Cover picture:

Vibrio cholerae, magnification approximately 10.000 ×

Vakgroep Analytische Chemie
INW
Proeftuinstraat 86
9000 Gent

Vakgroep Biochemie, Fysiologie en Microbiologie
Laboratorium voor Microbiologie (Lm-Ugent)
K.L. Ledeganckstraat 35
9000 Gent

Raman spectroscopy as a tool for studying bacterial cell compounds

Joke De Gelder

Academic year 2007-2008

Dissertation submitted in fulfillment of the requirements for the degree of Doctor (Ph.D.) in Sciences, Chemistry

UNIVERSITEIT GENT

Promotor: Prof. Dr. Luc Moens
Co-promotor: Prof. Dr. Peter Vandenabeele
Co-promotor: Prof. Dr. Paul De Vos

Content

Abbreviations and acronyms

General

A	adenine
acetyl-CoA	acetyl coenzyme A
AFLP	amplified fragment length polymorphism
ATP	adenosine triphosphate
C	cytosine
CaDPA	calcium dipicolinate
CCD	charge coupled device
CoA	coenzyme A
Cys	cysteine
def	deformation
EMSC	extended multiplicative signal correction
FAME	fatty acid methyl ester
G	guanine
Gly	glycine
glyc	glycosidic ring deformation
HB	β-3-hydroxybutyrate
HPLC	high pressure liquid chromatography
PCA	principal component analysis
PCR	polymerase chain reaction
PHB	poly-3-hydroxybutyrate
Phe	phenylalanine
PLS	partial least squares
RFLP	restricted fragment length polymorphism
SDS-PAGE	sodium dodecyl sulfate - polyacrylamide gel electrophoresis
SNV	standard normal variate
str	stretching

T	thymine
Tryp	tryptophan
Tyr	tyrosine
U	uracil

Spectral descriptions

vw	very weak
w	weak
mw	medium weak
m	medium
ms	medium strong
s	strong
vs	very strong
br	broad
sh	shoulder
ν or str	stretching
δ or def	deformation
ρ	rocking
sym	symmetric

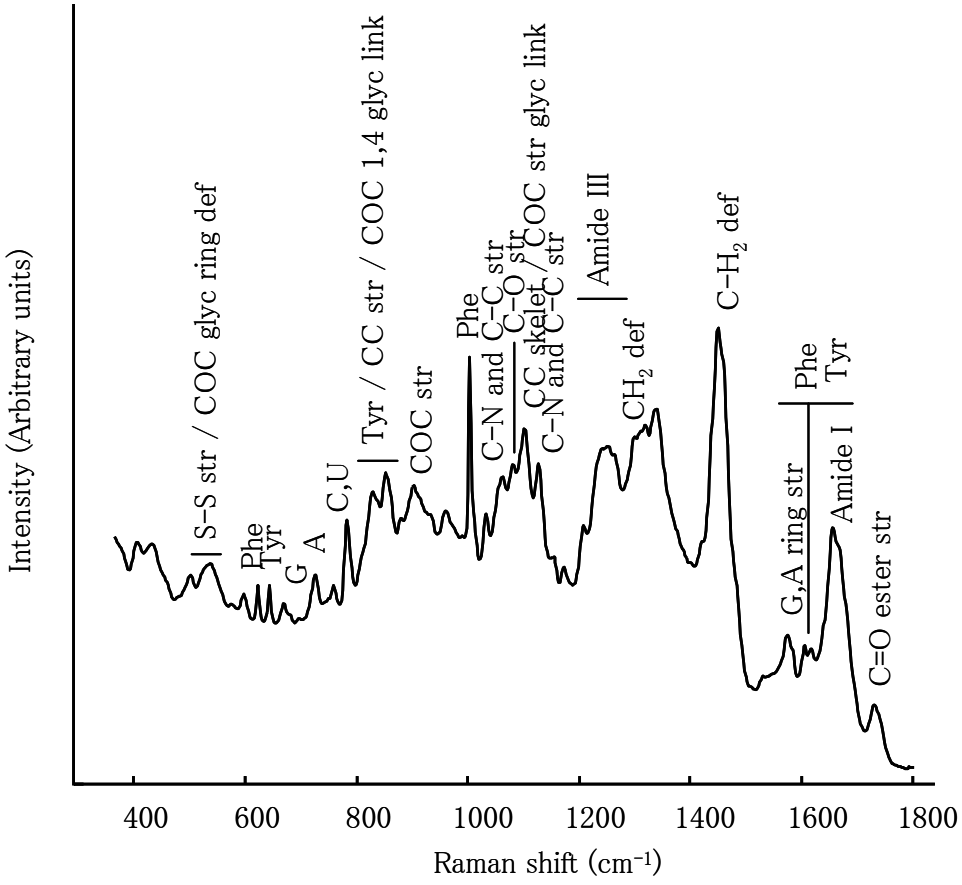
Chapter 1: Introduction and aim

Raman spectroscopy is a powerful technique as it is non-destructive, water interference is minimal and the required sample volumes are small. These features in combination with whole cell information as provided by Raman spectroscopy make this method of analysis attractive for use in microbiology. Indeed, a lot of current microbiological techniques of analysis focus on certain types of biomolecules (fatty acid methyl ester (FAME) analysis, sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) for proteins, etc.) and are labor-intensive. Identification of microorganisms on species[2,6] and even strain[1,4,7] level is possible. Bacterial Raman spectra can not only serve as a fingerprint pattern for identification, but they also contain important information about the cell's composition. However, because these spectra contain contributions of all Raman active molecules in the cell, they are complex. Bacterial Raman spectra usually have a typical shape (Fig. 1.1) because all cells contain the same basic building blocks: DNA, RNA, proteins, fatty acids, lipids, etc. Each band in a bacterial Raman spectrum is usually a superposition of contributions of several biomolecules. Therefore, in theory, we can not assign a single band to one biomolecule. However, in practice, some bands are dominated by the contribution of specific biomolecules and therefore some assignments to (groups of) biomolecules or specific vibrations have been made earlier in literature (Fig. 1.1)[3,5]. Although these assignments can lead to valuable information, they ignore possible contributions of other biomolecules.

