to work in an inspirational and successful research group such as ours in Biophysics of Bone and Cartilage. My warmest thanks go to all current and former members of our research group with whom I have worked. Particularly, I want to acknowledge my fellow spectroscopists Mikael Turunen, M.Sc, and Yevgeniya Kobrina, M.Sc. I also thank Mikko Nissi, Ph.D., for helping me with computer-related problems, and Petro Julkunen, Ph.D., for providing me help with MATLAB programming.

I wish to thank the personnel of the Institute of Biomedicine. Especially, Mrs. Eija Rahunen and Mr. Kari Kotikumpu are acknowledged for their help with sample processing. The personnel of Sib-Labs (formerly known as BioMater Centre) are also acknowledged.

This thesis work was financially supported by the strategic funding of the University of Eastern Finland, Ministry of Education, Academy of Finland, Kuopio University Hospital, the National Doctoral Programme of Musculoskeletal Disorders and Biomaterials (TBDP), The North Savo Regional Fund of the Finnish Cultural Foundation and Instrumentarium Science Foundation.

I owe my deepest gratitude to my parents Arja and Aimo and to my siblings Jussi and Lotta for the support and encouragement they have given to me during my whole life. Finally, I am grateful to my girlfriend Kia, who has given invaluable support to me during the writing process of this thesis.

Kuopio, 12th December 2012
Lassi Rieppo

ABBREVIATIONS

AC	articular cartilage
cm^{-1}	unit of wavenumber
DD	digital densitometry
DMMB	dimethylmethylene blue
FT-IR	Fourier transform infrared
GAG	glycosaminoglycan
IR	infrared
OA	osteoarthritis
PG	proteoglycan
PCR	principal component regression
PLSR	partial least squares regression
RMSECV	root-mean-square error of cross-validation

SYMBOLS AND NOTATIONS

A	absorbance or amplitude
a	radius of a sphere or linear baseline shift
b	multiplicative error
c	concentration
D	euclidean distance
d	linear baseline error
diam.	diameter
e	quadratic baseline error
I	intensity of light
I_0	intensity of light entering the sample
I_s	intensity of scattered light
l	path length
n	number of samples or refractive index
p	statistical significance or electric dipole moment
$\mathbf{P}$	loading matrix on $\mathbf{X}$

q	electric charge
Q	efficiency factor of scattering
$\mathbf{Q}$	loading matrix on $\mathbf{Y}$
r	Pearson's correlation coefficient or distance of the electric charge
s	sub-peak
T	transmittance
$\mathbf{T}$	score matrix
$\mathbf{W}$	additional set of loadings in partial least squares regression
$\mathbf{X}$	input variables (spectroscopic variables)
$\mathbf{Y}$	measured variables (predicted variables)
z	spectrum
$\boldsymbol{\beta}$	regression vector
ε	molecular absorption coefficient or residual term
λ	wavelength
$\tilde{\nu}$	wavenumber
σ	width of a gaussian peak

LIST OF PUBLICATIONS

This thesis consists of the following original articles, which are referred to by Roman numerals in the text:

I Rieppo L, Saarakkala S, Närhi T, Holopainen J, Lammi M, Helminen HJ, Jurvelin JS and Rieppo J, "Quantitative Analysis of Spatial Proteoglycan Content in Articular Cartilage with Fourier Transform Infrared Imaging Spectroscopy: Critical Evaluation of Analysis Methods and Specificity of the Parameters," *Microsc. Res. Tech* **73(5)** 503–12 (2010).

II Rieppo L, Saarakkala S, Närhi T, Helminen HJ, Jurvelin JS and Rieppo J, "Application of Second Derivative Spectroscopy for Increasing Molecular Specificity of Fourier Transform Infrared Imaging Spectroscopy of Articular Cartilage," *Osteoarthritis Cartilage* **20(5)** 451–9 (2012).

III Rieppo L, Rieppo J, Jurvelin JS and Saarakkala S, "Fourier Transform Infrared Imaging Spectroscopy and Partial Least Squares Regression for Prediction of Proteoglycan Content of Articular Cartilage," *PLoS ONE* **7(2)** e32344 (2012).

IV Rieppo L, Jurvelin JS, Saarakkala S and Rieppo J, "Prediction of Compressive Stiffness of Articular Cartilage Using Fourier Transform Infrared Spectroscopy," submitted.

AUTHOR'S CONTRIBUTION

The publications selected in this dissertation are original research papers on infrared spectroscopic characterization of articular cartilage. The author has contributed to the development of spectral analysis techniques and has carried out all spectroscopic measurements and analyses. The author was the main writer in the studies I-IV.

Contents

1 Introduction

Articular cartilage is connective tissue that covers the ends of long bones. Articular cartilage provides a nearly frictionless surface between the articulating bones and reduces the stresses applied to the subchondral bone. Articular cartilage is mainly composed of the fibrillar collagen network, proteoglycans (PGs), chondrocytes and interstitial water [1, 2]. The structure and composition of articular cartilage are inhomogeneous especially in the depthwise direction of the tissue [3–6]. The inhomogeneous distribution of the constituents and complex structure of AC are needed to achieve the unique biochemical properties of the tissue [7].

Osteoarthritis (OA) is globally the most common joint disease [8]. OA causes pain and impairs the joint function, making the daily life more difficult [8, 9]. OA is also responsible for significant financial losses due to reduced working ability and medical costs [8, 10]. The OA progresses slowly, and the early biochemical changes occur before there are any clinical signs of OA [11]. The degenerative changes include disruption of collagen network, loss of PGs and increase in water content [3, 7, 12–14]. These changes affect the biomechanical properties of AC [7, 15, 16]. In order to understand the structure-function relationships of AC, sensitive biochemical characterization methods are needed.

Infrared (IR) spectroscopic imaging opens new opportunities in AC research by combining biochemical analysis with microscopy [17], thus enabling the investigation of the spatial distribution of the tissue components at microscopic level [18, 19]. The first IR spectroscopic imaging studies of AC were published over a decade ago [20,21]. Since then, many studies have utilized IR spectroscopic imaging [22–33]. IR spectroscopic imaging has been used for analysis of collagen and PG contents and the collagen integrity. Furthermore, the orientation of the collagen fibrils can be determined using polarized IR light. There have been advances in the spectral analy-

sis methods, but the progress has been slow. The aim of this thesis was to introduce new spectral analysis methods for AC research and compare them with the previously used methods. In this thesis, curve fitting, second derivative spectroscopy and multivariate regression models were utilized for the determination of the composition of AC. Multivariate regression models were also used for predicting the compressive biomechanical properties of AC directly from their IR spectra.

2 Articular Cartilage

2.1 COMPOSITION AND STRUCTURE

Articular cartilage (AC) is an aneural and avascular specialized tissue which covers the ends of long bones. The function of AC is to reduce the stresses applied to bone ends and to provide a nearly frictionless surface for the articulating bones during locomotion. AC is composed of two phases: a fluid phase, which consists of water and electrolytes, and a solid phase, which is formed mainly by collagen fibrils, PGs, glycoproteins and chondrocytes. Collagen molecules account for 15-22% of the wet weight of AC [1, 2, 34–36]. Type II collagen forms the vast majority of collagens in AC (90-95% of total collagen amount). Types I, III, VI, IX, X, XI, XII and XIV are also found in AC, but they account only for 5-10% of the total amount of collagen in AC [1, 2, 37, 38]. Collagen forms a highly organized fibrillar network which entraps other matrix components within the tissue.

PGs are the second largest component of the solid phase of AC, as they account for 4-10% of the wet weight of AC [1, 2, 34, 36]. PGs are composed of a protein core to which numerous glycosaminoglycans (GAGs) are covalently attached. Aggrecan is the most common PG in AC. Aggrecan contains core protein (7%) and GAGs chondroitin sulphate (87%) and keratan sulphate (6%) [39]. Aggrecan forms large macromolecular units, *i.e.*, aggregates in AC (Figure 2.1). Aggregates are formed when a large number of PG monomer units become attached to hyaluronan chain through link proteins [40–43]. The negatively charged carboxyl and sulphate groups in GAGs attract positive Na^+ ions and water molecules into the tissue. The presence of the GAGs is the reason for the high water content of AC.

Chondrocytes are round or oval cells with a mean diameter of 13 μm [44, 45]. Their size and shape vary in the different layers

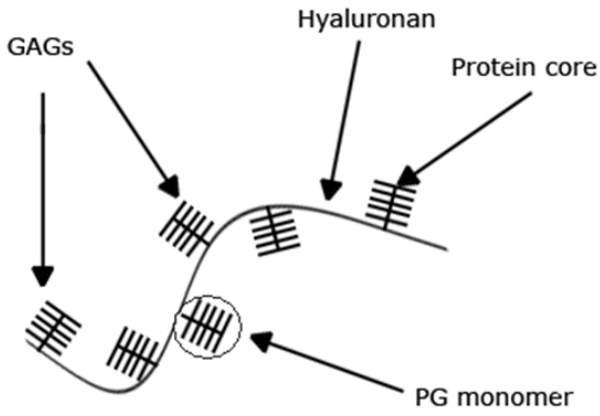


Figure 2.1: The structure of aggrecan. PG monomer is formed by a protein core and GAG side chains. A PG aggregate is formed when PG monomers attach to hyaluronan.

of AC. Chondrocytes occupy less than 5% of the total volume of AC [1,44]. The main function of chondrocytes is to synthesize and catabolize the extracellular matrix components [45,46].

Interstitial water is the main constituent of the fluid phase of AC. It constitutes 60-85% of the tissue wet weight [1, 2, 35]. The amount of water depends on the PG content and the properties of collagen network. The negative charge of GAG side chains in PGs attract water in the tissue while the collagen network limits the volume. Water content is the highest in the superficial zone of AC and decreases with the cartilage depth [5, 35]. Water plays an important role in the functional properties of AC.

Histologically, cartilage can be divided into four zones based on the collagen fibril orientation (Figure 2.2). The superficial zone is a thin zone on top of the cartilage. It comprises 5-10 % of the cartilage thickness [47]. The collagen content is usually very high in the superficial zone, while the PG content is at its lowest [4,5,48–51]. The collagen fibrils are oriented in parallel to the cartilage surface in the superficial zone [6,52]. Chondrocytes are small and elliptical, and the cell density is relatively high.

The middle zone accounts for 5-20% of the cartilage thickness

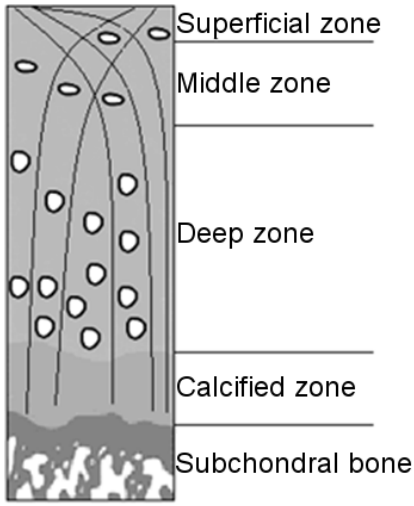


Figure 2.2: The structural layers of AC. AC is usually divided into four zones: superficial zone, middle zone, deep zone and calcified cartilage. The collagen fibril orientation in different layers is illustrated in the figure. The fibrils are attached to the subchondral bone.

[47]. On average, the orientation of collagen fibrils is random as the fibrils arch from a tangential to a radial orientation in the middle zone [6, 52]. The PG content increases in the middle zone [48–51]. The chondrocytes are round and the cell density is lower than in the superficial zone.

The deep zone occupies 70-90% of the total cartilage thickness [47]. The collagen fibril orientation is perpendicular to the cartilage surface in the deep zone [6, 52]. The PG content is highest in the deep zone [50, 51]. The chondrocytes are round and arranged in columns.

The calcified zone is adjacent to the subchondral bone. The separating line between the uncalcified and calcified cartilage is called the tidemark. Collagen fibrils are anchored to the bone by the calcified cartilage [2, 53]. The calcified zone contains only a few chondrocytes and they are metabolically inactive.

2.2 BIOMECHANICAL PROPERTIES

Biomechanical properties of AC are a result of its composition and structure [1]. PG molecules are entrapped within the collagen network and compressed to a fraction of their natural volume in aqueous solution [54]. PGs contain numerous negatively charged GAG side chains, and therefore attract free ions and water into the tissue. This creates a swelling pressure, which is resisted by the surrounding collagen network [55]. PGs are thought to be mainly responsible for the compressive stiffness of AC [1, 4], while the collagen fibrils determine the tensile stiffness of AC [56–58].

AC has to withstand both static and dynamic loading conditions in everyday life. When a compressive load is applied to AC, the pressure increases and the interstitial water flows within and out of the tissue as tissue is compressed. Due to the low permeability of the tissue, the water flow is relatively slow. During static loading, *e.g.*, as in knee cartilage when standing, an equilibrium can be reached after a sufficient amount of water has been squeezed out to achieve a balance between the external loading force and the osmotic pressure. The equilibrium modulus describes the stiffness of AC at equilibrium. After the load is released, the tissue becomes rehydrated to achieve its original state. With highly dynamic loading, *e.g.*, during locomotion, the interstitial water does not have enough time to flow out of the tissue. In this situation, the collagen network controls the behaviour of AC as it resists changes in tissue volume, creating a high hydrostatic pressure within the tissue [13, 16, 59]. The dynamic modulus describes the stiffness of AC during high-rate loading. Typically, the dynamic modulus is approximately ten times higher than the equilibrium modulus [60, 61].

Numerical biomechanical models may be used to understand the biomechanical behaviour of AC. The first proposed model was a single phasic elastic model [62]. Subsequently, a biphasic model which took into account both solid and fluid phases, was developed [63]. The biphasic model forms the basis for most of the current AC models. The isotropic biphasic model assumes that the solid

matrix is isotropic, linearly elastic and incompressible, while the fluid is assumed to be incompressible [53]. Transversely isotropic biphasic model assumes AC to be isotropic in the planes parallel to the cartilage surface [64]. In fibril-reinforced models, the collagen fibril network is separated from other solid matrix components (PGs), *i.e.*, fibril-reinforced models take into account the collagen network architecture [23, 65–68]. Triphasic model, which incorporates also ion flow, has been introduced [69, 70]. If one compares these models, then it seems that the triphasic fibril-reinforced models may be the most realistic [71].

2.3 OSTEOARTHRITIS

Osteoarthritis (OA) is a joint disease that causes pain and joint immobility, and is a major economic burden for society [8]. While the causes of OA are not fully understood, it seems that the most significant risk factors of OA are aging, obesity, joint injuries and genetic factors [72]. The main clinical signs of OA include joint pain and limitations of joint movement [8,14]. However, these signs usually do not occur until OA is already at an advanced phase, and irreversible damage has already taken place.

The degenerative signs of OA in AC are loss of superficial PGs, fibrillation of the superficial collagen network and an increase in the water content [9,73]. Compositional and structural changes lead to softening of the tissue [7], which makes it prone to suffer further damage provoked by mechanical loading. A thickening of subchondral bone is also associated with OA [73,74]. It has been speculated that OA might originate from subchondral bone. However, it seems more likely that degenerated AC distributes stresses differently to the subchondral bone, which reacts by becoming thicker [75].

AC has a limited capability to heal any degenerative changes caused by OA or injuries [76]. In OA, the chondrocytes are not able to synthesize the extracellular matrix molecules at the rate to compensate for their depletion [77]. Therefore, the cartilage degeneration triggered by OA can eventually lead to a complete loss of

cartilage. Currently, there is no cure for OA. In early OA, the focus is on pain relief and the preservation of joint mobility [72]. Surgical treatments can be used to repair more severe cartilage damage, although the outcome of the operation varies from patient to patient [78, 79]. A total joint replacement is used when the pain can no longer be relieved and the joint function is almost completely lost [72].

3 Infrared spectroscopy

Infrared (IR) spectroscopy is a traditional method used in the chemical sciences to determine the chemical composition of samples. It is suitable for solids, liquids as well as gases. In the traditional transmission mode, a spectrum of IR light passes through the sample, and the energy loss at different wavelengths is recorded.

3.1 PHYSICAL BACKGROUND

3.1.1 Absorption of infrared light

IR light is electromagnetic radiation with wavelengths longer than visible light. IR light covers wavelengths from $\lambda = 0.75$ μm to 1000 μm, but IR spectroscopy usually refers to the mid-IR region (2.5 - 25 μm). In IR spectroscopy, the wavenumber presentation is used instead of wavelengths. Wavenumber is the reciprocal of wavelength:

$$\tilde{\nu} = \frac{1}{\lambda}. \tag{3.1}$$

Every molecular bond has its characteristic resonance frequency. The resonance frequency depends on the structure of the molecule, most importantly on the type of the bond and the masses of the atoms. Since the energy of IR light is of the same magnitude as the resonance frequency, resonance frequencies can be studied by investigating the IR absorption properties of molecules.

There are three classes of molecular vibrations: stretching, bending and libration vibrations. The stretching vibration occurs when the length of a chemical bond changes. Stretching vibration can be symmetric or asymmetric. The angle of the bond changes in bending vibrations. There are six types of bending vibrations: deformation, rocking, wagging, twisting, out-of-plane bending and in-plane bending. Libration is a repetitive motion in which the molecule rotates back and forth in a nearly fixed orientation.

Two conditions must be fulfilled in order for a molecule to absorb IR light: 1) the energy of the IR light should be equal to the difference between an excited state and the ground state of the molecule and 2) the vibration should lead to a change in the net electric dipole moment:

$$\bar{p} = \sum_{i=1}^{\mathrm{n}} q_i r_i, \tag{3.2}$$

where q_i is the electric charge and r_i the distance of the electric charge from a reference point. When both of these conditions are fulfilled, the molecule will absorb IR energy which causes it to vibrate.

A simple example is a diatomic molecule constituting of atoms with equal magnitude but opposite electrical charges. In this case, the electric dipole moment changes as the bond stretches. Therefore, it is an infrared active vibration. However, if the charges are equal, the stretching does not change the electric dipole moment, and the vibration would be infrared inactive.

The Beer-Lambert law relates the amount of transmitted (or absorbed) IR light to the properties of the absorbing material. For transmittance (T), the Beer-Lambert law is

$$T = \frac{I}{I_0} = e^{-\varepsilon c l}. \tag{3.3}$$

where I_0 is the intensity of light entering the sample, I is the intensity of light transmitted through the sample, ε is the molecular absorption coefficient which describes the absorption properties of the molecule, c is the concentration of the absorbing molecule and l is the optical path length (or thickness of the sample). Often the absorbance (A) format is used in IR spectroscopy. For absorbance, the Beer-Lambert law is

$$A = -\ln \frac{I}{I_0} = \varepsilon c l. \tag{3.4}$$

In IR spectroscopy, multiple wavelengths are investigated at once. Therefore, the absorbance can be written as a function of wavenumber:

$$A(\tilde{\nu}) = \varepsilon(\tilde{\nu}) c l. \tag{3.5}$$

The relation between the absorbance and the concentration of absorbing material is linear according to the Beer-Lambert law. However, the Beer-Lambert law is valid only under ideal conditions. The sample needs to be homogeneous and there should not be any scattering. Therefore, it is not directly applicable in most situations when heterogeneous tissues are being studied.

3.1.2 Scattering of infrared light

In IR spectroscopy, the absorption is assumed to be the primary phenomenon that occurs when IR light interacts with the sample. However, the scattering effects are also seen in IR spectra and need to be taken into account. Scattering is mainly seen as baseline variations and as an increase in the optical path length, but also peak shifts might occur. Scattering is dependent on the particle size of the scatterer. This is traditionally minimized by homogenizing and grinding the sample to achieve a small particle size. However, grinding is not always an option. One particular case is IR microspectroscopy in which it is the heterogenic composition of the sample which is the focus of interest.

Elastic scattering can be divided into different sub-types based on the size of the scattering particle and wavelength of the light. Rayleigh scattering occurs when the particle is small as compared to wavelength of light. The intensity of Rayleigh scattered light is strongly dependent on the wavelength:

$$I_s \sim I_0 \frac{1}{\lambda^4}, \tag{3.6}$$

where I_s is the intensity of the scattered light. Rayleigh scattering is weak due relation of the inverse fourth power with wavelength [80].

Mie scattering occurs when the size of the scatterer is about the same size as the wavelength of light. Mie scattering has been shown to be problematic in IR microspectroscopic studies [81–85]. Mie scattering refers to scattering of electromagnetic radiation by spheres. Mie scattering can be approximated relatively simply by

the equation

$$Q(\lambda) = 2 - (4/\rho)\sin\rho + (4/\rho^2)(1 - \cos\rho), \tag{3.7}$$

where Q is the efficiency factor of scattering and

$$\rho(\lambda) = 4\pi a(n-1)/\lambda, \tag{3.8}$$

where a is the radius of the sphere, n is the ratio of refractive indices inside and outside of the sphere, and λ is the wavelength of the light. This approximation is accurate to within 1% of the results predicted by the full Mie theory [86]. Scattering factor curves calculated using equation (3.7) are shown in Figure 3.1.

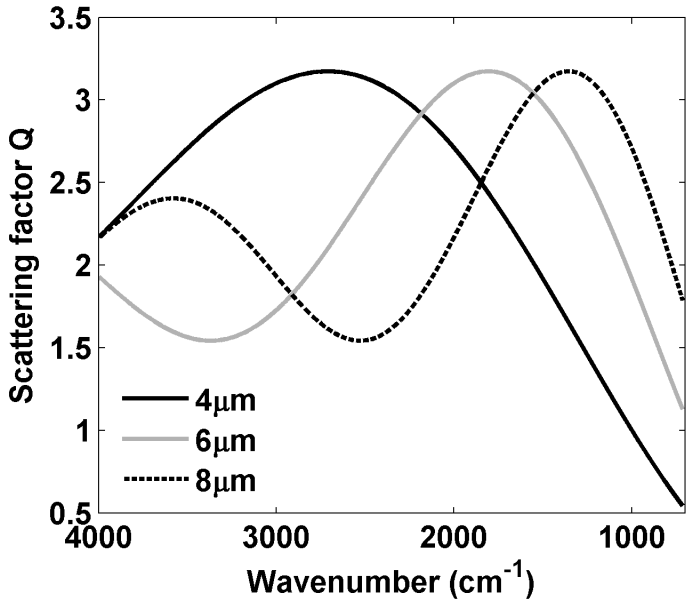


Figure 3.1: Mie scattering factors simulated with three different scatterer sizes. A refractive index of 1.3 and scatterer sizes of 4, 6 and 8 μm were used in the calculations.

Mie scattering is a particular problem in microspectroscopic studies of histological sections. The dimensions of cells and cell components are of the appropriate size that Mie scattering can occur when IR light interacts with them. In particular, chondrocytes or chondrocyte organelles may be significant sources of Mie scattering in IR microspectroscopic studies of AC. The mean diameter of chondrocytes is 13 μm [44, 45], which is within the range of Mie scattering when mid-IR light is considered.

3.2 INSTRUMENTATION

The first commercial IR spectrometers became available in the 1940s. The early systems were dispersive, *i.e.*, the broadband light was dispersed into separate wavelengths by using a prism or a diffraction grating. These systems were slow and had a poor signal-to-noise ratio. A major improvement was seen when the Fourier Transform spectrometers were introduced in the late 1960s [87]. Practically all modern IR spectrometers are Fourier Transform Infrared (FT-IR) spectrometers. Significant improvements in signal-to-noise ratio and the measurement time emerged with the use of the interferometer, which allowed simultaneous collection of all wavelengths [88].

The main components of a modern FT-IR spectrometer are schematically shown in Figure 3.2. IR radiation is produced by heating a radiation source. Nowadays the most common type of radiation source is Globar. This is constructed out of silicon carbide (SiC), and it acts approximately like a Planck radiator. Globar is typically heated to over 1,000 °C [87]. The generated radiation is passed through a semi-reflecting film called a beamsplitter at an 45 ° angle of incidence. Beamsplitters are typically made of potassium bromide (KBr) that has been coated with germanium (Ge) [87]. The beamsplitter reflects 50% of the radiation into a static mirror while the other 50% passes through the beamsplitter to a moving mirror. Both mirrors reflect the radiation back to the beamsplitter where they undergo interference. This interference can be constructive or

destructive depending on the position of the moving mirror and the phase difference of the interfering waves. The united beam is then passed through the sample. The signal collected by the detector is called an interferogram (Figure 3.3A). An IR absorption spectrum (Figure 3.3B) is obtained by calculating Fourier transformation of the interferogram.

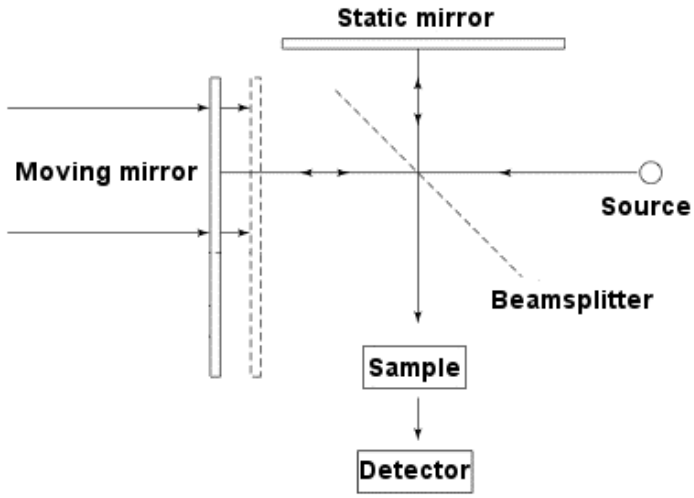


Figure 3.2: Schematic presentation of an FTIR spectrometer. Radiation is passed from the source to the beamsplitter. The radiation is reflected back from the mirrors and combined again. Interference occurs due to the pathlength difference. The radiation is then passed through the sample to the detector.

IR microspectroscopy is an extension of traditional IR spectroscopy. This combines an IR spectrometer with a microscope, thus enabling the study of samples at a spatial resolution of a few micrometers. The IR microscopes work also with visible light. Visible light image is used when the region of interest is defined. Single points or a larger region of interest can be selected and measured automatically [17, 89].

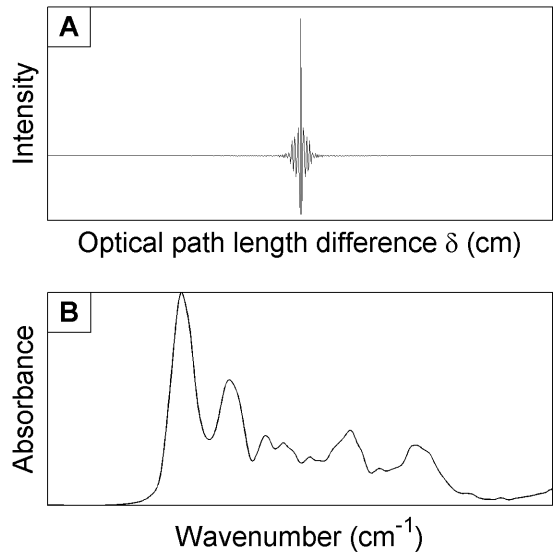


Figure 3.3: A) An interferogram and B) the corresponding IR spectrum of bovine AC.

3.3 SPECTRAL PREPROCESSING

The need for spectral pre-processing arises from the scattering effects, improper background correction and instrumental drift. The instrumental drift might originate from the shifts of the detector or the source, the changes in source temperature or functional errors in the interferometer [90]. These factors induce errors in the spectra that can be seen as baseline variations. The basic types of baseline errors are constant, linear and quadratic errors.

The constant error is removed simply by subtracting the minimum value of the spectrum from all wavenumber channels. The linear error can be removed by fitting a line between two points of the spectrum with no expected absorbance and subtracting the fitted line from the spectrum. In addition, polynomial baseline fits are used, but the proper choice of baseline points is not straightforward.

Model-based pre-processing, especially Extended Multiplicative

Signal Correction (EMSC), has become more popular in recent years [91–94]. This technique can be used to correct the typical baseline errors (constant, linear and quadratic) as well as other interference effects of which one has *a priori* knowledge, such as Mie scattering [81–84, 95, 96]. EMSC model is built around a reference spectrum $m(\tilde{\nu})$. The reference spectrum can be an error free spectrum chosen from the sample set or the average spectrum of the data set. Any spectrum can then be written as

$$z(\tilde{\nu}) = a + b \cdot m(\tilde{\nu}) + d \cdot \tilde{\nu} + e \cdot \tilde{\nu}^2 + \epsilon_r(\tilde{\nu}), \tag{3.9}$$

where the spectrum $z(\tilde{\nu})$ is a linear combination of a baseline shift a, a multiplicative effect b times a reference spectrum $m(\tilde{\nu})$, linear and quadratic wavenumber-dependent effects $d \cdot \tilde{\nu}$ and $e \cdot \tilde{\nu}^2$. The term $\varepsilon_r(\tilde{\nu})$ is the residual [90,97]. The parameters a, b, d and e can be estimated by the least-squares method, and the corrected spectrum can then be calculated with the following equation:

$$z_{corr}(\tilde{\nu}) = (s - a - d \cdot \tilde{\nu} - e \cdot \tilde{\nu}^2)/b. \tag{3.10}$$

3.4 ANALYSIS TECHNIQUES

3.4.1 Univariate analysis

Univariate analysis is the simplest spectral analysis method. A single variable at time is investigated in a univariate analysis [88]. The variable can be height, width or integrated area of an absorption peak (Figure 3.4). In addition, a ratio of the heights or areas of two different absorption peaks can be calculated. Univariate analysis methods offer a fast and straightforward way to visualize large IR imaging data sets. The univariate analysis assumes that an absorption peak can be linked to a single chemical component. Unfortunately, this is most often not the case in biological samples because of significant spectral overlap between different tissue constituents. Therefore, univariate analysis is not always an appropriate option.

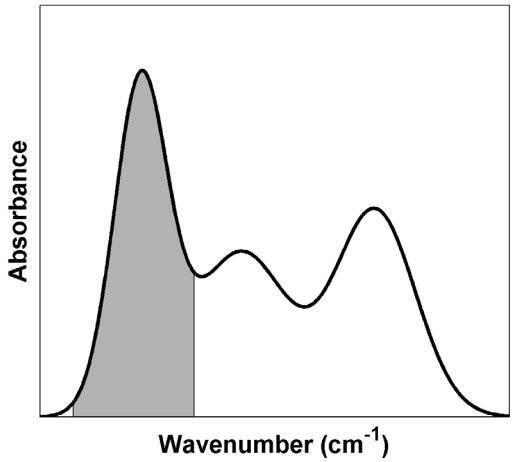


Figure 3.4: Univariate analysis can be carried out, *e.g.*, by calculating the integrated area of an absorption peak.

3.4.2 Curve fitting

Curve fitting is a technique that can be used to decode the overlapping absorption peaks into sub-peaks that might contain more detailed information about the composition of the sample (Figure 3.5). The sub-peak shape is usually approximated by a Gaussian, Lorentzian or a Gaussian-Lorentzian mixture peak shape. The locations of the sub-peaks can be found by locating the local minima in the second derivative spectrum. The other parameters of the sub-peaks (width and height) are optimized to minimize the root-mean-square difference between the measured spectrum and the sum of the fitted sub-peaks [88]:

$$\min\left(\sqrt{\left(z(\tilde{\nu})-\sum_{i=1}^{\mathrm{n}} s_i(\tilde{\nu})\right)^2}\right), \tag{3.11}$$

where z is the measured spectrum and s_{i} are the fitted sub-peaks. This technique has been utilized in different experiments, *e.g.*, in secondary structure analysis of proteins [98–101], for determining

biochemical changes in tumors [102], for determining collagen type [103,104] and in collagen cross-link analysis [105–110].

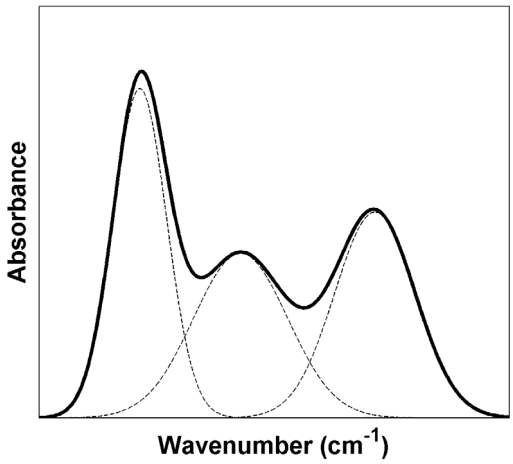


Figure 3.5: An illustration of curve fitting of three overlapping peaks. The fitted peaks can be analyzed separately to obtain more detailed information.

3.4.3 Second derivative spectroscopy

Derivative spectra are routinely used in IR spectroscopy [91,92,111–114]. In particular, second derivative spectra have been found to be useful since the second derivative can resolve adjacent overlapping absorption peaks. The peaks of the original absorption spectrum are seen as local minima in the second derivative spectrum (Figure 3.6). The use of second derivative spectra also reduces the need for baseline correction as the differentiation removes some of the baseline errors. Differentiation can be thought as a pre-processing method rather than as an actual analysis method. Second derivative spectra can be analyzed using the same methods as the absorption spectra.