The aim of this work is to explore the possibilities to obtain information about cell compounds from bacterial Raman spectra. This includes the study whether these Raman spectra reflect differences in culturing conditions and whether these differences provide useful information about the cell's composition or metabolism.

This thesis starts with some theoretical considerations about microbiology (Chapter 2), Raman spectroscopy (Chapter 3) and data processing (Chapter 4). In order to get a better idea of the possible contribution of several cell compounds to a bacterial Raman spectrum, a database of reference Raman spectra of biomolecules present in cells was constructed and is discussed in Chapter 5. These spectra serve as basic knowledge for the extraction of

information about bacterial compounds. Whether these biomolecules will contribute significantly to bacterial Raman spectra depends on their intrinsic Raman activity as well as on their concentration in the cell. Therefore, in some cases, Raman bands in bacterial spectra can approximately be assigned to a specific biomolecule. For these cases, the possibilities of extracting information from bacterial Raman spectra are illustrated in Chapter 6. For example, the intensity of these bands can be directly compared to each other in different spectra, in order to deduce changes between different populations of cells. However, in many cases, it is incorrect to assume that a bacterial Raman band is approximately caused by only one biomolecule. Then, more complex approaches are necessary to obtain information about specific cell compounds, which are illustrated in Chapter 7. One of the illustrated methods, namely the use of coefficients from an EMSC (extended multiplicative signal correction) procedure, is extended in Chapter 8 to obtain information about structurally very similar biomolecules.



Figuur 1.1 Tentative band assignments in a typical bacterial Raman spectrum according to Maquelin *et al.*[5], here applied on a spectrum of *Cupriavidus metallidurans* LMG 1195.

Chapter 9 explores the potential of Raman spectroscopy to monitor a Raman active cell compound (semi-)quantatively. A case study in Chapter 10 applies some of the described methods to compare the behavior of a bacterial strain cultured in microgravity conditions and in a control setup. Finally, in Chapter 11, some conclusions and suggestions for further research are made.

References

1. Hutsebaut D., Vandroemme J., Heyrman J., Dawyndt P., Vandenabeele P., Moens L., De Vos P. (2006) Raman microspectroscopy as an identification tool within the phylogenetically homogeneous '*Bacillus subtilis*'-group. *Syst. Appl. Microbiol.* 29: 650-660.
2. Kirschner C., Maquelin K., Pina P., Thi N.A.N., Choo-Smith L.P., Sockalingum G.D., Sandt C., Ami D., Orsini F., Doglia S.M., Allouch P., Mainfait M., Puppels G.J., Naumann D. (2001) Classification and identification of *enterococci*: a comparative phenotypic, genotypic, and vibrational spectroscopic study. *J. Clin. Microbiol.* 39: 1763-1770.
3. Lopez-Diez E.C., Goodacre R. (2004) Characterization of microorganisms using UV resonance Raman spectroscopy and chemometrics. *Anal. Chem.* 76: 585-591.
4. Maquelin K., Dijkshoorn L., van der Reijden T.J.K., Puppels G.J. (2006) Rapid epidemiological analysis of *Acinetobacter* strains by Raman spectroscopy. *J. Microbiol. Meth.* 64: 126-131.
5. Maquelin K., Krischner C., Choo-Smith L.P., van den Braak N., Endtz H.Ph., Naumann D., Puppels G.J. (2002) Identification of medically relevant microorganisms by vibrational spectroscopy. *J. Microbiol. Meth.* 51: 255-271.
6. Naumann D., Keller S., Helm D., Schultz C., Schrader B. (1995) FT-IR spectroscopy and FT-Raman spectroscopy are powerful analytical tools for the noninvasive characterization of intact microbial cells. *J. Mol. Struct.* 347: 399-504.
7. Oust A., Moretro T., Naterstad K., Sockalingum G.D., Adt I., Manfait M., Kohler A. (2006) Fourier transform infrared and Raman spectroscopy for characterization of *Listeria monocytogenes* strains. *App. Environ. Microb.* 72: 228-232.

Chapter 2: Microbiological aspects

In this chapter, some theoretical considerations are presented concerning microbiology. General aspects of the bacterial cell, population growth and taxonomy are discussed. In addition, some specific characteristics of bacteria that are of importance for this work are treated, such as sporulation, PHB formation and the nomenclature and characteristics of some specific species.

2.1 Bacteria in the pool of living organisms

The universal tree of life was built on basis of (merely 16S) rRNA gene sequence comparisons. In 1978, Carl R. Woese proposed three domains based on the cell type of organisms: Bacteria, Archaea and Eukarya. Both Bacteria and Archaea display a prokaryotic cell type in contrast to the eukaryotic cells of the Eukarya. In addition to differences in rRNA gene sequences, the organisms of these domains differ in cell composition, membrane lipid structure, sensitivity to antibiotics, etc. The domain Bacteria includes pathogenic bacteria as well as the vast majority of nonpathogenic ones, each displaying wide metabolic diversity. The domain Archaea includes prokaryotes that lack peptidoglycan in their cell walls and often live in extreme environments. According to the endosymbiotic hypotheses Eukarya have developed by nucleus formation in a pro-eukaryotic cell, due to fusion of ancient bacteria and Archaea or nucleus formation from the Golgi apparatus in Archaea[28,39].

2.2 The bacterial cell

2.2.1 General bacterial cell constitution

In general, a cell consists of a cell membrane enclosing the cytoplasma containing macromolecules, ribosomes (protein-synthetizing factories that consist of rRNA and proteins), small organic molecules and inorganic ions[18]. These entities are surrounded by a

cell wall to give structural strength. More specifically, bacterial cells are prokaryotic and differ from eukaryotic cells by the absence of a membrane-enclosed nucleus. They have a simple internal structure and are in general very small (typically one to a few μm in length). They are surrounded by a cell wall from which the rigidity is caused by a substance called peptidoglycan or murein. The basic structure of peptidoglycan is a sheet where glycan chains (of N-acetylglucosamine and N-acetylmuramic acids) are connected by peptide cross-links (of L-alanine, D-alanine, D-glutamic acid and lysine or diaminopimelic acid). The degree of cross-linking determines the rigidity and varies amongst different bacteria. Structural differences in the cell wall divide the Bacteria in Gram-positive and Gram-negative (Fig. 2.1). The Gram-negative cell wall is complex and multilayered, while the Gram-positive cell wall consists mainly of a thick layer of peptidoglycan.