One major drawback in the use of derivative spectra is that the noise is amplified in the differentiation process [115]. Because of this the derivatives are practically always calculated using Savitzky-

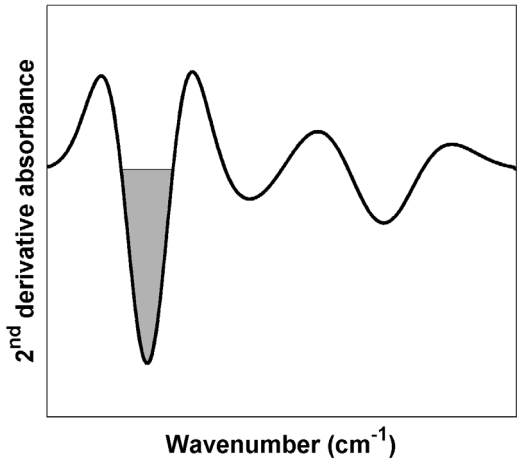


Figure 3.6: Second derivative spectrum of that shown in Figure 3.4. The negative peaks correspond with the positive peaks in the original spectrum. Peaks are narrower and better separated from each other in the second derivative spectrum.

Golay algorithm in order to reduce noise. The Savitzky-Golay algorithm performs a least-squares fit of a polynomial of degree k over at least $k+1$ data points around each point in the spectrum to smooth the data. The value of derivative at the respective point is then found by differentiating the fitted polynomial at each point [116, 117]. The smallest sub-peaks might be lost if an excessively wide fitting window is used. Therefore, the parameters have to be chosen carefully.

3.4.4 Multivariate regression

Multivariate analysis methods utilize more than one variable of the spectra. The simplest multivariate regression method is Multiple Linear Regression (MLR), which can be expressed as

$$y_i = \beta_0 + \beta_1 \cdot x_{i1} + \cdots + \beta_m \cdot x_{im} + \varepsilon_i, \tag{3.12}$$

where β are the regression coefficients, x are the input variables and y is the measured variable. This can also be written using matrix

presentation:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \varepsilon. \tag{3.13}$$

MLR is an effective regression method, but there is a risk of multicollinearity, *i.e.*, two or more of the chosen variables are highly correlated. This might lead to overfitting, making the model instable.

Principal Component Regression (PCR) is similar to MLR, but instead of using the measured variables directly, it uses so called principal components (PCs). PCR is a bilinear model. A bilinear calibration can be described with two matrix equations:

$$\begin{aligned} \mathbf{X} &= \mathbf{T}\mathbf{P}' + \mathbf{E} \\ \mathbf{Y} &= \mathbf{T}\mathbf{Q}' + \mathbf{F}, \end{aligned} \tag{3.14}$$

where $\mathbf{T}$ is the score matrix, loading matrices $\mathbf{P}$ and $\mathbf{Q}$ represent the regression coefficients of $\mathbf{X}$ and $\mathbf{Y}$ on $\mathbf{T}$, and $\mathbf{E}$ and $\mathbf{F}$ are the residuals. A linear transformation is performed in principal component analysis so that a new set of uncorrelated variables (PCs) are found. These uncorrelated variables are used in the regression model and the multicollinearity problem is avoided. The PCs are constructed to explain variance in measured data, and they are ordered so that the first PC explains the most of the variance, the second PC the second most and so on. A consequence of this arrangement is that usually only a couple of the first PCs can be considered to contain actual information while the rest can be considered to be noise [118].

Partial Least Squares Regression (PLSR) is another bilinear calibration model. Unlike in PCR, also $\mathbf{Y}$ variables are used when the decomposition is performed. In PLSR, the variables $\mathbf{T}$ are constructed to explain co-variance between the independent data (spectra) and the dependent data (information predicted from the spectra) [118, 119]. Consequently, PLSR tolerates a measurement error in $\mathbf{Y}$ better than PCR.

Different methods have been developed to decompose the matrices $\mathbf{X}$ and $\mathbf{Y}$ into the form of equations (3.14). In the following section, an algorithm in case of one $\mathbf{y}$ variable (PLS1) is described.

The matrix $\mathbf{Y}$ is now replaced by the vector $\mathbf{y}$. First, $\mathbf{X}$ and $\mathbf{y}$ are mean-centered. For the first PLS component, $j = 1$, $\mathbf{X_1} = \mathbf{X}$ and $\mathbf{y_1} = \mathbf{y}$. The following algorithm is run for the desired number (g) of PLS components:

$$
\begin{aligned}
&1)\mathbf{w_j} = \mathbf{X_j'y_j}/\|\mathbf{X_j'y_j}\| \\
&2)\mathbf{t_j} = \mathbf{X_j w_j} \\
&3)\mathbf{q_j} = \mathbf{y_j't_j}/(\mathbf{t_j't_j}) \\
&4)\mathbf{p_j} = \mathbf{X_j't_j}/(\mathbf{t_j'tj}) \\
&5)\mathbf{X_{j+1}} = \mathbf{X_j} - \mathbf{t_j p_j'} \text{ and } \mathbf{y_{j+1}} = \mathbf{y_j} - \mathbf{t_j q_j}. \\
&6)\text{Stop if } j = g; \text{ otherwise } j = j + 1, \text{ return to 1.}
\end{aligned}
\tag{3.15}
$$

Matrices $\mathbf{W}$ (an additional set of loadings), $\mathbf{P}$ and $\mathbf{T}$ are then formed by the calculated vectors $\mathbf{w_j}$, $\mathbf{p_j}$ and $\mathbf{t_j}$, and vector $\mathbf{Q}$ is formed by $\mathbf{q_j}$. The regression vector can now be calculated:

$$\beta = \mathbf{W}(\mathbf{P'W})^{-1}\mathbf{Q}. \tag{3.16}$$

3.4.5 Genetic algorithm

Genetic algorithms are variable selection methods inspired by the theory of evolution. Genetic algorithms can be used with multivariate regression methods. The genetic algorithm tries to find the most useful variables for the regression problem instead of using the full spectral window. The genetic algorithm begins with an initial population consisting of multiple possible solutions to the variable selection problem. These solutions are called chromosomes. Chromosomes are binary vectors consisting of ones and zeros, where 1 means that the corresponding variable is selected. Each one or zero is called a gene. The population size is typically between 20-500 and this stays constant during calculations. Each chromosome is evaluated mathematically, *e.g.*, by calculating the root-mean-square error of the prediction for the regression model made using the variables of the chromosome. The initial population produces an offspring by recombining the initial chromosomes. The

recombination is made so that the best chromosomes have a better chance of being copied. In recombination, cross-over of chromosomes and mutation are used to produce new chromosomes. In the cross-over, randomly selected parts of two chromosomes are interchanged. A mutation is a change in a single gene. If elitism is used, the best solution of each generation is passed on to the next generation without any changes. The genetic algorithm is usually run for a pre-defined number of generations or until some other stop criterion is fulfilled [120–122].

3.4.6 Cluster analysis

Cluster analysis techniques can be used to reveal qualitative differences between the spectra. Cluster analysis divides data into groups so that the samples inside a group are as similar as possible while the data between the groups differ from each other. K-means, Fuzzy c-means and Hierarchical Cluster Analysis are some of the most popular clustering methods in use in IR spectroscopic studies [123,124]. For example, cluster analysis can be used to separate healthy and diseased specimens [125–127], or to reveal tissue morphology in IR microspectroscopic studies [128–130] solely based on the spectral information.

3.5 INFRARED SPECTROSCOPY IN CARTILAGE RESEARCH

IR microspectroscopic study of AC began in 2001 when two pioneering studies were published by Camacho *et al* and Potter *et al* [20, 21]. The first study presented univariate parameters with which to quantify collagen and PG contents in AC. The amide I (1584 - 1720 cm^{-1}) was shown to correlate with the collagen content and the carbohydrate region (984-1140 cm^{-1}) correlated with the PG content in pure compound mixtures of collagen and aggrecan [20]. A subsequent study suggested that PG quantification could be improved by normalizing the carbohydrate region by amide I in order to reduce the thickness variation in the prepared cartilage sec-

tions [29]. That study used tissue engineered cartilage and presented correlations with optical density of Alcian blue staining and dimethylmethylene blue (DMMB) staining. A statistically significant correlation was found with Alcian blue staining but not with DMMB [29]. Later, a statistically significant correlation was found between the DMMB staining method and the integrated area of the carbohydrate region [131]. Since their introduction, the univariate methods have been applied in several AC studies. Depletion of PGs and decreased integrity of collagen has been seen in OA studies using univariate parameters [22, 25, 26, 132]. Decreased integrity in arthritic human AC was revealed also by an intra-articular fiber optic probe [133]. Furthermore, the clinical outcome of autologous chondrocyte implantation in human AC was shown to correlate with the PG content and the collagen integrity [134]. The specificity of the univariate parameters in human AC has been recently questioned [135]. In an attempt to increase the specificity for collagen, enzymatic removal of PGs can be used before conducting the measurements [30, 136, 137].

The second pioneering approach used pure compound spectra of collagen and PGs (chondroitin sulphate or aggrecan) to decompose measured IR spectra of AC. The first method used the euclidean distance between a cartilage spectrum and pure compound spectra to obtain relative concentration of collagen and PGs. The spectra are normalized before the calculations. In general, euclidean distance between spectra z_1 and z_2 is calculated as follows

$$D(z_1, z_2) = \sqrt{\sum_{i=\tilde{\nu}_1}^{\tilde{\nu}_n} [z_1(i) - z_2(i)]^2}, \tag{3.17}$$

where $[\tilde{\nu}_1, \tilde{\nu}_n]$ is the wavenumber range in use. Euclidean distance is small when the spectra are similar to each other [21].

The second multivariate method uses the linear combination of chosen pure compound spectra to decompose cartilage spectrum ($z_{\text{cartilage}}$). When two pure compounds, z_{collagen} and z_{PG}, are used,

the equation is

$$z_{\text{cartilage}} = c_{\text{collagen}} \cdot z_{\text{collagen}} + c_{\text{PG}} \cdot z_{\text{PG}} + \epsilon, \tag{3.18}$$

where c_{collagen} and c_{PG} are the concentrations of corresponding pure compounds, and ϵ is the unmodelled residual.

In two studies, type II collagen (s_{collagen}) and chondroitin sulphate (s_{CS}) were used as pure compounds in the linear combination model [21,27]. The tissue-engineered cartilage was found to contain more collagen and less PGs than the native cartilage [21] and focal degenerative lesions in human osteoarthritic AC contained less PGs than the surrounding healthy tissue [27].

Polarized IR light can be used to detect orientation of molecular bonds. The polarized IR light studies of AC have revealed that the intensities of amide I, amide II and amide III regions vary strongly when polarization plane is altered [20, 24, 31, 32, 138], whereas the sugar region shows only weak anisotropy in the radial zone of AC [32,139]. It is known that the transition moments of the amide I and II bonds are qualitatively perpendicular to each other [24,140]. This has been utilized to assess the orientation of the collagen fibrils by calculating the ratio of amide I to amide II peaks under polarized IR light. The collagen fibril orientation was seen to be abnormal in equine repair cartilage after a full-thickness chondral defect, as the orientation of the collagen fibrils was random in all regions except in the superficial layer [26].

The relative collagen and PG contents in bovine nasal cartilage were predicted by building a PCR model using mixtures of collagen and chondroitin sulphate. Biochemical analysis was also performed for cartilage samples in order to confirm these results [33]. Later the same PCR model was applied to AC to examine depth-dependent concentration profiles of collagen and PGs in AC [141]. A PLSR model was used in an intra-articular fiber optic probe study when early-stage degradation of human AC was evaluated. A strong correlation between the PLSR model and the histological OA grading was revealed [142]. A PLSR model was also created to monitor the OA progression in a rabbit model after ligament transection and

medial menisectomy [25].

The peak height ratio of 1660/1690 cm^{-1} has been used for analyzing collagen maturity in bone [19, 105–110]. Recently, the peak height ratio was used to evaluate the maturity of cross-links in repair tissue in rabbit AC following healing of full-thickness osteochondral defects [143]. The maturity was initially greater in the repair tissue before reaching the levels present in control tissue. However, the result was inconsistent with biochemically determined cross-link levels. Later, the peak height ratio was also used for characterization of a cartilage-like engineered biomass in an attempt to identify calcification of the tissue by comparing this ratio with the values from normal cortical bone [144].

Cluster analysis was recently used to reveal histological layers of AC based on IR microspectroscopic data. The fuzzy C-means algorithm was applied to the IR spectra of bovine and rabbit AC samples. The results were similar to the structural layers found using polarized light microscopy. It was speculated that the clustering was mainly a result of varying collagen-to-PG ratio in the different layers of AC [145].

The origins of IR absorption peaks have been characterized for biological tissues. Some uncertainty and overlap exist in cases where there are many peaks. Therefore, the peak assignments should only be regarded as suggestive. A list of possible peak assignments in AC is shown in Table 3.1.

Table 3.1: Assignment of second derivative IR peaks in AC.

Wavenumber (cm^{-1})	**Assignment of second derivative peaks**
1700-1600	Amide I region (C-O stretch)
1600-1500	Amide II region (C-N stretch + N-H bend)
1448	CH_3 asymmetric bending vibrations [91,146]
1400	COO^- stretch of amino side chains [146]
1374	CH_3 symmetric bending vibration of GAGs [147]
1336	CH_2 side chain vibrations of collagen [146]
1280	Collagen amide III vibration with significant mixing with CH_2 wagging vibration from the glycine backbone and proline sidechain [146]
1228	SO_3^- asymmetric stretching vibration of sulphated GAGs [35]
1200	Collagen amide III vibration with significant mixing with CH_2 wagging vibration from the glycine backbone and proline sidechain [146]
1120	C-O-S asymmetric stretching [148]
1080	C-O stretching vibrations of the carbohydrate residues in collagen and PGs [91,146]
1062	C-O stretching vibrations of the carbohydrate residues in PGs [91,146] / SO_3^- symmetric stretching vibration of sulphated GAGs [148]
1032	C-O stretching vibrations of the carbohydrate residues in collagen and PGs [91,146]

4 Aims of the study

IR microspectroscopic studies of AC have been performed for over 10 years and this technique has been taken into routine use in some laboratories. This thesis work evaluates the quality of the spectral analysis techniques and introduces new methods to enhance the possibilities for using IR microspectroscopy in AC research.

The specific aims of this thesis were:

- to evaluate the specificity of current univariate IR spectral analysis methods in the compositional analysis of AC,
- to investigate the IR spectroscopic changes caused by PG depletion in AC,
- to improve the IR spectroscopic analysis of AC composition through the use of curve fitting, second derivative spectroscopy and multivariate models,
- to determine whether it is possible to predict the compressive biomechanical properties of AC samples based solely on their IR spectra.

5 Materials and methods

This thesis consists of four independent studies (I - IV), with the focus on the development of analysis techniques for IR microspectroscopic data of AC. Digital densitometry (DD), biomechanical testing and biochemical analysis are used as reference techniques. All samples, with the exception of the cryosectioned samples in study II, have been extracted from earlier studies [4,15,60]. A summary of the methods used in the independent studies is presented in Table 5.1.

Table 5.1: Materials and methods used in the studies I-IV. All AC samples were prepared from bovine patellae.

Study	Samples	n	Methods	Parameters
I	Intact	8	IR	Univariate analysis
	Enzymatically		DD	Curve fitting
	degraded	8		Pure compound fitting
II	Fixed sections	8	IR	Univariate analysis
	Cryosections	6		2nd derivative spectroscopy
III	Intact	8	IR	Univariate analysis
	Enzymatically		DD	2nd derivative spectroscopy
	degraded	8		Multivariate analysis
IV	Spontaneously		IR	Multivariate analysis
	degraded	32	Biomechanical testing	
			Biochemical analysis	

5.1 SAMPLE PREPARATION

Bovine patellar cartilage of 1–3-year-old specimen obtained from a local slaughterhouse (Atria Oyj, Kuopio, Finland) was used in all studies. Knee joints were opened within a few hours *post mortem*. IR microspectroscopy was conducted in all studies. DD was conducted in studies I and III and biomechanical testing was conducted in study IV.

Studies I and III: Osteochondral plugs (*diam.* = 13 mm, n = 16) were prepared from the lateral upper quadrant of the patellae. The samples were kept moist with physiological saline during the sample preparation. Control samples (n = 8) were subjected to no additional processing. The other samples (n = 8) were subjected to an enzymatic degradation of PGs. The samples were incubated at 37°C for 44 h in 5% CO_2 atmosphere in a cell culture medium with antibiotics. Chondroitinase ABC enzyme was added to the medium to degrade the superficial PGs [149]. An osteochondral plug (*diam.* = 6 mm) was punched out from the center of the original sample after incubation to ensure that the enzyme degrades the PGs only from the superficial AC. Samples were fixed with 10% formalin, decalcified, dehydrated in an increasing series of ethanol solutions and embedded in paraffin (Paraplast Plus, Lance Division of Sherwood medical, Kildare, Ireland). Multiple 5-μm-thick sections were cut perpendicular to the cartilage surface with a microtome (LKB 2218 HistoRange microtome, LKB produkter AB, Bromma, Sweden). Sections were placed on standard microscope slides and immersed in xylene to remove the paraffin. Xylene was washed out by using a descending series of ethanol and distilled water. One section from each sample was placed on 2-mm-thick ZnSe window, while another section from each sample was first treated with hyaluronidase (type IV, H-3884, Sigma, St. Louis, MO, USA) for 18 h to remove the PGs [150, 151] before it was placed on ZnSe window for IR microscopic measurements.

Study II: Control samples in studies I and III were also used in this study. Additional samples (n = 6) from bovine patellae were prepared for cryosectioning in order to evaluate whether formalin-fixation affects the enzymatic removal of PGs. Samples were kept moist with physiological saline during the sample preparation. Cartilage samples were detached from the underlying subchondral bone with a razorblade. Subsequently, the samples were embedded into Tissue Tek Optimal Cutting Temperature (OCT) embedding medium (Sakura Finetek, Torrence, CA, USA). Five μm thick cryosec-

tions were cut (Reichert-Jung Frigocut 2800, Nussloch, Germany) and OCT was removed with water from the sections before transferring them onto 2-mm-thick ZnSe windows for the IR microspectroscopic measurements. After the measurements were conducted for both cryosections and formalin-fixed sections, all sections were placed back on microscope slides for the enzymatic removal of PGs. The sections were treated with hyaluronidase (type IV, H-3884, Sigma) enzyme for 18h to remove PGs [150, 151]. After the enzymatic treatment, the sections were rinsed with distilled water and transferred back on ZnSe windows. The measurements were repeated using identical measurement parameters.

Study IV: Knee joints obtained from a slaughterhouse were opened within 5 h of post mortem and thereafter the lateral facets of patellar cartilage surfaces were visually classified by two experts to four different degenerative grades: grade 0=intact cartilage surface (n = 13), grade 1=slightly discoloured but otherwise smooth (n = 5), grade 2=superficial defect in cartilage (n = 6) and grade 3=deep defect in cartilage (n = 8). Subsequently, a cylindrical osteochondral sample (*diam.* = 19 mm) was drilled from each patella and split into two halves. The first block was used for biomechanical reference measurements whereas the second block was fixed with 10% formalin, decalcified, dehydrated and embedded in paraffin. Five μm thick sections were cut perpendicular to the cartilage surface with a microtome from each sample and placed on the 2-mm-thick ZnSe window.

5.2 IR MICROSPECTROSCOPY

Measurements were conducted with a Perkin Elmer Spotlight 300 FT-IR imaging system (Perkin Elmer, Shelton, CO, USA). A CO_2-free dry air purge system (FT-IR purge gas generator, Parker Hannifin Corporation, Haverhill, MA, USA) was used during all measurements to standardize the experimental conditions.

Pure compound spectra of type II collagen, chondroitin sulphate

and aggrecan were measured and used in multivariate analyses in study I and as qualitative references in studies II-IV. The purified compound (1 mg) was mixed together with KBr powder (200 mg) and homogenized manually. The homogenized mixture was compressed with a manual press. Spectra were measured using a Perkin Elmer Spotlight 300 FT-IR imaging system in the point mode, using 4 cm^{-1} spectral resolution, 100 μm aperture and 128 repeated scans.

In study I, the cartilage sections were measured using 6.25 μm pixel size and 4 cm^{-1} spectral resolution and 4 scans per pixel. The small pixel size was used in order to image the thin superficial layer of AC accurately. In other studies (II-IV), the pixel size of 25 μm and 8 scans per pixels were used to achieve a good signal-to-noise ratio.

Pre-processing

In study I, the adjacent spectra from the 200 μm wide region-of-interest were averaged to obtain only one spectrum for every 6.25 μm thick layer in the depth-wise direction of AC. The baseline offsets of the spectra were then corrected so that the minimum value of the spectra were set to zero.

In study II, spectra of each measured section were averaged since only average changes were studied.

In study III, a data set consisting of 294 data points was assembled so that PG concentration levels according to the safranin O reference information were evenly presented. Second derivative spectra were calculated using the Savitzky-Golay algorithm with 7 smoothing points.

In study IV, the spectra of each measured section were averaged. Second derivative spectra were calculated with Savitzky-Golay algorithm with 7 smoothing points and EMSC correction was applied using equations (3.9) and (3.10).

Curve fitting

Curve fitting was performed point-by-point using a custom-made Matlab (Ver. R2007b, MathWorks Inc., Sherborn, MA, USA) software. Sub-peaks were modeled using a Gaussian peak shape:

$$G(\tilde{\nu}) = A \cdot \exp\left(-\frac{(\tilde{\nu} - \tilde{\nu}_0)^2}{2 \cdot \sigma^2}\right) \quad (5.1)$$

where A is the amplitude of the peak, $\tilde{\nu}_0$ is the location of the peak and σ is the width of the peak. The locations of the sub-peaks were found from the local minima in second derivative spectra. The other parameters were obtained by minimizing the root-mean-squared difference of the measured spectrum and the sum of the fitted peaks. The number of sub-peaks was assumed to be the same for all spectra, but the locations of the peaks were allowed to change ($\pm$ 8 cm^{-1}) from the initial values provided that the second derivative spectrum indicated a peak shift. The spectral region of 1300 - 900 cm^{-1} was used for curve fitting. The integrated absorption of each sub-peak was plotted from the superficial cartilage to the cartilage-bone junction and compared to the safranin O distribution profiles.

Univariate methods

In studies I, II and III, the integrated absorbances of amide I (1720-1585 cm^{-1}) and carbohydrate region (1140-984 cm^{-1}) were calculated to quantify collagen and PG content, respectively, in AC (Figure 5.1A). In study I, the amide I absorbance was calculated also after enzymatic removal of PGs to serve as a reference for collagen distribution in AC. In study II, both amide I and carbohydrate region absorbances were calculated also after enzymatic removal of PGs.

Pure compound methods

In study I, two pure compound-based multivariate methods, the euclidean distance and linear combination, were used for collagen and PG analysis. Type II collagen and either aggrecan or chondroitin sulphate were used as pure compounds.

Second derivative spectroscopy

In study II, the changes caused by enzymatic removal of PGs were evaluated by calculating the relative changes in second derivative

peak heights in both formalin-fixed and cryosectioned sample groups. The peaks that showed the most significant changes were assumed to be PG-related peaks, whereas the peaks that showed only minimal or no changes were considered as collagen-related peaks (Figure 5.1B). The depth-wise distribution profiles of the most interesting peaks were plotted as group means of formalin-fixed sections. For comparison, a difference spectrum was calculated by subtracting the mean absorption spectrum after the removal of PGs from the mean spectrum of the same samples before the treatment to show the changes seen in the absorption spectrum.

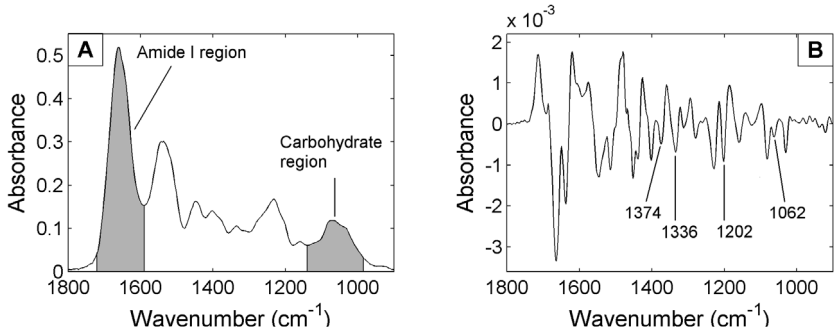


Figure 5.1: A) IR absorption spectrum and B) second derivative spectrum of bovine AC. The peaks used in the analyses are marked in the spectra.

Multivariate regression

In studies III and IV, multivariate regression models were used to predict PG content and biomechanical properties of AC from IR spectra. In study III, the optical density of safranin O was used as reference data. In study IV, the equilibrium modulus and dynamic modulus obtained from biomechanical testing were used as reference data.

In study III, spectral regions of 1000-1440 cm^{-1} and 1480-1700 cm^{-1} were used in multivariate models, whereas in study IV, spectral regions of 900-1440 cm^{-1} and 1480-1800 cm^{-1} were used in multivariate models. The region of 1440-1480 cm^{-1} was omitted since the absorption bands of paraffin residues are present in this region.

The optimal number of variables for the regression models was chosen based on the root-mean-square error of the cross-validation (RMSECV):

$$RMSECV = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n}} \tag{5.2}$$

where $\hat{y}_i$ is the predicted value and y_i is the observed value and n is the number of samples [118,152]. In leave-one-out cross-validation, each sample in turn is removed from the data to be used as a validation data. The number of variables is optimal when increasing the number of variables no longer significantly decreases the RMSECV. The performance of final models was evaluated by RMSECV and Pearson's correlation coefficient. In study III, both PCR and PLSR models were used. In study IV, PLSR model was used. In addition, the genetic algorithm was used for the variable selection.

Genetic algorithm

In study IV, a genetic algorithm was used for variable selection when the multivariate models were built. The parameters used in the genetic algorithm were as follows; the population size: 100, gene initialization probability: 5%, cross-over method: one-point, cross-over probability: 80%, mutation probability: 1%, number of generations: 100, response to be minimized: RMSECV of the prediction of the multivariate model.

The number of PCR or PLSR components for equilibrium modulus was chosen based on the full spectrum model. In the dynamic modulus, the full spectrum model used a relatively high number of components. A simpler model was preferred when the genetic algorithm was used. Therefore, the same number of components was used for both the equilibrium modulus and the dynamic modulus when the genetic algorithm was used.

There is a risk of overfitting when variables/objects ratio is too large. As a rule of thumb, the performance of genetic algorithm decreases when more than 200 variables are used [121]. Originally, the spectra contained 450 variables and there was 32 samples. To avoid the problem with overfitting, the spectra were averaged with

a window size of 5, which resulted in 90 variables. The genetic algorithm was run for 100 times and the selection frequencies of the variables were calculated. When the final model was built, variables were added to the model according to their selection frequencies. The variable combination that resulted in the minimum RMSECV was chosen as the final model.

5.3 DIGITAL DENSITOMETRY

In studies I and III, DD of Safranin O -stained AC sections was used to assess the spatial distribution profiles of PGs. Multiple three-μm-thick sections were cut from each sample. The sections were stained with 0.5% safranin O as described in our earlier study [151]. Safranin O is a cationic dye that binds stoichiometrically to the negatively charged GAGs in aggrecan [49, 153]. Therefore, the safranin O staining is an indirect estimate of the PG content in AC. The intensity of safranin O staining was measured using monochromatic light (492$\pm$5 nm), a microscope (Leitz Ortholux-II, Leitz, Wetzlar, Germany) and a 12-bit CCD camera (CH250, Photometrics, Tucson, AZ, USA). The system was calibrated using neutral density filters (Schott, Mainz, Germany). Three sections per sample were measured to reduce the sample thickness variation. The rows of the measured image were averaged to obtain a depth-dependent safranin O distribution profile for each measured section. Then all profiles of the same sample were averaged.

5.4 BIOMECHANICAL TESTING

In study IV, a custom-made material testing instrument (with resolutions of 5 mN and 0.1 μm for the force and position, respectively) was used for biomechanical reference measurements. Testing was performed using a stress-relaxation (10% prestrain followed by 10% strain with 2 mm/s ramp speed and relaxation time of 2400 s) protocol in unconfined compression geometry. Young's modulus at equilibrium and dynamic modulus were calculated as a stress-

strain ratio after the relaxation and instantaneously after a 10% step, respectively [15].

5.5 BIOCHEMICAL ANALYSIS

In study IV, uronic acid (PG content) [154] and hydroxyproline contents (collagen content) [155] were determined using biochemical reference methods.

5.6 STATISTICAL ANALYSIS

In study I, Pearson's correlation coefficient was calculated between the IR spectroscopy-derived parameters and the optical density of Safranin O for each sample pair to compare the shape of the PG distribution profiles. The non-parametric Wilcoxon signed-rank test was used to test the significance of the differences in the correlation coefficients between the different parameters.

In study II, the non-parametric Wilcoxon signed-rank test was used when the statistical significance in the observed changes caused by the removal of PGs was evaluated. The non-parametric Mann-Whitney U-test was used when the statistical significance in differences between formalin-fixed sections and cryosections was evaluated.

In study III, Pearson's correlation coefficients were calculated between the observed values and different IR spectroscopic parameters. The statistical difference between the correlation coefficients was tested by using a test described by Steiger [156]. Bonferroni correction was applied because of multiple comparisons ($n = 7$ comparisons). Therefore, the level of significance was $p < 0.05/7 = 0.007$.

In study IV, Pearson's correlation coefficients were calculated between the observed values and values predicted by the multivariate models.

In all studies, SPSS 14.0 (SPSS Inc., Chicago, IL, USA) or Matlab 2010b (Mathworks Inc., Natick, MA, USA) software was used for statistical analyses.

6 Results

6.1 UNIVARIATE METHODS

The difference spectrum revealed that the most significant changes occur in the carbohydrate region when the PGs were enzymatically removed (II). The carbohydrate region changed 35.9% ($p < 0.01$) and 31.0% ($p < 0.05$) in formalin-fixed sections and cryosections, respectively. However, the depth-wise distributions of the integrated absorbance of the carbohydrate region did not agree closely with the safranin O distributions (I). The mean ($\pm$ standard deviation) correlation for depthwise data of individual samples was 0.69 ± 0.11 and 0.55 ± 0.18 for chondroitinase ABC-treated samples and control samples, respectively (I). The deviation from the reference distributions (Figure 6.1A) was most evident in the deep region of AC (Figure 6.2A). The overall correlation with the optical density of safranin O was $r = 0.605$ ($p < 0.001$) (Figure 6.3A) (III). Correlation analysis showed that some of the largest PG concentrations were extensively overestimated by the carbohydrate region absorbance.

When the PGs were removed, the amide I region, traditionally used for collagen analysis, changed by 8.2% ($p < 0.01$) and 11.2% ($p < 0.05$) in formalin-fixed sections and cryosections, respectively (II). The depth-wise distributions of amide I after the removal of PGs are shown in Figure 6.1. When the carbohydrate region was normalized with the amide I, overestimation of high PG concentration values was not present, but the smallest PG concentrations were overestimated (Figure 6.2B). The mean correlation of individual samples was 0.65 ± 0.24 and 0.37 ± 0.19 for chondroitinase ABC-treated samples and control samples, respectively (I). The overall correlation with the optical density of safranin O was $r = 0.379$ ($p < 0.001$) (Figure 6.3B) (III).

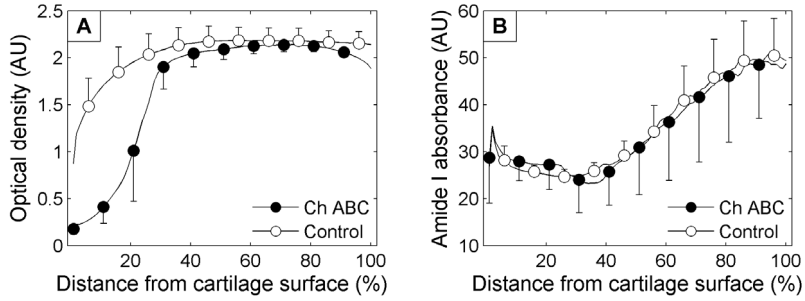


Figure 6.1: A) Safranin O distributions of control and enzymatically modified cartilage samples. B) Amide I absorbance after enzymatic removal of PGs of control and enzymatically modified samples.