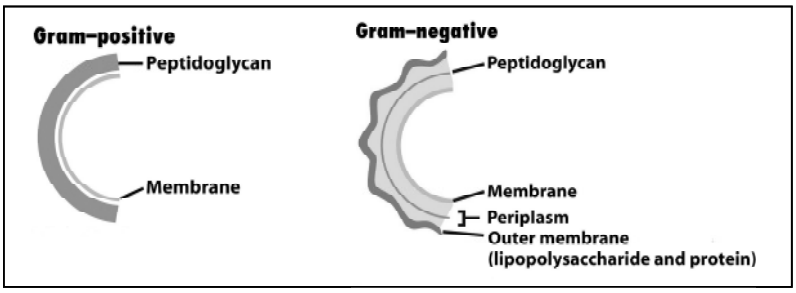


Figure 2.1 Composition of the cell wall of Gram-positive and Gram-negative bacteria[18]

In a prokaryotic cell, there is generally one large double-stranded circular DNA molecule, called the bacterial chromosome, which aggregates to form the nucleoid. More recently, it has been shown that various bacteria (e.g. *Burkholderia cepacia*[46]) contain more than one chromosome. Because the chromosome contains only one copy of most genes, these genes are genetically haploid. In addition to the bacterial chromosome(s), prokaryotes may contain extrachromosomal (circular) DNA, called plasmids. The genes on plasmids usually encode special properties, such as antibiotic resistance, in contrary to the essential housekeeping genes located on the chromosome. The chemical composition of a prokaryotic cell is given in Table 2.1.

Table 2.1 Chemical composition of a prokaryotic cell[18]

Molecule	Percent of dry weight
Total macromolecules	96
Protein	55
Polysaccharide	5
Lipid	9.1
Lipopolysaccharide	3.4
DNA	3.1
RNA	20.5
Total monomers	3
Amino acids and precursors	0.5
Sugars and precursors	2
Nucleotides and precursors	0.5
Inorganic ions	1

2.2.2 Nutrition and metabolic systems

The most important macronutrients for prokaryotes are carbon and nitrogen containing compounds[18]. Carbon is necessary for the assimilation of many cell compounds, such as fatty acids, amino acids, etc. Nitrogen is mainly present in nature under inorganic form (mostly NH_4^+, NO_3^-, N_2). Most bacteria only need NH_3 as nitrogen source, but many of them can also use NO_3^- and NO_2^-. Only a limited number of bacteria, called nitrogen-fixers, can use N_2 as a nitrogen source. Other nutrients are listed in Table 2.2.

Table 2.2 Nutrients for microbial growth, apart from C and N sources[18] (essential nutrients are marked with *)

Nutrient	Function in the cell
P* (organic and inorganic)	In nucleic acids, phospholipids and ATP
S* (sulfate (SO_4^{2-}) or sulphide (HS^-))	In amino acids (cysteine, methionine), vitamins (thiamin, biotin, lipoic aicd) and coenzyme A
K*	In enzymes
Mg*	For the acivity of enzymes and stabilization of ribosomes, cell membranes and nucleic acids
Ca	Stabilization of cell wall and endospores
Na	Can be required in certain habitats
Fe*	Major role in cellular respiration: cytochromes and iron-sulphur proteins for electron transport
Trace elements*	In enzymes
Growth factors (vitamins, amino acids, purines and pyrimidines)	Vitamins → cofactors Amino acids → proteins Purines and pyrimidines → DNA and RNA

In biological systems, energy conservation involves mainly oxidation-reduction reactions[18]. In general, the formation of adenosine triphosphate (ATP) is the central step in energy conservation. ATP is a high-energy compound (high free energy of hydrolysis) that conserves the energy released by oxidation-reduction reactions during catabolism. Subsequently, ATP can be used as an energy source for biosynthesis (anabolism) and other cell activities. According to the energy source used for ATP production, we distinguish chemotrophs and phototrophs.

Chemotrophs use chemicals as energy source for their metabolism, via different energy conservation mechanisms according to the type of electron donor en electron acceptor.

Focusing on the differences in terminal electron acceptor, we distinguish:

- Fermentation: ATP is produced only by substrate-level phosphorylation, thus during the catabolism of an organic compound. The redox processes occur in absence of a terminal electron acceptor. In fermentation, oxidation of a compound is coupled to reduction of compounds generated from the initial substrate to balance the electrons. Fermentation occurs frequently through glycolysis with the conversion of glucose to fermentation products (e.g. ethanol and CO_2, or lactic acid).
- Respiration: the carbon flow in respiration occurs usually through glycolysis (formation of pyruvate) and subsequently the citric acid cycle (oxidation of pyruvate to CO_2 and ATP formation by substrate-level phosphorylation). During electron transport over (membrane associated) carriers, such as the coenzymes nicotinamide-adenine dinucleotide (NAD^+) and NAD-phosphate ($NADP^+$), flavoproteins (riboflavin-containing electron carriers), iron-sulphur proteins, quinones and cytochromes, a pH gradient is created across the membrane by translocation of protons. ATP is then produced by oxidative phosphorylation, thus at the expense of the proton motive force. Molecular oxygen (aerobic respiration)

or other molecules such as NO_3^-, Fe^{3+}, SO_4^{2-}, CO_3^{2-} and organic compounds (anaerobic respiration) are used as final electron acceptor. During electron transport, protons originating from the proton donating substance (e.g. sugars) are translocated through the membrane resulting in a pH gradient. The proton motive force can then be used to form ATP via ATP synthase. The presence of a final electron acceptor allows the full oxidation of pyruvate to CO_2, which results in a larger energy yield compared to fermentation. Furthermore, different intermediates from the citric acid cycle can be redrawn for biosynthesis (anabolism).