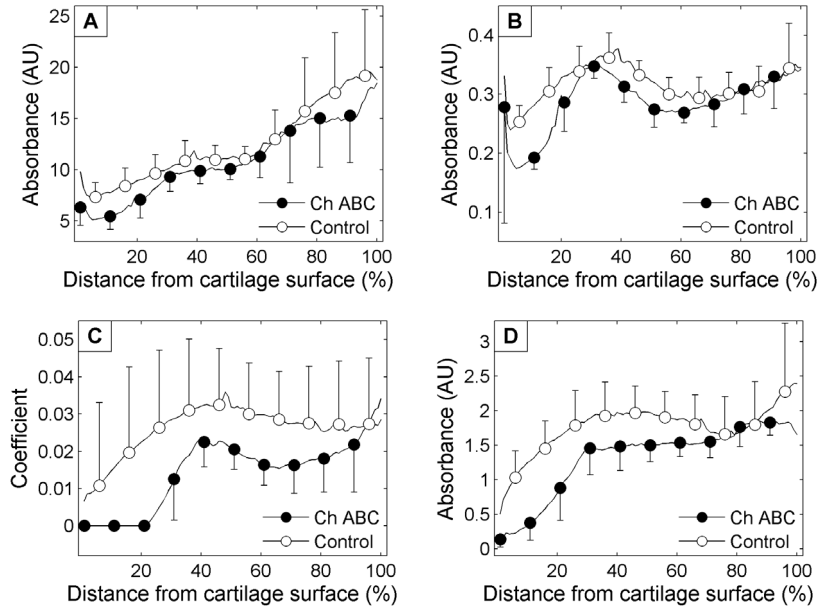


Figure 6.2: Depth-wise distributions of different IR spectroscopic PG parameters and mean correlation ($\pm$ standard deviation) with the reference safranin O distributions in chondroitinase ABC-treated samples and control samples, respectively. A) Carbohydrate region ($r = 0.69 \pm 0.11$ and $r = 0.55 \pm 0.18$), B) carbohydrate region normalized with amide I ($r = 0.65 \pm 0.24$ and $r = 0.37 \pm 0.19$), C) least-squares fit using pure compound spectra of chondroitin sulphate ($r = 0.69 \pm 0.09$ and $r = 0.55 \pm 0.31$) and type II collagen and D) curve fitting ($r = 0.91 \pm 0.06$ and $r = 0.76 \pm 0.13$).

6.2 PURE COMPOUND METHODS

The euclidean distance of aggrecan or chondroitin sulphate and AC spectra was used to estimate the relative amount of PGs in AC (I). Both spectra failed to show the complete loss of PGs in the superficial layers of chondroitinase ABC-treated samples, provided that the off-set corrected spectra were used. However, when a polynomial baseline correction was applied to the spectrum of chondroitin sulphate, the loss of PGs was revealed. Otherwise, all spectra produced similar results. No correlation analysis was performed since only the relative amount of PGs was investigated.

The linear combination of offset-corrected collagen and aggrecan or chondroitin sulphate spectra provided no meaningful results for PGs since the coefficient for PGs was systematically zero. When the polynomial baseline correction was applied to the spectrum of chondroitin sulphate, the loss of superficial PGs was revealed with the linear combination method (Figure 6.2C). However, the amount of PGs in the deep layers of AC was overestimated. The mean correlation of individual samples was 0.69 ± 0.09 and 0.55 ± 0.31 for chondroitinase ABC-treated samples and control samples, respectively.

6.3 CURVE FITTING

Curve fitting was used to decompose measured AC spectra (I). One peak, located at 1060 cm^{-1}, produced similar depth-wise distributions to that of safranin O staining (Figure 6.2D). The mean correlation of individual samples was 0.91 ± 0.06 and 0.76 ± 0.13 for chondroitinase ABC-treated samples and control samples, respectively. The other peaks in the investigated spectral region (1300-900 cm^{-1}) were strongly affected by collagen and were not considered as suitable peaks for PG analysis.

6.4 SECOND DERIVATIVE SPECTROSCOPY

The enzymatic removal of PGs was used to locate the PG-related and collagen-related second derivative peaks in AC (II). Only two peaks, located at 1062 cm^{-1} and 1374 cm^{-1}, were seen to change extensively. The peak 1062 cm^{-1} changed 49.9% and 58.0% in formalin-fixed sections and cryosections, respectively, while the peak 1374 cm^{-1} changed 61.0% and 83.7%. The PG-related second derivative peaks were correlated with safranin O-staining of parallel sections. The correlation coefficients were $r = 0.701$ $(p < 0.01)$ and $r = 0.766$ $(p < 0.01)$ between safranin O staining density and the peaks at 1062 cm^{-1} and 1374 cm^{-1}, respectively (Figures 6.3C and 6.3D) (III).

Several peaks showed only minimal changes when PGs were removed, and therefore were regarded as collagen-related peaks (II). These peaks were located at 1638 cm^{-1}, 1514 cm^{-1}, 1448 cm^{-1}, 1336 cm^{-1}, 1202 cm^{-1}. The observed changes in these peaks were not statistically significant except for the peak at 1202 cm^{-1} in formalin-fixed sections, which increased by 4.7% ($p < 0.01$).

6.5 MULTIVARIATE REGRESSION

Two multivariate models, PCR and PLSR, were built to predict PG content of AC (III). RMSECV was used to select the optimal number of components for the models. Seven components were considered to be optimal as RMSECV did not decrease significantly any longer when more components were added to the model. PCR model showed a high linear correlation with the optical density of safranin O ($r = 0.903$, $p < 0.001$) (Figure 6.3E). PLSR model even improved the correlation with the reference information ($r = 0.943$, $p < 0.001$) (Figure 6.3F).

PLSR models were built to predict equilibrium modulus and dynamic modulus of AC samples (IV). First, the models were built using the full spectrum. The best model for the equilibrium modulus was achieved with three PLS components. The correlation coefficient of the predicted values with the reference values was

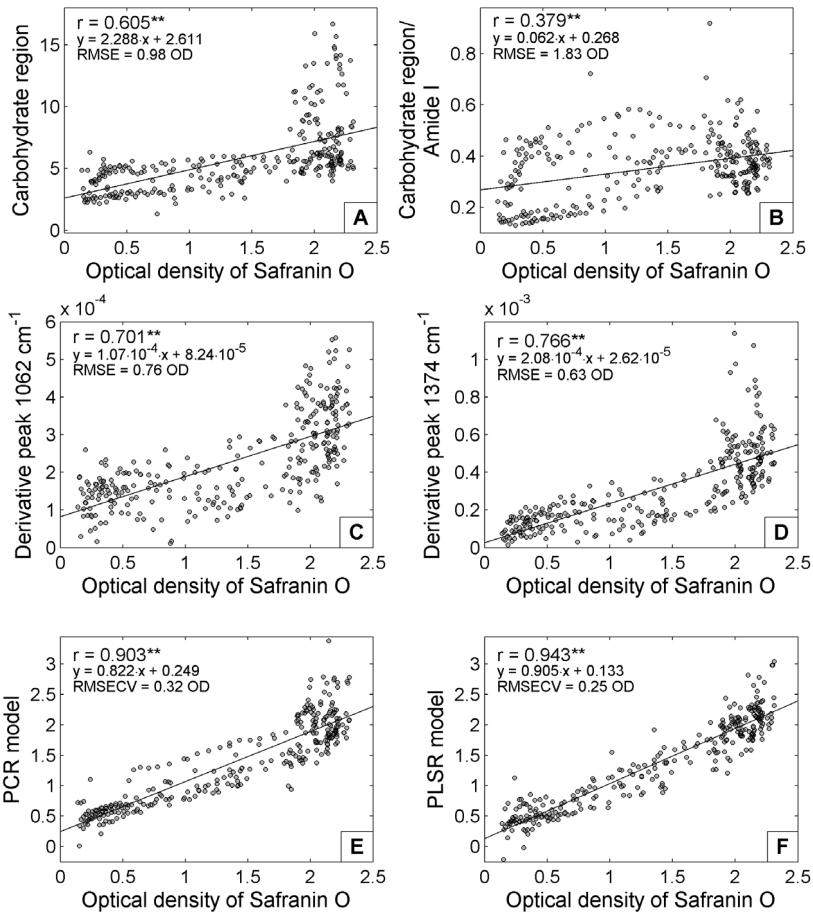


Figure 6.3: Scatter plots between the reference PG information and different IR spectroscopic parameters: A) the carbohydrate region, B) the carbohydrate region normalized with the amide I, C) derivative peak 1062 cm^{-1}, D) derivative peak 1374 cm^{-1} E) PCR model and F) PLSR model.

$r = 0.828$ ($p < 0.001$). The best model for the dynamic modulus required six PLS components. The correlation coefficient between the predicted values and the reference values was $r = 0.893$ ($p < 0.001$). However, eight of the samples were significantly stiffer than the others which increased the correlation coefficient. When the small and high values were separated from each other, the correlations were lower ($r = 0.63$ and $r = 0.55$ for lower and higher value clusters, respectively). Another model was built for a sample set from which the stiffest samples were omitted. This model predicted the dynamic moduli better ($r = 0.725$, $p < 0.001$) than the earlier model in the range of 0 - 2.5 MPa.

In order to improve the results, a genetic algorithm was used to select the variables for PLSR models. Four PLS components were used for both equilibrium modulus and dynamic modulus. A total of 16 variables (Figure 6.4A) resulted in the best prediction of equilibrium modulus ($r = 0.866$, $p < 0.001$) (Figure 6.5A). This result is slightly better than that which was obtained with the full spectrum model. In the case of the dynamic modulus (with stiffest samples removed from the data) 17 variables (Figure 6.4B) were used. A more significant improvement was revealed in the prediction of the dynamic modulus, *i.e.*, the correlation coefficient between the predicted values and the reference values was $r = 0.898$ ($p < 0.001$) (Figure 6.5B).

PLSR models were built to predict bulk composition (collagen and PGs) of bovine AC samples (unpublished). The models were first built using the full spectra and then using the genetic algorithm to select optimal variables. Four PLS components were optimal for the prediction of the PG content in the full spectral model ($r = 0.836$, $p < 0.001$). The genetic algorithm selected 11 out of 90 (Figure 6.4C) possible spectral variables to achieve the best model ($r = 0.875$, $p < 0.001$) (Figure 6.5C). The collagen content was best predicted when three PLS components were used in the full spectral model ($r = 0.765$, $p < 0.001$). The genetic algorithm selected 13 out of 90 (Figure 6.4D) possible 10-wavenumber-wide spectral variables for the best model ($r = 0.876$, $p < 0.001$) (Figure 6.5D).

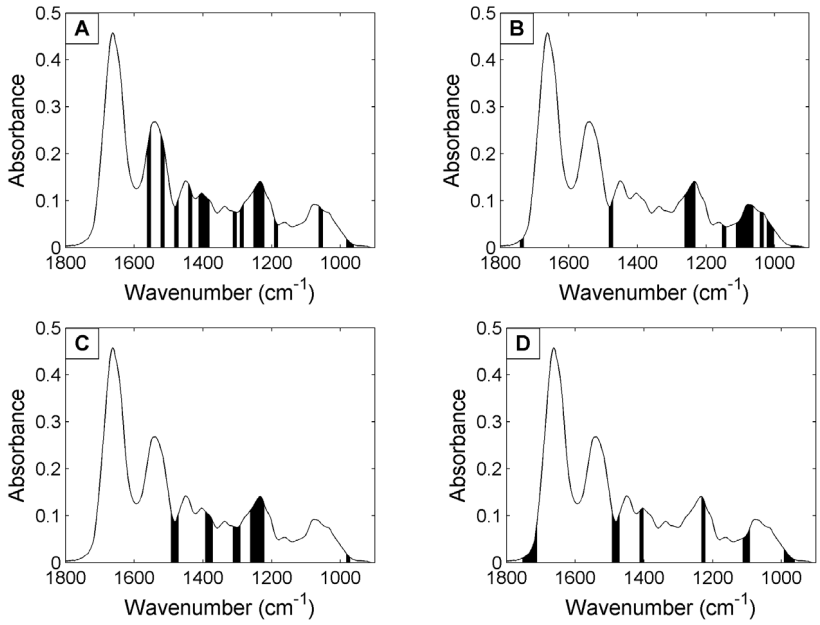


Figure 6.4: The wavenumbers selected by genetic algorithm for PLSR models to predict A) Young's modulus, B) dynamic modulus, C) PG content and D) collagen content.

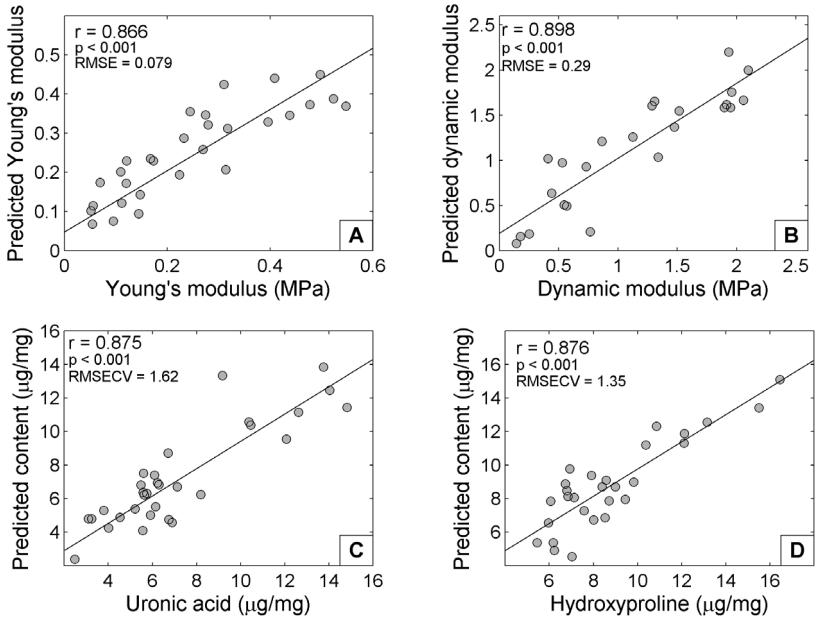


Figure 6.5: Scatter plots between the PLSR models utilizing the variables selected by the genetic algorithm and A) Young's modulus, B) dynamic modulus, C) PG content and D) collagen content.

7 Discussion

7.1 UNIVARIATE METHODS

Univariate methods are easy to implement and can be used for fast visualization of spectroscopic data sets. However, they do not utilize the measured spectral data as efficiently as would be possible. This was demonstrated in studies I, II and III. The problems originate from the significant overlap of collagen and PG vibrations as seen from their pure compound spectra. In earlier studies, the amide I correlated with the collagen content and the carbohydrate region with the PG content in pure compound preparates [20]. Later, the carbohydrate content normalized with amide I has been shown to correlate with the optical density of Alcian blue [29]. However, in study II, it was shown experimentally that the amide I absorbance contains a significant contribution from PGs, *i.e.*, amide I is not a pure measure of collagen. Similarly, although the carbohydrate region is the best option for PG quantification from the absorbance spectrum, it is still not fully specific for PGs in all situations (I, III). Instead, the depth-wise distributions of the carbohydrate region (with or without amide I normalization) differed markedly from those of safranin O, a well validated method for the determination of PG content [49,153]. One possible explanation for this discrepancy is that this study I utilized bovine AC whereas the earlier study used tissue engineered cartilage. Nonetheless, one must conclude that the univariate methods do not provide the most effective analysis of multivariate data.

7.2 PURE COMPOUND METHODS

Two methods based on the use of pure compound spectra were investigated in study I. The euclidean distance measure is suitable mainly for qualitative analyses as it evaluates the similarity between two spectra. The linear combination of pure compound spectra can

be used for quantitative analysis. However, the results were poor unless a polynomial baseline correction was applied to the chondroitin sulphate spectrum. There are two explanations for the difficulties encountered with the linear combination methods. First, the dry matrix of AC was assumed to be composed of only collagen and PGs. This is not strictly correct. Second, it was assumed that the isolated pure compounds would be spectroscopically equal to the compounds in the tissue. The isolation of pure compounds requires multiple purification steps with different chemicals which might alter their original structure. In addition, interactions between different compounds in the tissue might introduce spectroscopical differences compared to the isolated compounds. For these reasons the direct use of pure compound spectra might not always be appropriate. The need for a strong baseline correction supports the belief that the isolated pure compounds are not spectroscopically equivalent to the compounds in the tissue. The pure compound spectra offer a starting point for the model, but because of the aforementioned reasons, they do not always represent the compounds in the tissue perfectly. Possibly, a better result could be achieved if pure compound spectra would be treated only as initial approximations. For example, this can be done by using a chemometric method called Multivariate curve resolution (MCR), which adjusts the initial spectral approximations to solve the problem more efficiently [157, 158].

7.3 CURVE FITTING

In study I, curve fitting was used to decompose the overlapping absorption peaks in the carbohydrate region into separate sub-peaks in order to improve the specificity for PGs. Interestingly, only one peak in this region was seen to be strongly related to PGs. The subpeak located at 1060 cm^{-1} correlated with the safranin O staining better than the earlier univariate or pure compound methods. This peak has been assigned to both C-O stretching vibrations of the carbohydrate residues in PGs [91, 146] and SO_3^- symmetric stretching

vibration of sulphated GAGs [148]. As safranin O attaches to the negatively charged GAGs, and the sub-peak correlated with the safranin O reference information, it appears that the sub-peak 1060 cm^{-1} is more strongly related to sulphate vibrations than to C-O vibrations.

Although curve fitting improved the accuracy of PG analysis, the technique has some major limitations, especially in imaging studies. First, the second derivative spectra are typically used to locate the sub-peaks. This poses demands on the high signal-to-noise ratio as the differentiation amplifies also the noise. Noise introduces additional local minima to second derivative spectra that can be interpreted as locations of actual peaks. However, a high signal-to-noise ratio is not so critical if pre-defined peak locations are used. Second, the optimization process for obtaining height and width parameters of sub-peaks is computationally demanding. Even though the time that the optimization requires to handle one spectrum might only take one second or even less, the total time for a single sample can quickly be transformed to minutes or even hours as one IR spectroscopic image can contain thousands of spectra. Finally, the optimized curve fitting solution of experimental spectrum is probably never unique, but slightly different results will be obtained depending on which optimization algorithm is being used as well as the number of peaks to be fitted and the peak shape. Curve fitting is an effective spectral analysis method, but it requires care and experience when it is applied, and is not a feasible method for the analysis of large data sets because of the long analysis time.

7.4 SECOND DERIVATIVE SPECTROSCOPY

Second derivative spectroscopy was used for resolving nearby lying bands (II and III). It improved the separation of absorption peaks. The enzymatic removal of PGs revealed that multiple peaks were strongly linked to collagen vibrations as they showed only minor or no changes in the experiment (II). The changes reached no level of statistical significance in three of these peaks (except for 1202

cm^{-1} in formalin-fixed sections): 1202 cm^{-1} (amide III vibration of collagen), 1336 cm^{-1} (CH_2 side chain vibrations of collagen) and 1448 cm^{-1} (asymmetric bending of CH_3). Therefore, they provided the best peaks for collagen analysis. Depth-wise distribution patterns of the collagen-related peaks were similar before and after the enzyme treatment, which further indicates that these peaks originate from the collagen vibrations and not from PGs.

Many peaks have been assigned to GAGs in the earlier literature (Table 3.1). However, only two of the peaks (1062 cm^{-1} and 1374 cm^{-1}) displayed an extensive reduction in their intensity when PGs were removed. The peak located at 1374 cm^{-1}, which is assigned to CH_3 stretching vibrations of GAGs, displayed the most extensive change in the experiment. The other peak with extensive changes, 1062 cm^{-1}, has been assigned to both C-O stretching and SO_3- stretching vibrations. Interestingly, these two peaks produced slightly different depth-wise distribution patterns. Despite the fact that for the first 60% of the cartilage thickness the patterns were very similar, they differed in the deep cartilage. A strong increase was seen in the intensity of the peak 1374 cm^{-1} while the intensity of the peak 1062 cm^{-1} slightly decreased. Non-sulphated glycoproteins could explain the difference. It has been suggested that the deep cartilage contains significant amounts of non-sulphated glycoproteins, which would explain the increase in CH_3 vibration in the deep tissue [159]. The CH_3 vibration does not differentiate sulphated PGs and non-sulphated glycoproteins from each other, whereas SO_3- stretching vibration is able to distinguish between these forms. Therefore, it can be speculated that the peak at 1374cm^{-1} is related to all GAGs and glycoproteins while the peak at 1062 cm^{-1} is related directly to GAGs.

Second derivative spectroscopy offers one way to increase the specificity for collagen and PGs as compared to earlier univariate methods. However, this requires a high signal-to-noise ratio as the noise is amplified in the differentiation of the spectra. The problem becomes more significant if the analyzed peak is weak, as is the case with the sulphate peak at 1062 cm^{-1}. The signal-to-noise ratio can

be increased by collecting more scans per pixel, increasing the pixel size or reducing the spectral resolution in use. However, some of the smaller peaks will be lost if the spectral resolution is reduced. Further, sometimes a small pixel size is needed if highly spatial compositional changes are studied. The number of scans per pixel can be increased, but since this increases the measurement time, there has to be a compromise between the quality of spectra and measurement time. When the large sample pools are considered, the total measurement time might become excessive if one attempts to analyse weak peaks with a small pixel size.

7.5 MULTIVARIATE REGRESSION

Multivariate models (PCR and PLSR) were found as the most useful methods for exploiting the spectroscopic data efficiently (III, IV). In the simplest multivariate method, MLR, a major problem encountered is the multicollinearity of measured variables. This is avoided by transferring the original variables into new orthogonal variables. This enables the building of multivariate models without the multicollinearity problem. On average, a fairly low number of new variables (4-7) were needed to predict the biochemical composition or biomechanical function of AC. It can be assumed that the multivariate models utilize the spectroscopic data efficiently and that the IR absorption spectra of AC contain very detailed information on AC composition. Therefore, the composition can be predicted with high specificity using IR spectra. However, in the case of biomechanical properties, a good result based on the average spectrum of the sample is not necessarily expected. The biomechanical properties of AC are determined by its heterogeneous composition and structure. The equilibrium modulus of AC was predicted with good accuracy by a relatively simple model in the sample set used (3 components). It was more difficult to predict the dynamic modulus from IR spectra. Some of the samples were shown to be significantly stiffer than the others. Ideally, a multivariate model should be trained by using a sample set with evenly distributed reference values. This was not

the case with the dynamic modulus, and the model was not feasible when all samples were used. This is because of the uneven distribution of dynamic modulus values, and furthermore, the dynamic moduli of the stiffest samples are most likely strongly influenced by the collagen network architecture. IR spectra mainly contain information on the biochemical composition of the sample. Although some structural information can be achieved by IR spectroscopy, the present spectra were average spectra of the samples. Therefore, information on the collagen network structure is not included in the spectra. When the stiffest samples were omitted, the accuracy of the model was significantly improved.

In study IV, a genetic algorithm was used to further improve the PLSR models for prediction of the biomechanical properties of AC. The genetic algorithm is an iterative algorithm that preselects the most useful variables from the measured spectra to build the most optimal multivariate model for the problem. The variables were averaged (window size of 5 variables) before the genetic algorithm in order to avoid overfitting. Nonetheless, a relatively low number of original variables was needed to predict the equilibrium (16 variables) and dynamic modulus (17 variables) to achieve the best model. One reason for this might be the low number of samples.

Equilibrium modulus is known to be strongly linked to PG content, while the collagen network plays an important role in dynamic loading. This might explain why the models did not use identical spectral regions. Water is also an important factor in the biomechanical behaviour of AC. Information on water was not included in the IR spectra of AC, as the measurements were conducted using dry sections. However, the amount of water in AC is dependent on the PG content, since PGs attract water in the tissue, and the collagen content, as the collagen network limits the volume that water can take. Therefore, the volume of water can be approximated if the amounts PG and collagen contents are known. This might explain the obtained good correlations with the biomechanical reference measurements, even though the IR spectra contained direct information only on the dry matrix content of AC.

In the future, the multivariate models for the prediction of biomechanical properties could be enhanced in two ways. First, the spatial variation in composition could be taken into account. For example, this could be achieved by utilizing ten evenly divided layers of the cartilage to produce ten different spectra for each sample and to use so-called multi-block multivariate methods [160]. The second important thing is to include the collagen network architecture into the model. This could be measured by using polarized light microscopy [52] or polarized IR spectroscopy [20, 24, 31, 32, 138] and taken into account by including the depth-wise collagen orientation in the same multi-block model as the IR spectra. In principle, this kind of model should explain very accurately the specific biomechanical parameters of AC.

7.6 COMPARISON OF ANALYSIS METHODS

The results of this thesis show that while the univariate methods are straightforward in use, they do not exploit the full potential present in IR spectroscopic imaging. Univariate methods typically represent the first step when new data sets are investigated. They are suitable for visualizing the data and in some cases they do produce sufficiently specific information about the composition of the samples. However, this is usually not the case with biological samples, because absorption peaks overlap with each other and it is difficult to identify specific peaks.

Curve fitting or second derivative spectroscopy can be used to increase the separation between overlapping absorption peaks. The results show that this is beneficial when the composition of AC is analyzed. However, they do not resolve the problem completely. In addition, both methods have their own inherent weaknesses especially when imaging studies are concerned, as curve fitting is computationally demanding, and second derivative spectroscopy requires a high signal-to-noise ratio. The use of pure compound spectra is intended to decompose the measured spectrum into component spectra instead of resolving all separate absorption peaks.

In principle, this approach is feasible, but the results of this study were not as good as expected. This is most likely related to the difficulty of obtaining a proper pure compound spectra from which to model the compounds in the tissue.

Multivariate regression methods offer a means to achieve very good results in both prediction of composition and functional properties of AC. PCR and PLSR form new uncorrelated variables that can be used to predict the desired properties of the samples. By adopting this approach, the problems associated with multicollinearity can be avoided. Very good results can be obtained in many cases by directly using either full spectra or certain spectral regions. Variable selection techniques, such as genetic algorithm, can be utilized to select the most relevant spectral regions and to improve the results even further. Multivariate regression methods require the collection of a representative sample pool and the calibration of the IR spectra using the reference information obtained from other measurements. Computational techniques are then used to find the corresponding information from the IR spectra. Accurate qualitative and quantitative analyses can be conducted in this way from IR spectra.

8 Conclusions

IR microspectroscopy represents a powerful tool for biochemical characterization of AC at microscopic level. However, the results are strongly dependent on the spectral analysis methods. In order to utilize the spectroscopic data effectively, one must be aware of the potential and limitations of these analysis methods.

The most important findings of this thesis may be summarized as follows:

- The specificity of the previously used univariate parameters for collagen and PGs is limited.
- The separation of overlapping peaks can be increased by curve fitting or second derivative spectroscopy. This also improves the biochemical specificity of the peaks.
- Multivariate regression models provided the most effective analysis of biochemical composition of AC.
- Multivariate regression models can be further enhanced with the variable selection algorithms, such as the genetic algorithm.
- Prediction with high accuracy of the specific biomechanical parameters of AC is possible solely from IR spectra by applying multivariate regression models.

Bibliography

[1] V. Mow, A. Ratcliffe, and A. R. Poole, "Cartilage and diarthrodial joints as paradigms for hierarchical materials and structures.," *Biomaterials* **13,** 67–97 (1992).

[2] J. Martel-Pelletier, C. Boileau, J.-P. Pelletier, and P. J. Roughley, "Cartilage in normal and osteoarthritis conditions.," *Best Pract Res Clin Rheumatol* **22,** 351–384 (2008).

[3] A. I. Maroudas, "Balance between swelling pressure and collagen tension in normal and degenerate cartilage.," *Nature* **260,** 808–809 (1976).

[4] J. Rieppo, J. Töyräs, M. T. Nieminen, V. Kovanen, M. M. Hyttinen, R. K. Korhonen, J. S. Jurvelin, and H. J. Helminen, "Structure-function relationships in enzymatically modified articular cartilage.," *Cells Tissues Organs* **175,** 121–132 (2003).

[5] M. Venn and A. Maroudas, "Chemical composition and swelling of normal and osteoarthrotic femoral head cartilage. I. Chemical composition.," *Ann Rheum Dis* **36,** 121–129 (1977).

[6] Y. Xia, J. B. Moody, H. Alhadlaq, N. Burton-Wurster, and G. Lust, "Characteristics of topographical heterogeneity of articular cartilage over the joint surface of a humeral head.," *Osteoarthritis Cartilage* **10,** 370–380 (2002).

[7] C. G. Armstrong and V. C. Mow, "Variations in the intrinsic mechanical properties of human articular cartilage with age, degeneration, and water content.," *J Bone Joint Surg Am* **64,** 88–94 (1982).

[8] J. A. Buckwalter and J. A. Martin, "Osteoarthritis.," *Adv Drug Deliv Rev* **58,** 150–167 (2006).

[9] J. A. Buckwalter, H. J. Mankin, and A. J. Grodzinsky, "Articular cartilage and osteoarthritis.," *Instr Course Lect* **54,** 465–480 (2005).

[10] R. D. Altman, "Early management of osteoarthritis.," *Am J Manag Care* **16 Suppl Management,** S41–S47 (2010).

[11] C. R. Chu, A. A. Williams, C. H. Coyle, and M. E. Bowers, "Early diagnosis to enable early treatment of pre-osteoarthritis.," *Arthritis Res Ther* **14,** 212 (2012).

[12] H. J. Mankin and A. Z. Thrasher, "Water content and binding in normal and osteoarthritic human cartilage.," *J Bone Joint Surg Am* **57,** 76–80 (1975).

[13] V. Mow, D. Fithian, and M. Kelly, Chap Fundamentals of articular cartilage and meniscus biomechanics in *Articular cartilage and knee joint function: basic science and arthroscopy* (Raven Press Ltd., 1990).

[14] J. A. Buckwalter and J. Martin, "Degenerative joint disease.," *Clin Symp* **47,** 1–32 (1995).

[15] S. Saarakkala, M. S. Laasanen, J. S. Jurvelin, K. Törrönen, M. J. Lammi, R. Lappalainen, and J. Töyräs, "Ultrasound indentation of normal and spontaneously degenerated bovine articular cartilage.," *Osteoarthritis Cartilage* **11,** 697–705 (2003).

[16] P. Kiviranta, E. Lammentausta, J. Töyräs, I. Kiviranta, and J. S. Jurvelin, "Indentation diagnostics of cartilage degeneration.," *Osteoarthritis Cartilage* **16,** 796–804 (2008).

[17] I. W. Levin and R. Bhargava, "Fourier transform infrared vibrational spectroscopic imaging: integrating microscopy and molecular recognition.," *Annu Rev Phys Chem* **56,** 429–474 (2005).

[18] P. Lasch and D. Naumann, "Spatial resolution in infrared microspectroscopic imaging of tissues.," *Biochim Biophys Acta* **1758,** 814–829 (2006).

[19] A. Boskey and N. P. Camacho, "FT-IR imaging of native and tissue-engineered bone and cartilage," *Biomaterials* **28,** 2465–2478 (2007).

[20] N. P. Camacho, P. West, P. A. Torzilli, and R. Mendelsohn, "FTIR microscopic imaging of collagen and proteoglycan in bovine cartilage," *Biopolymers* **62,** 1–8 (2001).

[21] K. Potter, L. H. Kidder, I. W. Levin, E. N. Lewis, and R. G. Spencer, "Imaging of collagen and proteoglycan in cartilage sections using Fourier transform infrared spectral imaging," *Arthritis Rheum* **44,** 846–855 (2001).

[22] N. Camacho, P. West, M. Griffith, R. Warren, and C. Hidaka, "FT-IR imaging spectroscopy of genetically modified bovine chondrocytes," *Materials Science and Engineering* **17,** 3–9 (2001).

[23] P. Julkunen, P. Kiviranta, W. Wilson, J. S. Jurvelin, and R. K. Korhonen, "Characterization of articular cartilage by combining microscopic analysis with a fibril-reinforced finite-element model.," *J Biomech* **40,** 1862–1870 (2007).

[24] X. Bi, G. Li, S. B. Doty, and N. P. Camacho, "A novel method for determination of collagen orientation in cartilage by Fourier transform infrared imaging spectroscopy (FT-IRIS)," *Osteoarthritis Cartilage* **13,** 1050–1058 (2005).

[25] X. Bi, X. Yang, M. P. G. Bostrom, D. Bartusik, S. Ramaswamy, K. W. Fishbein, R. G. Spencer, and N. P. Camacho, "Fourier transform infrared imaging and MR microscopy studies detect compositional and structural changes in cartilage in a rabbit model of osteoarthritis," *Anal Bioanal Chem* **387,** 1601–1612 (2007).

[26] X. Bi, X. Yang, M. P. G. Bostrom, and N. P. Camacho, "Fourier transform infrared imaging spectroscopy investigations in the pathogenesis and repair of cartilage.," *Biochim Biophys Acta* **1758,** 934–941 (2006).

[27] E. David-Vaudey, A. Burghardt, K. Keshari, A. Brouchet, M. Ries, and S. Majumdar, "Fourier Transform Infrared Imaging of focal lesions in human osteoarthritic cartilage," *Eur Cell Mater* **10,** 51–60; discussion 60 (2005).

[28] E. Jiang and J. Rieppo, "Enhancing FTIR imaging capabilities with two-dimensional correlation spectroscopy (2DCOS): A study of concentration gradients of collagen and proteoglycans in human patellar cartilage," *Journal of Molecular Structure* **799,** 196–203 (2006).