Differences in electron donor molecules lead to a second distinction within the chemotrophs:

- Chemoorganotrophy: organic compounds are used as electron donors.
- Chemolithotrophy: inorganic compounds such as H_2S, H_2, Fe^{2+} and NH_3 instead of organic compounds are used as electron donors. As the electron donors are inorganic, chemolithotrophs need another source of carbon for biosynthesis. Most chemolithotrophs use atmospheric CO_2 and hence are refered to as autotrophs. However, some chemolithotrophs lack enzymes of the calvin cycle and need organic compounds as a carbon source. These are called mixotrophs because their electron donor is inorganic while their carbon source is organic.

Phototrophs use light as an energy source to create a proton motive force and hence to produce ATP (light reaction). For biosynthesis (dark reaction), they can either use CO_2 (photoautotrophs) or organic compounds as carbon sources (photoheterotrophs). To drive autotrophic reactions, some bacteria obtain reducing power from electron donors such as H_2O (oxygenic) or reduced sulphur sources, H_2, H_2S, Fe^{2+}, etc. (anoxygenic). Prefixes and terms used to indicate types of metabolism are summarized in Figure 2.2.

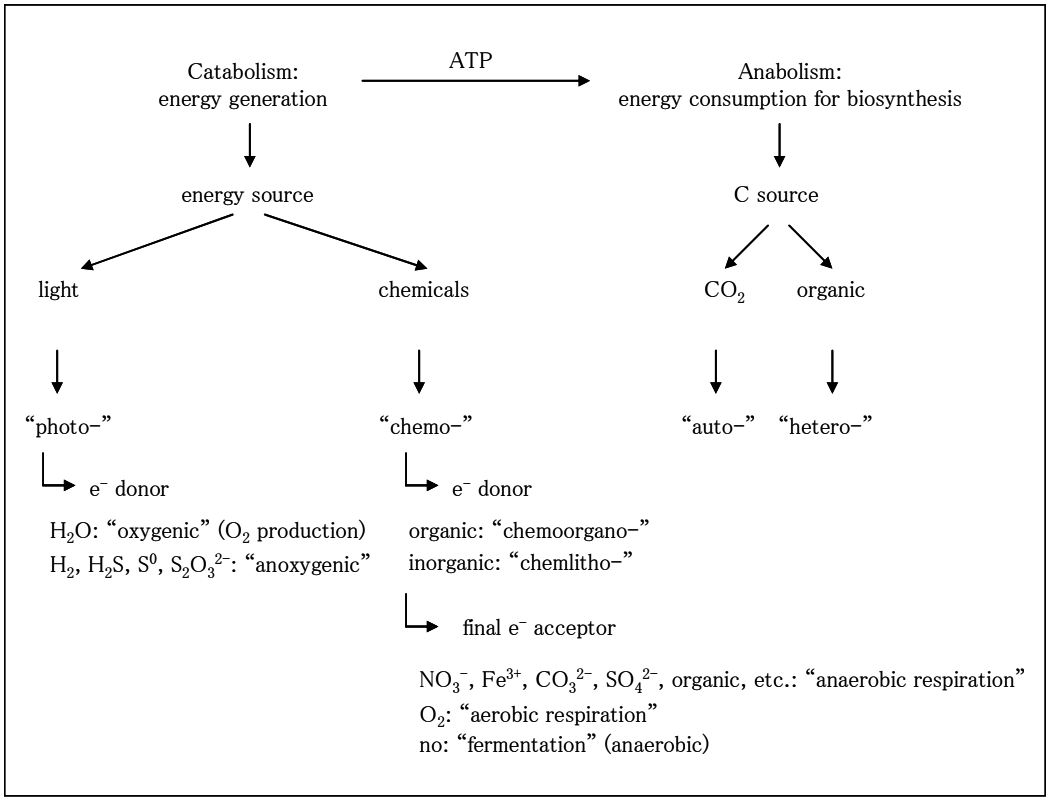


Figure 2.2 Summary of prefixes and terms used to describe different types of metabolism[18]

Combination of the type of energy source, electron donor, electron acceptor and carbon source, leads to a specific type of metabolism, from which some examples are given in Table 2.3.

Table 2.3 Examples of bacterial metabolism[18]

metabolism	E source	e– donor	e– acceptor	C source
denitrification	chemo	organic	$NO_3^-/NO_2^- \rightarrow N_2$	organic (hetero)
nitrification	chemo	$NH_4^+/NO_2^- \rightarrow NO_3^-$	O_2	CO_2 (auto)
anammox	chemo	$NH_4^+ \rightarrow N_2$	NO_3^-	CO_2 (auto)
Sulphate en S reduction	chemo	H_2	$SO_4^{2-}/S \rightarrow H_2S$	CO_2 (auto)
	chemo	organic	$SO_4^{2-}/S \rightarrow H_2S$	organic (hetero)
S oxidation	chemo	$H_2S/S_2O_3/S \rightarrow SO_4^{2-}$	O_2/ NO_3^-	CO_2 (auto)
methanogenesis	chemo	H_2	CO_2	CO_2 (auto)
acetogenesis	chemo	H_2	CO_2	CO_2 (auto)
methanotrophy	chemo	$CH_4 \rightarrow CO_2$	O_2	CH_4
cyanobacteria	photo	H_20	($NADP^+$)	CO_2 (auto)

2.2.3 Sporulation

Certain *Bacteria* form endospores within their cells. Endospore-forming bacteria are found most commonly in soil and the best studied endospore-forming bacteria are members of *Bacillus*, *Clostridium* and relatives[18].

Endospores are differentiated cells that are structurally more complex than vegetative cells and consist of (Fig. 2.3):

- The exosporium: the outermost layer is thin and constituted of proteins.
- Spore coat: layers of spore-specific proteins.
- The cortex: loosely cross-linked peptidoglycan.
- The core of the spore: it contains the usual cell structures such as the cell wall (core wall), cytoplasmic membrane, cytoplasm, nucleoid, etc.

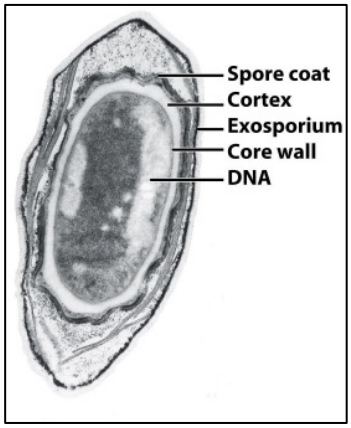


Figure 2.3 Schematic composition of an endospore[18]

The structure of an endospore thus differs from that of a vegetative cell by the layers outside the core wall. In addition, pyridine-2,6-dicarboxylic acid (Appendix) or dipicolinic acid (DPA) is an important spore compound which is not present in vegetative cells. It is located in the spore core, where it forms a complex with calcium ions. The calcium dipicolinate complex (CaDPA) represents about 10% of the endospores dry weight. Other spore-specific compounds are small acid soluble proteins (SASPs). They bind tightly to DNA to protect it and serve as a carbon energy source when the spore germinates. The

core is in a partially dehydrated state as it contains only 10 to 30% of the water content of a vegetative cell. Its pH is about one unit lower than that of a vegetative cell.

Endospore formation (sporulation) occurs when growth ceases because essential nutrients, such as carbon and/or nitrogen, become depleted. The synthesis of some proteins necessary in certain vegetative cell functions cease, while specific spore proteins are synthesized. Sporulation involves a series of cellular differentiation and can be divided into 7 stages that are illustrated in Figure 2.4.

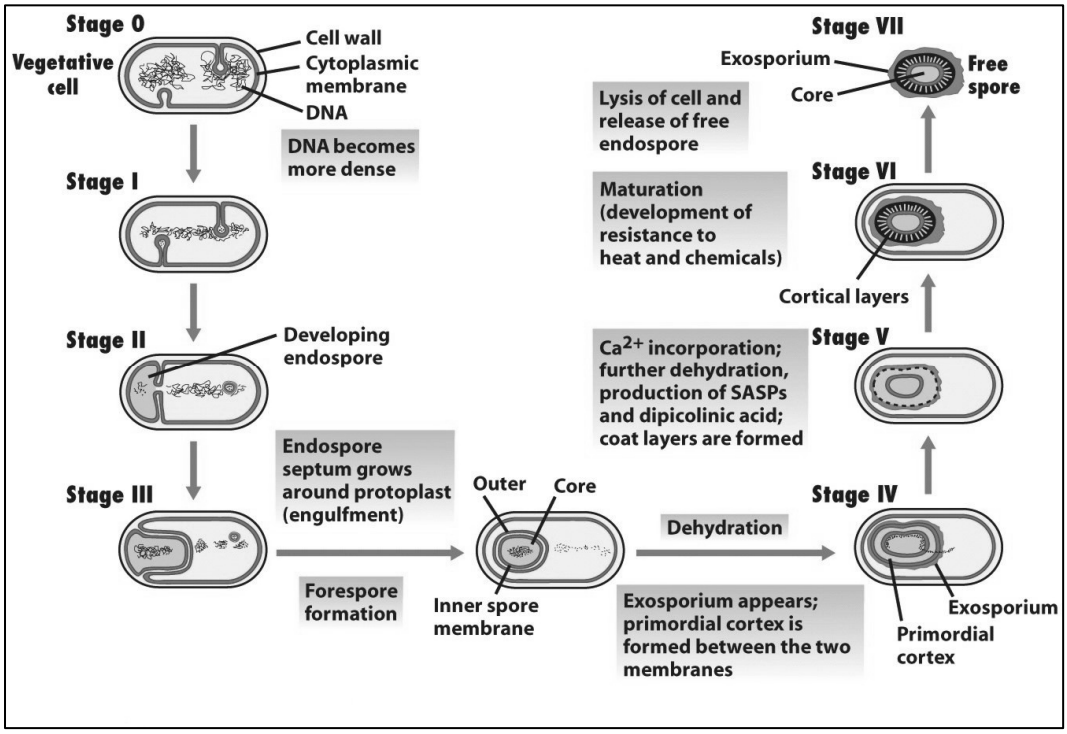


Figure 2.4 Seven steps towards endospore formation[18]

A free endospore can show resistance to dry heat, wet heat, UV radiation and chemicals such as hydrogen peroxide, nitrous acid and formaldehyde (up to certain concentrations, which can differ per species). Setlow[33] reported that (a combination of) properties such as the content of water, SASPs, CaDPA and core minerals can enhance resistance to these environmental conditions. This way, a spore can remain dormant for many years and can be rapidly converted into a vegetative cell in three steps. The first step includes activation of the spore, which is followed by germination. During germination the spore loses CaDPA,

cortex compounds and its resistance. SASPs are degraded and the third stage, outgrowth, can begin. The cell swells by water uptake and synthesis of DNA, RNA and proteins, and is ready to divide.

2.2.4 PHB production

Certain bacteria synthesize storage materials which are often included in granules or inclusions[18]. One of the most common storage materials is poly-3-hydroxybutyrate (PHB), which is formed by many bacteria when growth conditions become unfavorable, such as a high C/N ratio[1]. PHB is a biodegradable polyester and its production by bacteria is of importance for the industry. *Cupriavidus necator* (previously known as *Ralstonia eutropha* and *Alcaligenes eutrophus*) is the most widely used organism for industrial PHB production[25]. PHB consists of β-3-hydroxybutyric acid units that are connected by ester linkages (Fig. 2.5). The length of the polymer can vary from 4 up to 18 monomer units in certain organisms[18].

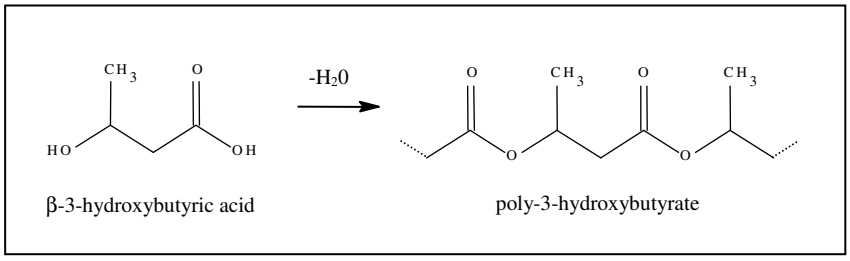


Figure 2.5 Chemical structure of β-3-hydroxybutyric acid and poly-3-hydroxybutyrate

Carbon sources such as glucose, acetate, lactate and butyrate are favorable for PHB production. Figure 2.6 shows the metabolic flux of this species grown on several carbon sources[34]. During PHB accumulation, cells do not grow or divide. Other polyhydroxyalkanoates (PHA's) can be formed when carbon sources with different chain lengths are included in the medium, such as the formation of a copolymer of 3-hydroxybutyrate and 3-hydroxyvalerate units from glucose and propionic acid. The composition of these polyesters is determined by the relative concentration of the carbon sources present in the medium[1].