[29] M. Kim, X. Bi, W. E. Horton, R. G. Spencer, and N. P. Camacho, "Fourier transform infrared imaging spectroscopic analysis of tissue engineered cartilage: histologic and biochemical correlations.," *J Biomed Opt* **10,** 031105 (2005).

[30] J. Rieppo, M. M. Hyttinen, E. Halmesmäki, H. Ruotsalainen, A. Vasara, I. Kiviranta, J. S. Jurvelin, and H. J. Helminen, "Changes in spatial collagen content and collagen network architecture in porcine articular cartilage during growth and maturation.," *Osteoarthritis Cartilage* (2008).

[31] Y. Xia, H. Alhadlaq, N. Ramakrishnan, A. Bidthanapally, F. Badar, and M. Lu, "Molecular and morphological adaptations in compressed articular cartilage by polarized light microscopy and Fourier-transform infrared imaging.," *J Struct Biol* **164,** 88–95 (2008).

[32] Y. Xia, N. Ramakrishnan, and A. Bidthanapally, "The depth-dependent anisotropy of articular cartilage by Fourier-transform infrared imaging (FTIRI).," *Osteoarthritis Cartilage* **15,** 780–788 (2007).

[33] J. Yin and Y. Xia, "Macromolecular concentrations in bovine nasal cartilage by Fourier transform infrared imaging and principal component regression.," *Appl Spectrosc* **64,** 1199–1208 (2010).

[34] J. A. Buckwalter and H. J. Mankin, "Articular cartilage: tissue design and chondrocyte-matrix interactions.," *Instr Course Lect* **47,** 477–486 (1998).

[35] M. F. Venn, "Variation of chemical composition with age in human femoral head cartilage.," *Ann Rheum Dis* **37,** 168–174 (1978).

[36] P. D. Byers, A. Maroudase, F. Oztop, R. A. Stockwell, and M. F. Venn, "Histological and biochemical studies on cartilage from osteoarthrotic femoral heads with special reference to surface characteristics.," *Connect Tissue Res* **5,** 41–49 (1977).

[37] R. Mayne, "Cartilage collagens. What is their function, and are they involved in articular disease?," *Arthritis Rheum* **32,** 241–246 (1989).

[38] D. Eyre, "Collagen of articular cartilage.," *Arthritis Res* **4,** 30–35 (2002).

[39] V. C. Hascall and S. W. Sajdera, "Physical properties and polydispersity of proteoglycan from bovine nasal cartilage.," *J Biol Chem* **245,** 4920–4930 (1970).

[40] H. Muir, "Proteoglycans as organizers of the intercellular matrix.," *Biochem Soc Trans* **11,** 613–622 (1983).

[41] J. Dudhia, "Aggrecan, aging and assembly in articular cartilage.," *Cell Mol Life Sci* **62,** 2241–2256 (2005).

[42] A. R. Poole, T. Kojima, T. Yasuda, F. Mwale, M. Kobayashi, and S. Laverty, "Composition and structure of articular cartilage: a template for tissue repair.," *Clin Orthop Relat Res* S26–S33 (2001).

[43] T. E. Hardingham and A. J. Fosang, "Proteoglycans: many forms and many functions.," *FASEB J* **6,** 861–870 (1992).

[44] E. B. Hunziker, T. M. Quinn, and H.-J. Häuselmann, "Quantitative structural organization of normal adult human articular cartilage.," *Osteoarthritis Cartilage* **10,** 564–572 (2002).

[45] Z. Lin, C. Willers, J. Xu, and M.-H. Zheng, "The chondrocyte: biology and clinical application.," *Tissue Eng* **12,** 1971–1984 (2006).

[46] K. D. Jadin, B. L. Wong, W. C. Bae, K. W. Li, A. K. Williamson, B. L. Schumacher, J. H. Price, and R. L. Sah, "Depth-varying density and organization of chondrocytes in immature and mature bovine articular cartilage assessed by 3d imaging and analysis.," *J Histochem Cytochem* **53,** 1109–1119 (2005).

[47] J. P. Arokoski, M. M. Hyttinen, T. Lapveteläinen, P. Takács, B. Kosztáczky, L. Módis, V. Kovanen, and H. Helminen, "Decreased birefringence of the superficial zone collagen network in the canine knee (stifle) articular cartilage after long distance running training, detected by quantitative polarised light microscopy.," *Ann Rheum Dis* **55,** 253–264 (1996).

[48] H. Muir, P. Bullough, and A. Maroudas, "The distribution of collagen in human articular cartilage with some of its physiological implications.," *J Bone Joint Surg Br* **52,** 554–563 (1970).

[49] I. Kiviranta, J. Jurvelin, M. Tammi, A. M. Säämänen, and H. J. Helminen, "Microspectrophotometric quantitation of glycosaminoglycans in articular cartilage sections stained with Safranin O.," *Histochemistry* **82,** 249–255 (1985).

[50] C. Ficat and A. Maroudas, "Cartilage of the patella. Topographical variation of glycosaminoglycan content in normal and fibrillated tissue.," *Ann Rheum Dis* **34,** 515–519 (1975).

[51] A. Maroudas, M. T. Bayliss, and M. F. Venn, "Further studies on the composition of human femoral head cartilage.," *Ann Rheum Dis* **39,** 514–523 (1980).

[52] J. Rieppo, J. Hallikainen, J. S. Jurvelin, I. Kiviranta, H. J. Helminen, and M. M. Hyttinen, "Practical considerations in the use of polarized light microscopy in the analysis of the collagen network in articular cartilage.," *Microsc Res Tech* **71,** 279–287 (2008).

[53] V. C. Mow, S. C. Kuei, W. M. Lai, and C. G. Armstrong, "Biphasic creep and stress relaxation of articular cartilage in compression? Theory and experiments.," *J Biomech Eng* **102,** 73–84 (1980).

[54] V. C. Mow, M. H. Holmes, and W. M. Lai, "Fluid transport and mechanical properties of articular cartilage: a review.," *J Biomech* **17,** 377–394 (1984).

[55] A. Maroudas and C. Bannon, "Measurement of swelling pressure in cartilage and comparison with the osmotic pressure of constituent proteoglycans.," *Biorheology* **18,** 619–632 (1981).

[56] M. B. Schmidt, V. C. Mow, L. E. Chun, and D. R. Eyre, "Effects of proteoglycan extraction on the tensile behavior of articular cartilage.," *J Orthop Res* **8,** 353–363 (1990).

[57] G. E. Kempson, H. Muir, C. Pollard, and M. Tuke, "The tensile properties of the cartilage of human femoral condyles related to the content of collagen and glycosaminoglycans.," *Biochim Biophys Acta* **297,** 456–472 (1973).

[58] W. C. Bae, V. W. Wong, J. Hwang, J. M. Antonacci, G. E. Nugent-Derfus, M. E. Blewis, M. M. Temple-Wong, and R. L. Sah, "Wear-lines and split-lines of human patellar cartilage: relation to tensile biomechanical properties.," *Osteoarthritis Cartilage* **16,** 841–845 (2008).

[59] J. K. Suh, Z. Li, and S. L. Woo, "Dynamic behavior of a biphasic cartilage model under cyclic compressive loading.," *J Biomech* **28,** 357–364 (1995).

[60] J. Töyräs, M. S. Laasanen, S. Saarakkala, M. J. Lammi, J. Rieppo, J. Kurkijärvi, R. Lappalainen, and J. S. Jurvelin, "Speed of sound in normal and degenerated bovine articular cartilage.," *Ultrasound Med Biol* **29,** 447–454 (2003).

[61] J. E. Kurkijärvi, M. J. Nissi, I. Kiviranta, J. S. Jurvelin, and M. T. Nieminen, "Delayed gadolinium-enhanced MRI of cartilage (dGEMRIC) and T2 characteristics of human knee articular cartilage: topographical variation and relationships to mechanical properties.," *Magn Reson Med* **52,** 41–46 (2004).

[62] W. C. Hayes, L. M. Keer, G. Herrmann, and L. F. Mockros, "A mathematical analysis for indentation tests of articular cartilage.," *J Biomech* **5,** 541–551 (1972).

[63] E. M. Hasler, W. Herzog, J. Z. Wu, W. Mller, and U. Wyss, "Articular cartilage biomechanics: theoretical models, material properties, and biosynthetic response.," *Crit Rev Biomed Eng* **27,** 415–488 (1999).

[64] W. Wilson, C. C. van Donkelaar, R. van Rietbergen, and R. Huiskes, "The role of computational models in the search for the mechanical behavior and damage mechanisms of articular cartilage.," *Med Eng Phys* **27,** 810–826 (2005).

[65] W. Wilson, J. M. Huyghe, and C. C. van Donkelaar, "A composition-based cartilage model for the assessment of compositional changes during cartilage damage and adaptation.," *Osteoarthritis Cartilage* **14,** 554–560 (2006).

[66] P. Julkunen, J. S. Jurvelin, and H. Isaksson, "Contribution of tissue composition and structure to mechanical response of articular cartilage under different loading geometries and strain rates.," *Biomech Model Mechanobiol* **9,** 237–245 (2010).

[67] P. Julkunen, W. Wilson, J. S. Jurvelin, J. Rieppo, C.-J. Qu, M. J. Lammi, and R. K. Korhonen, "Stress-relaxation of human patellar articular cartilage in unconfined compression: prediction of mechanical response by tissue composition and structure.," *J Biomech* **41,** 1978–1986 (2008).

[68] R. K. Korhonen, M. S. Laasanen, J. Töyräs, R. Lappalainen, H. J. Helminen, and J. S. Jurvelin, "Fibril reinforced poroelastic model predicts specifically mechanical behavior of normal, proteoglycan depleted and collagen degraded articular cartilage.," *J Biomech* **36,** 1373–1379 (2003).

[69] D. D. Sun, X. E. Guo, M. Likhitpanichkul, W. M. Lai, and V. C. Mow, "The influence of the fixed negative charges on mechanical and electrical behaviors of articular cartilage under unconfined compression.," *J Biomech Eng* **126,** 6–16 (2004).

[70] W. M. Lai, J. S. Hou, and V. C. Mow, "A triphasic theory for the swelling and deformation behaviors of articular cartilage.," *J Biomech Eng* **113,** 245–258 (1991).

[71] W. Wilson, C. C. van Donkelaar, B. van Rietbergen, and R. Huiskes, "A fibril-reinforced poroviscoelastic swelling model for articular cartilage.," *J Biomech* **38,** 1195–1204 (2005).

[72] J. Arokoski, A. Malmivaara, M. Manninen, E. Moilanen, R. Ojala, P. Paavolainen, J. Ruuskanen, P. Virolainen, H. Virtapohja, K. Vuolteenaho, and H. Osterman, "Polvi- ja lonkkanivelrikon hoito," *Duodecim* **123,** 1759–1760 (2007).

[73] J. P. Arokoski, J. S. Jurvelin, U. Väätäinen, and H. J. Helminen, "Normal and pathological adaptations of articular cartilage to joint loading.," *Scand J Med Sci Sports* **10,** 186–198 (2000).

[74] E. L. Radin, I. L. Paul, and M. J. Tolkoff, "Subchondral bone changes in patients with early degenerative joint disease.," *Arthritis Rheum* **13,** 400–405 (1970).

[75] D. T. Felson and T. Neogi, "Osteoarthritis: is it a disease of cartilage or of bone?," *Arthritis Rheum* **50,** 341–344 (2004).

[76] A. P. Newman, "Articular cartilage repair.," *Am J Sports Med* **26,** 309–324 (1998).

[77] M. B. Mueller and R. S. Tuan, "Anabolic/Catabolic balance in pathogenesis of osteoarthritis: identifying molecular targets.," *PM R* **3,** S3–11 (2011).

[78] M. Brittberg, A. Lindahl, A. Nilsson, C. Ohlsson, O. Isaksson, and L. Peterson, "Treatment of deep cartilage defects in the knee with autologous chondrocyte transplantation.," *N Engl J Med* **331,** 889–895 (1994).

[79] L. Hangody and P. Fles, "Autologous osteochondral mosaicplasty for the treatment of full-thickness defects of weight-bearing joints: ten years of experimental and clinical experience.," *J Bone Joint Surg Am* **85-A Suppl 2,** 25–32 (2003).

[80] J. Lee, E. Gazi, J. Dwyer, M. D. Brown, N. W. Clarke, J. M. Nicholson, and P. Gardner, "Optical artefacts in transflection mode FTIR microspectroscopic images of single cells on a biological support: the effect of back-scattering into collection optics.," *Analyst* **132,** 750–755 (2007).

[81] P. Bassan, A. Sachdeva, A. Kohler, C. Hughes, A. Henderson, J. Boyle, J. H. Shanks, M. Brown, N. W. Clarke, and P. Gardner, "FTIR microscopy of biological cells and tissue: data analysis using resonant Mie scattering (RMieS) EMSC algorithm.," *Analyst* **137,** 1370–1377 (2012).

[82] P. Bassan, H. J. Byrne, F. Bonnier, J. Lee, P. Dumas, and P. Gardner, "Resonant Mie scattering in infrared spectroscopy of biological materials–understanding the 'dispersion artefact'," *Analyst* **134,** 1586–1593 (2009).

[83] P. Bassan, A. Kohler, H. Martens, J. Lee, H. J. Byrne, P. Dumas, E. Gazi, M. Brown, N. Clarke, and P. Gardner, "Resonant Mie scattering (RMieS) correction of infrared spectra from highly scattering biological samples," *Analyst* **135,** 268–277 (2010).

[84] A. Kohler, J. Sul-Suso, G. D. Sockalingum, M. Tobin, F. Bahrami, Y. Yang, J. Pijanka, P. Dumas, M. Cotte, D. G. van Pittius, G. Parkes, and H. Martens, "Estimating and correcting mie scattering in synchrotron-based microscopic fourier transform infrared spectra by extended multiplicative signal correction.," *Appl Spectrosc* **62,** 259–266 (2008).

[85] B. Mohlenhoff, M. Romeo, M. Diem, and B. R. Wood, "Mie-type scattering and non-Beer-Lambert absorption behavior of human cells in infrared microspectroscopy.," *Biophys J* **88,** 3635–3640 (2005).

[86] P. Walstra, "Approximation formulae for the light scattering coefficient of dielectric spheres," *British Journal of Applied Physics* **15,** 1545–1552 (1964).

[87] A. Christy, Y. Ozaki, and V. Gregoriou, *Modern Fourier Transform Infrared Spectroscopy* (Wiley, 2002).

[88] B. Stuart, *Infrared Spectroscopy: Fundamentals and Applications* (Wiley, 2004).

[89] R. Bhargava and I. W. Levin, "Fourier transform infrared imaging: theory and practice.," *Anal Chem* **73,** 5157–5167 (2001).

[90] I. Salomaa and J. Kauppinen, "Origin of and Compensation for the Baseline Errors in Fourier Transform Spectra," *Applied Spectroscopy* **52,** 579–586(8) (1 April 1998).

[91] A. Kohler, D. Bertrand, H. Martens, K. Hannesson, C. Kirschner, and R. Ofstad, "Multivariate image analysis of a set of FTIR microspectroscopy images of aged bovine muscle tissue combining image and design information," *Anal Bioanal Chem* **389,** 1143–1153 (2007).

[92] A. Kohler, C. Kirschner, A. Oust, and H. Martens, "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.," *Appl Spectrosc* **59,** 707–716 (2005).

[93] N. Afseth and A. Kohler, "Extended multiplicative signal correction in vibrational spectroscopy, a tutorial," *Chemometrics and Intelligent Laboratory Systems* (2012).

[94] A. Kohler, U. Böcker, J. Warringer, A. Blomberg, S. W. Omholt, E. Stark, and H. Martens, "Reducing inter-replicate variation in fourier transform infrared spectroscopy by extended multiplicative signal correction.," *Appl Spectrosc* **63,** 296–305 (2009).

[95] K. R. Bambery, B. R. Wood, and D. McNaughton, "Resonant Mie scattering (RMieS) correction applied to FTIR images of biological tissue samples.," *Analyst* **137,** 126–132 (2012).

[96] P. Bassan, A. Kohler, H. Martens, J. Lee, E. Jackson, N. Lockyer, P. Dumas, M. Brown, N. Clarke, and P. Gardner, "RMieS-EMSC correction for infrared spectra of biological cells: extension using full Mie theory and GPU computing.," *J Biophotonics* **3,** 609–620 (2010).

[97] A. Kohler, D. Bertrand, H. Martens, K. Hannesson, C. Kirschner, and R. Ofstad, "Multivariate image analysis of a set of FTIR microspectroscopy images of aged bovine muscle tissue combining image and design information," *Analytical and Bioanalytical Chemistry* **389,** 1143–1153(11) (October 2007).