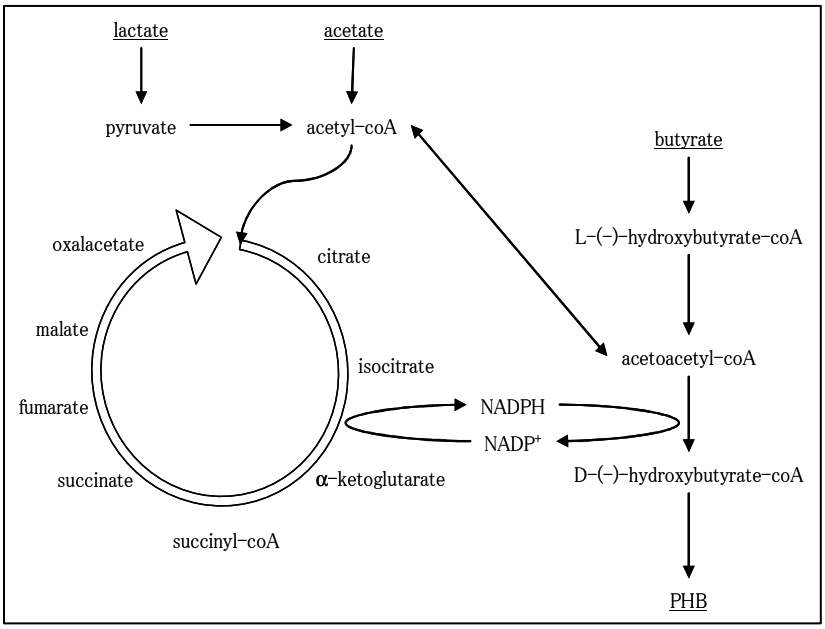


Figure 2.6 Metabolic cycle for PHB production in *C. necator* from various carbon sources[34]

Jurasek and Marchessault[16] gave an overview of granule formation in *C. necator*. As PHB chains grow, agglomerations are formed attached to PHB synthase. Newly formed PHB synthase molecules are swept up by the PHB surfaces and so granules start to form. Collisions of several granules may lead to coalescence and thus growth of the granules. However, by growing, the mobility of the granules decreases. In addition, the protein phasing, which has a high affinity for PHB surfaces, is synthesized in large amounts and covers a large part of the PHB granule surface. By these two effects, the average granule diameter reaches a plateau in the later stage of the accumulation process. Eventually, there is a dynamic equilibrium on the granule's surface between synthase, phasing and PhaR protein that blocks the transcription of the phasing gene. In addition, the granule surface contains a small number of phospholipids and PHB depolymerase. The latter converts PHB into digestible products when for example growth conditions are favorable and/or when lacking essential nutrients become available. The final number of PHB granules is 10 on average with each a diameter of about 500 nm, leading to PHB yield of up to 90% of the cell's dry weight. An electron micrograph of bacterial cells containing PHB granules is shown in Figure 2.7.

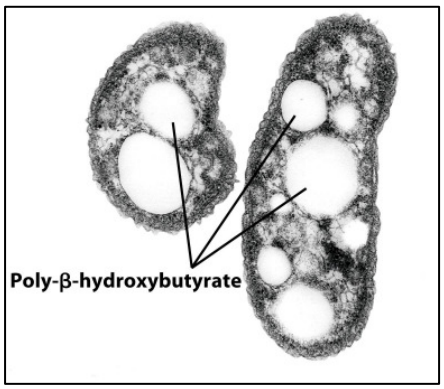


Figure 2.7 Electron micrograph of a thin section of cells of *Rhodovibrio sodomensis* containing PHB granules[18]

2.3 Bacterial growth

Bacterial growth is defined as an increase in the number of microbial cells in a population[18]. The time required for the formation of two cells from one cell (cell division) is called the generation time. Although this parameter can vary substantially, most bacteria have a generation time of 1 to 3 hours. After each time interval corresponding to the generation time, the number of cells in a population doubles, which is called exponential growth. This implies that the initial increase in absolute cell number is low, but it increases per generation to a high number. In a batch culture (i.e. a closed system with a finite amount of nutrients), there are typically four stages of growth that describe a growing population (not individual cells) (Fig. 2.8). When fresh medium is inoculated, growth usually does not start immediately, but after a lag phase. This phase is necessary when the inoculum contains cells of an old culture or when the population is transferred from a rich to a poor medium. The lag phase is then respectively the time needed to resynthesize essential constituents or to synthesize new enzymes that can metabolize the metabolites in the medium. The exponential phase can start after this lag phase or immediately after inoculation. Exponential growth cannot occur indefinitely because essential nutrients become depleted and/or waste products build up and have an inhibitory toxic effect. The stationary phase begins at the point where there is no net increase in cell number. Either there is no growth or growth and death balance out (cryptic growth). Even when there is no growth, cell

functions such as energy metabolism and some biosynthetic processes may continue. When incubation continues, cells may start to die, usually at a slow rate but faster than they divide. This is called the death phase.

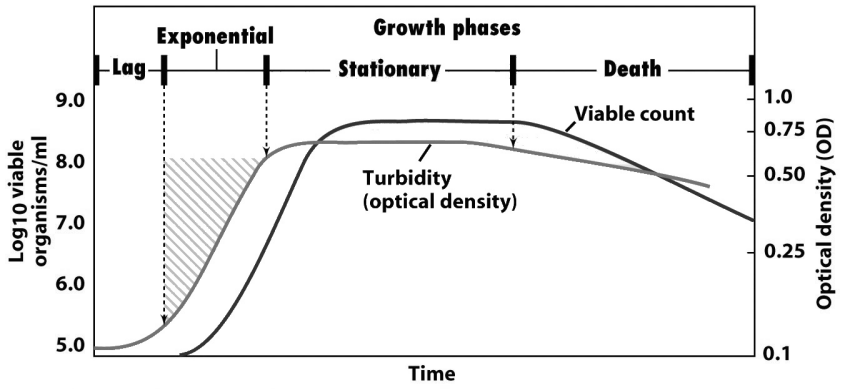


Figure 2.8 A typical growth curve of a bacterial population[18]

2.4 Taxonomy

2.4.1 Definition and techniques

Taxonomy includes three aspects: (i) the classification of bacteria in groups on the basis of similarities, (ii) the naming of these groups (nomenclature) and (iii) verification whether a bacterium belongs to one of these groups classified in (i) and named in (ii) (identification)[41]. Today, microbiologists use a combination of methods to differentiate organisms on both genetic and phenotypic ground, named polyphasic taxonomy[18]. Vandamme *et al.* gave an overview of the techniques that can be used in polyphasic taxonomy and evaluated them for various taxonomic levels[41].

Classical bacterial taxonomy relies on several characteristics such as various aspects of morphology, nutrition, physiology and habitat[18]. **Morphological characteristics** have helped taxonomists to classify organisms in the past and some of them, such as the presence of endospores and flagella, can still be useful[39]. **Differential staining** (Gram stain or acid-fast stain) is based on the chemical composition of cell walls and is used to obtain information quickly in a chemical environment[39]. **Biochemical tests** that evaluate the enzymatic activities of an organism in specific culturing conditions can be performed using

commercially available kits developed for medically important bacteria[39]. Next to these phenotypic characteristics, determination of the **GC ratio** is also part of classical bacterial taxonomy[18]. This ratio reflects the guanine plus cytosine content of an organism's genomic DNA and is expressed as (G+C)/(A+T+G+C)×100%. GC ratios can vary from 20 to 80 % in the microbial world and can be identical for unrelated organisms. Therefore the GC ratio of an organism is only of limited taxonomic value.

Molecular taxonomy is based on DNA related techniques of analysis. Comparative analysis of **16S rRNA gene** sequences reflects the bacterial phylogeny and forms the basis for bacterial systematics. Phylogeny is defined as the study of the evolutionary history[28]. It reveals the evolutionary relatedness between bacteria rather than their general resemblance. Today, most scientists perform this 16S rRNA gene sequencing by using the polymerase chain reaction (PCR) with a specific primer to amplify the genomic DNA genes that encode 16S rRNA and then sequence the PCR product by standard dideoxy (Sanger) sequencing. **DNA-DNA hybridization**, which measures the degree of DNA sequence similarity, is useful for differentiation of closely related organisms at the species level. Practically, an organism's isolated DNA is sheared, radioactively (^{32}P or ^{3}H) or fluorescently labeled and denatured. This DNA is mixed with a similarly prepared sample from another organism and cooled down so the DNA can reanneal. The double stranded DNA is separated from the unhybridized DNA and its radioactivity or fluorescence is determined compared to the control (100%). In **MLSA** (multilocus sequence analysis), several conserved housekeeping genes are sequenced. This technique was proposed as an alternative for DNA:DNA hybridization to detect differences at the species level[7].

Several characterization techniques rely on the generation of a DNA fingerprint. **Ribotyping** is one of the methods used for DNA-based typing. After restriction enzyme digestion of DNA from a particular organism, the fragments are separated and probed with a labeled ribosomal RNA probe, for example 16S rRNA or a fragment of it[18,41]. For **AFLP**[15] (amplified fragment length polymorphism) analysis, two restriction enzymes are used that cut specific sticky ends. Next, adapters are ligated to these sticky ends. Subsequently, the DNA fragments are selectively amplified using primers that are slightly longer than the length of

the adapter and the sticky end sequences. In this way, the PCR reaction and subsequent electrophoresis result in a fingerprint pattern. A fingerprint pattern can also be generated without the use of a restriction enzyme, when choosing an appropriate primer for PCR. In **rep-PCR** (repetitive element primed - polymerase chain reaction) one or two conservative primers are used (e.g. REP-PCR[44] and BOX-PCR[43]), while in random primed PCR random primers of variable length (5- to 20-mers) are used. Both techniques result after electrophoresis in a characteristic pattern that may be strain specific.

C**hemotaxonomic techniques** are analytical techniques that provide phenotypic information. One of the most important chemotaxonomic techniques is **FAME** (fatty acid methyl ester) analysis which determines the fatty acid composition of the membrane and outer membrane lipids of cells. Therefore, fatty acid are extracted from cell hydrolysates and derivatized to methyl esters which are volatile and can be detected by gas chromatography[18]. FAME easily allows grouping of many organisms according to the resemblance of their fatty acid pattern. In addition, it can be used for identification when standardized culturing conditions are applied and an extensive database of FAME data is available.

2.4.2 Taxonomic ranks and Bergey's manual

In prokaryotic taxonomy, the commonly used ranks are (in descending order): phylum, class, order, family, genus and species. The basic taxonomic group is the species and every organism is named according to the **binomial nomenclature** (genus, species). Until 1987, a species was described as a collection of strains that share an important number of stable (genotypic and/or phenotypic) characteristics and that differ significantly from other groups of strains[28,39]. In 1978, Wayne *et al.*[45] described that the complete bacterial genome should be the reference for phylogeny and taxonomy. At that moment whole genome sequencing was not possible and whole genome DNA-DNA hybridization was found to be the best approach for analyzing whole genome similarity. The phylogenetic definition of a species includes strains with at least 60-70% DNA-DNA hybridization similarity. Phenotypic characteristics should agree with this definition. Bacteria that show less than 70%

relatedness in DNA-DNA hybridization experiments with known species, have generally less than 97% 16S ribosomal RNA gene sequences in common, as illustrated by the green box in Figure 2.9. Therefore, a bacterial strain whose 16S ribosomal RNA sequence differs by more than 3% from that of all other bacteria should be considered as belonging to a new species (more recently, even a 2% limit is used).