[98] D. M. Byler and H. Susi, "Examination of the secondary structure of proteins by deconvolved FTIR spectra.," *Biopolymers* **25,** 469–487 (1986).

[99] J. L. Arrondo and F. M. Goi, "Structure and dynamics of membrane proteins as studied by infrared spectroscopy.," *Prog Biophys Mol Biol* **72,** 367–405 (1999).

[100] M. C. Manning, "Use of infrared spectroscopy to monitor protein structure and stability.," *Expert Rev Proteomics* **2,** 731–743 (2005).

[101] C. Petibois and G. Déléris, "Chemical mapping of tumor progression by FT-IR imaging: towards molecular histopathology.," *Trends Biotechnol* **24,** 455–462 (2006).

[102] C. Petibois, B. Drogat, A. Bikfalvi, G. Déléris, and M. Moenner, "Histological mapping of biochemical changes in solid tumors by FT-IR spectral imaging.," *FEBS Lett* **581,** 5469–5474 (2007).

[103] C. Petibois, G. Gouspillou, K. Wehbe, J.-P. Delage, and G. Déléris, "Analysis of type I and IV collagens by FT-IR spectroscopy and imaging for a molecular investigation of skeletal muscle connective tissue.," *Anal Bioanal Chem* **386,** 1961–1966 (2006).

[104] K. Belbachir, R. Noreen, G. Gouspillou, and C. Petibois, "Collagen types analysis and differentiation by FTIR spectroscopy.," *Anal Bioanal Chem* **395,** 829–837 (2009).

[105] E. P. Paschalis, D. B. Burr, R. Mendelsohn, J. M. Hock, and A. L. Boskey, "Bone mineral and collagen quality in humeri of ovariectomized cynomolgus monkeys given rhPTH(1-34) for 18 months," *J Bone Miner Res* **18,** 769–775 (2003).

[106] E. P. Paschalis, R. Recker, E. DiCarlo, S. B. Doty, E. Atti, and A. L. Boskey, "Distribution of collagen cross-links in normal human trabecular bone.," *J Bone Miner Res* **18,** 1942–1946 (2003).

[107] E. P. Paschalis, E. Shane, G. Lyritis, G. Skarantavos, R. Mendelsohn, and A. L. Boskey, "Bone fragility and collagen cross-links.," *J Bone Miner Res* **19,** 2000–2004 (2004).

[108] E. P. Paschalis, K. Verdelis, S. B. Doty, A. L. Boskey, R. Mendelsohn, and M. Yamauchi, "Spectroscopic characterization of collagen cross-links in bone.," *J Bone Miner Res* **16,** 1821–1828 (2001).

[109] M. J. Turunen, S. Saarakkala, H. J. Helminen, J. S. Jurvelin, and H. Isaksson, "Age-related changes in organization and content of the collagen matrix in rabbit cortical bone.," *J Orthop Res* **30,** 435–442 (2012).

[110] M. J. Turunen, S. Saarakkala, L. Rieppo, H. J. Helminen, J. S. Jurvelin, and H. Isaksson, "Comparison between infrared and Raman spectroscopic analysis of maturing rabbit cortical bone.," *Appl Spectrosc* **65,** 595–603 (2011).

[111] R. Cheheltani, J. M. Rosano, B. Wang, A. K. Sabri, N. Pleshko, and M. F. Kiani, "Fourier transform infrared spectroscopic imaging of cardiac tissue to detect collagen deposition after myocardial infarction.," *J Biomed Opt* **17,** 056014 (2012).

[112] A. C. Leskovjan, A. Kretlow, and L. M. Miller, "Fourier transform infrared imaging showing reduced unsaturated lipid content in the hippocampus of a mouse model of Alzheimer's disease.," *Anal Chem* **82,** 2711–2716 (2010).

[113] A. M. Croxford, D. Crombie, D. McNaughton, R. Holmdahl, K. S. Nandakumar, and M. J. Rowley, "Specific antibody protection of the extracellular cartilage matrix against collagen antibody-induced damage.," *Arthritis Rheum* **62,** 3374–3384 (2010).

[114] A. M. Croxford, K. S. Nandakumar, R. Holmdahl, M. J. Tobin, D. McNaughton, and M. J. Rowley, "Chemical changes demonstrated in cartilage by synchrotron infrared microspectroscopy in an antibody-induced murine model of rheumatoid arthritis.," *J Biomed Opt* **16,** 066004 (2011).

[115] H. Mark and J. Workman, "Derivatives in Spectroscopy, Part II - The True Derivative," *Spectroscopy* **18,** 25–28 (2003).

[116] A. Savitzky and M. J. E. Golay, "Smoothing and Differentiation of Data by Simplified Least Squares Procedures.," *Analytical Chemistry* **36,** 1627–1639 (1964).

[117] J. Steinier, Y. Termonia, and J. Deltour, "Smoothing and differentiation of data by simplified least square procedure," *Analytical Chemistry* **44,** 1906–1909 (1972).

[118] H. Martens and T. Naes, *Multivariate Calibration* (John Wiley & Sons Ltd., Chichester, 1989).

[119] S. Wold, M. Sjöström, and L. Eriksson, "PLS-regression: a basic tool of chemometrics," *Chemometrics and Intelligent Laboratory Systems* **58,** 109–130 (2001).

[120] C. Andersen and R. Bro, "Variable selection in regression - a tutorial," *Journal* **24,** 728–737 (2010).

[121] R. Leardi, M. Seasholtz, and R. Pell, "Variable selection for multivariate calibration using a genetic algorithm: prediction of additive concentrations in polymer films from Fourier transform-infrared spectral data," *Analytica Chimica Acta* **461,** 189–200 (2002).

[122] A. Niazi and R. Leardi, "Genetic algorithms in chemometrics," *Journal of Chemometrics* **26,** 345–351 (2012).

[123] C. Krafft, G. Steiner, C. Beleites, and R. Salzer, "Disease recognition by infrared and Raman spectroscopy.," *J Biophotonics* **2,** 13–28 (2009).

[124] L. Mutihac and R. Mutihac, "Mining in chemometrics.," *Anal Chim Acta* **612,** 1–18 (2008).

[125] E. Peuchant, S. Richard-Harston, I. Bourdel-Marchasson, J.-F. Dartigues, L. Letenneur, P. Barberger-Gateau, S. Arnaud-Dabernat, and J.-Y. Daniel, "Infrared spectroscopy: a reagent-free method to distinguish Alzheimer's disease patients from normal-aging subjects.," *Transl Res* **152,** 103–112 (2008).

[126] P. Lasch, J. Schmitt, M. Beekes, T. Udelhoven, M. Eiden, H. Fabian, W. Petrich, and D. Naumann, "Antemortem identification of bovine spongiform encephalopathy from serum using infrared spectroscopy.," *Anal Chem* **75,** 6673–6678 (2003).

[127] H. Fabian, P. Lasch, and D. Naumann, "Analysis of biofluids in aqueous environment based on mid-infrared spectroscopy.," *J Biomed Opt* **10,** 031103 (2005).

[128] B. R. Wood, L. Chiriboga, H. Yee, M. A. Quinn, D. McNaughton, and M. Diem, "Fourier transform infrared (FTIR) spectral mapping of the cervical transformation zone, and dysplastic squamous epithelium.," *Gynecol Oncol* **93,** 59–68 (2004).

[129] P. Lasch, W. Haensch, D. Naumann, and M. Diem, "Imaging of colorectal adenocarcinoma using FT-IR microspectroscopy and cluster analysis.," *Biochim Biophys Acta* **1688,** 176–186 (2004).

[130] H. Fabian, P. Lasch, M. Boese, and W. Haensch, "Mid-IR microspectroscopic imaging of breast tumor tissue sections.," *Biopolymers* **67,** 354–357 (2002).

[131] D. Baykal, O. Irrechukwu, P.-C. Lin, K. Fritton, R. G. Spencer, and N. Pleshko, "Nondestructive assessment of engineered cartilage constructs using near-infrared spectroscopy.," *Appl Spectrosc* **64,** 1160–1166 (2010).

[132] P. A. West, P. A. Torzilli, C. Chen, P. Lin, and N. P. Camacho, "Fourier transform infrared imaging spectroscopy analysis of collagenase-induced cartilage degradation.," *J Biomed Opt* **10,** 14015 (2005).

[133] P. A. West, M. P. G. Bostrom, P. A. Torzilli, and N. P. Camacho, "Fourier transform infrared spectral analysis of degenerative cartilage: an infrared fiber optic probe and imaging study.," *Appl Spectrosc* **58,** 376–381 (2004).

[134] A. Hanifi, J. B. Richardson, J. H. Kuiper, S. Roberts, and N. Pleshko, "Clinical outcome of autologous chondrocyte implantation is correlated with infrared spectroscopic imaging-derived parameters.," *Osteoarthritis Cartilage* **20,** 988–996 (2012).

[135] S. Saarakkala and P. Julkunen, "Specificity of Fourier Transform Infrared (FTIR) Microspectroscopy to Estimate Depth-Wise Proteoglycan Content in Normal and Osteoarthritic Human Articular Cartilage," *Cartilage* **1,** 262–269 (2010).

[136] M. S. Laasanen, S. Saarakkala, J. Töyräs, J. Rieppo, and J. S. Jurvelin, "Site-specific ultrasound reflection properties and superficial collagen content of bovine knee articular cartilage.," *Phys Med Biol* **50,** 3221–3233 (2005).

[137] P. A. J. Brama, J. Holopainen, P. R. van Weeren, E. C. Firth, H. J. Helminen, and M. M. Hyttinen, "Influence of exercise and joint topography on depth-related spatial distribution of proteoglycan and collagen content in immature equine articular cartilage.," *Equine Vet J* **41,** 557–563 (2009).

[138] N. Ramakrishnan, Y. Xia, and A. Bidthanapally, "Polarized IR microscopic imaging of articular cartilage," *Phys Med Biol* **52,** 4601–4614 (2007).

[139] Y. Xia, D. Mittelstaedt, N. Ramakrishnan, M. Szarko, and A. Bidthanapally, "Depth-dependent anisotropies of amides and sugar in perpendicular and parallel sections of articular cartilage by Fourier transform infrared imaging.," *Microsc Res Tech* **74,** 122–132 (2011).

[140] A. M. Coats, D. W. L. Hukins, C. T. Imrie, and R. M. Aspden, "Polarization artefacts of an FTIR microscope and the consequences for intensity measurements on anisotropic materials," *J Microsc* **211,** 63–66 (2003).

[141] J. Yin, Y. Xia, and M. Lu, "Concentration profiles of collagen and proteoglycan in articular cartilage by Fourier transform infrared imaging and principal component regression.," *Spectrochim Acta A Mol Biomol Spectrosc* **88,** 90–96 (2012).

[142] G. Li, M. Thomson, E. Dicarlo, X. Yang, B. Nestor, M. P. G. Bostrom, and N. P. Camacho, "A chemometric analysis for evaluation of early-stage cartilage degradation by infrared fiber-optic probe spectroscopy.," *Appl Spectrosc* **59,** 1527–1533 (2005).

[143] M. Terajima, S. Damle, M. Penmatsa, P. West, M. Bostrom, C. Hidaka, M. Yamauchi, and N. Pleshko, "Temporal Changes in Collagen Cross-Links in Spontaneous Articular Cartilage Repair," *Cartilage* **3,** 278–287 (2012).

[144] M. Kim, J. J. Kraft, A. C. Volk, J. Pugarelli, N. Pleshko, and G. R. Dodge, "Characterization of a cartilage-like engineered biomass using a self-aggregating suspension culture model: molecular composition using FT-IRIS.," *J Orthop Res* **29,** 1881–1887 (2011).

[145] Y. Kobrina, L. Rieppo, S. Saarakkala, J. S. Jurvelin, and H. Isaksson, "Clustering of infrared spectra reveals histological zones in intact articular cartilage.," *Osteoarthritis Cartilage* **20,** 460–468 (2012).

[146] M. Jackson, L. P. Choo, P. H. Watson, W. C. Halliday, and H. H. Mantsch, "Beware of connective tissue proteins: assignment and implications of collagen absorptions in infrared spectra of human tissues.," *Biochim Biophys Acta* **1270,** 1–6 (1995).

[147] M. Jackson, M. G. Sowa, and H. H. Mantsch, "Infrared spectroscopy: a new frontier in medicine.," *Biophys Chem* **68,** 109–125 (1997).

[148] R. Servaty, J. Schiller, H. Binder, and K. Arnold, "Hydration of polymeric components of cartilage–an infrared spectroscopic study on hyaluronic acid and chondroitin sulfate.," *Int J Biol Macromol* **28,** 121–127 (2001).

[149] T. Yamagata, H. Saito, O. Habuchi, and S. Suzuki, "Purification and properties of bacterial chondroitinases and chondrosulfatases.," *J Biol Chem* **243,** 1523–1535 (1968).

[150] C. L. Borders and M. A. Raftery, "Purification and partial characterization of testicular hyaluronidase.," *J Biol Chem* **243,** 3756–3762 (1968).

[151] K. Király, M. M. Hyttinen, T. Lapveteläinen, M. Elo, I. Kiviranta, J. Dobai, L. Módis, H. J. Helminen, and J. P. Arokoski, "Specimen preparation and quantification of collagen birefringence in unstained sections of articular cartilage using image analysis and polarizing light microscopy.," *Histochem J* **29,** 317–327 (1997).

[152] H. Martens and P. Dardenne, "Validation and verification of regression in small data sets," *Chemometrics and Intelligent Laboratory Systems* **44,** 99–121 (1998).

[153] K. Király, T. Lapveteläinen, J. Arokoski, K. Törrönen, L. Módis, I. Kiviranta, and H. J. Helminen, "Application of selected cationic dyes for the semiquantitative estimation of glycosaminoglycans in histological sections of articular cartilage by microspectrophotometry.," *Histochem J* **28,** 577–590 (1996).

[154] N. Blumenkrantz and G. Asboe-Hansen, "New method for quantitative determination of uronic acids.," *Anal Biochem* **54,** 484–489 (1973).

[155] D. E. Schwartz, Y. Choi, L. J. Sandell, and W. R. Hanson, "Quantitative analysis of collagen, protein and DNA in fixed, paraffin-embedded and sectioned tissue.," *Histochem J* **17,** 655–663 (1985).

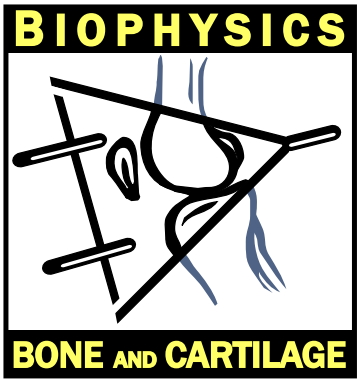
[156] J. Steiger, "Tests for Comparing Elements of a Correlation Matrix," *Psychological Bulletin* **87,** 245–251 (1980).

[157] R. Tauler, A. Smilde, and B. Kowalski, "Selectivity, local rank, three-way data analysis and ambiguity in multivariate curve resolution," *Journal of Chemometrics* **9,** 31–58 (1995).

[158] J. Lundbom, M. Kolehmainen, L. Pulkkinen, P. Soininen, M. Tiainen, U. Schwab, M. Uusitupa, and M. Tammi, "Mid-infrared spectroscopy and multivariate curve resolution for analyzing human adipose tissue triacylglycerols," *European Journal of Lipid Science and Technology* **112,** 1308–1314 (2010).

[159] H. Muir, Chap Biochemistry in *Adult articular cartilage* (Pittman Medical, 1979).

[160] S. Hassani, H. Martens, E. Qannari, M. Hanafi, G. Borge, and A. Kohler, "Analysis of -omics data: Graphical interpretation- and validation tools in multi-block methods," *Chemometrics and Intelligent Laboratory Systems* **104,** 140–153 (2010).



Lassi Rieppo

Infrared Spectroscopic Characterization of Articular Cartilage

Infrared spectroscopic imaging is a powerful tool for chemical analysis at the microscopic level. The technique has been used for the characterization of articular cartilage (AC). However, the development of data analysis methods has been slow. This thesis work aimed at developing novel infrared spectroscopic analysis techniques to help characterize AC. The novel methods included curve fitting, second derivative spectroscopy and multivariate regression models. The thesis work reveals that infrared spectra can provide detailed information on the composition and the biomechanical properties of AC when the spectroscopic data is exploited efficiently.

UNIVERSITY OF
EASTERN FINLAND

Publications of the University of Eastern Finland
Dissertations in Forestry and Natural Sciences

ISBN 978-952-61-0963-3