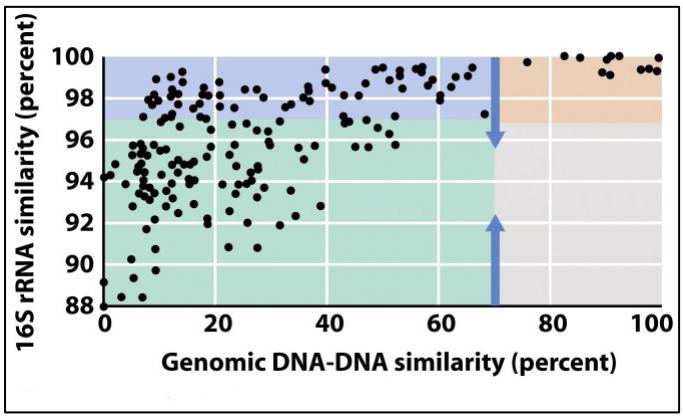


Figure 2.9 Relation between 16S rRNA sequence similarity and genomic DNA hybridization between different pairs of organisms[18]

Nevertheless some species show more than 97% similarity in 16S rRNA gene sequence, while their genomic DNA-DNA relatedness is low (blue box in Fig. 2.9). This illustrates that 16S rRNA gene sequencing alone is not sufficiently precise as a taxonomic marker for species delineation and that a polyphasic approach is necessary. The orange box in Figure 2.9 represents data from identical species.

Bergey's Manual of Systematic Bacteriology[6] is considered as the primary reference for bacterial taxonomy. The first edition of the book is a four volume set and gives a detailed description of species according to a taxonomic outline. Bergey's Manual of Determinative Microbiology[14] is dedicated to bacterial identification. It does not classify bacteria according to evolutionary relatedness, but provides identification schemes.

2.4.3 Taxonomic position of species studied in this work

The species studied in this work belong to the following genera:

Table 2.4 Taxonomic situation of the genera from which strains are used in this work

Genus	Family	Order	Class	Phylum
Cupriavidus	Burkholderiaceae	Burkholderiales	β-Proteobacteria	Proteobacteria
Bacillus	Bacillaceae	Bacillales	Bacilli	Firmicutes
Paenibacillus	Paenibacillaceae	Bacillales	Bacilli	Firmicutes
Enterococcus	Enterococcaceae	Lactobacillales	Bacilli	Firmicutes

Cupriavidus necator

Cupriavidus necator was first described by Makkar and Casida[19] in 1987. It is a Gram-negative aerobic, mesophilic rod and a non-obligate bacterial predator in soil. It has simple nutritive requirements when it is not acting as a predator and uses fructose (not glucose) or specific amino acids as a carbon source. It is highly resistant to copper, by which its growth is stimulated. This species has some characteristics in common with the genus *Alcaligenes*, such as testing negative for glucose degradation and a similar DNA G+C content. *C. necator* was classified in a separate genus because of its predatory activity, its use of fructose and resistance to and growth stimulation by copper. However, Makkar and Casida only studied morphological, biochemical, respiratory and nutritional characteristics of this organism and molecular characteristics.

In 2004, Vandamme and Coenye[40] reported that *Wautersia eutropha* is a later synonym of *Cupriavidus necator*. They compared results of several analyses of both taxa such as 16S rRNA gene sequencing, SDS-PAGE, DNA base ratio and finally performed DNA:DNA hybridization. *Wautersia eutropha* itself was formerly known as *Ralstonia eutropha* which was in its turn a reclassification of *Alcaligenes eutrophus* performed by Yabuuchi *et al.*[47] who also proposed *Burkholderia pickettii* and *Burkholderia solanacearum as Ralstonia* members. This generic reclassification was based on phenotypic characterization, cellular lipid and fatty acid analysis, phylogenetic analysis of 16S rDNA nucleotide sequences and

rRNA-DNA hybridization. Vaneechoutte *et al.*[42] showed that based on 16S rRNA gene sequence analysis two sublineages exist within the genus *Ralstonia*: the *Ralstonia eutropha* lineage and the *Ralstonia picketti* lineage. This observation was supported by phenotypic differences concerning the presence of flagella, the production of acid from glucose and the susceptibility to colistin. Therefore, it was proposed that the species of the *R. eutropha* lineage should be reclassified in a novel genus, *Wautersia*. When Vandamme and Coenye reported that *Wautersia eutropha* is a latter synonym of *Cupriavidus necator*, it became clear that the isolate described by Makkar and Casida in 1987 should have been originally classified as *Alcaligenes eutrophus*. Because the names *Ralstonia* and *Wautersia* were published later then *Cupriavidus necator*, the latter has priority.

Cupriavidus metallidurans

In the late 1970s, the strain CH34, a Gram-negative rod, was isolated from sediments of a decantation basin of a zinc factory near Liège (Belgium)[21]. This strain showed a heavy-metal resistance, which was shown to be a plasmid born characteristic. In 1985, this strain was considered as a member of *Alcaligenes eutrophus*, because of its hydrogen-oxidizing characteristics[22]. In 1995, the taxonomic status of *A. eutrophus* and some related strains was revised and these bacteria were reclassified as *Ralstonia*[47]. In 2001, Goris *et al.*[8] subjected 31 heavy metal resistant isolates from industrial biotopes to a polyphasic characterization. All strains belonged to the genus *Ralstonia* and two groups of 8 and 17 strains respectively were clearly delineated from the other species and showed low similarity to the existing *Ralstonia* strains. These isolates were proposed as two new *Ralstonia* species, of which one was named *Ralstonia metallidurans* with CH34 as type strain. When Vaneechoutte *et al.*[42] discovered that there were two lineages within the genus *Ralstonia*, *Ralstonia metallidurans* was reclassified in the genus *Wautersia*. When Vandamme and Coenye[40] showed that *Wautersia eutropha* was a synonym of *Cupriavidus necator*, they proposed that the name *Wautersia* was replaced by *Cupriavidus*, and thus that all species of the genus *Wautersia* should be considered as species of the genus *Cupriavidus*, including *Cupriavidus metallidurans*. The epithet *metallidurans* indicates that

this species is able to survive high heavy-metal concentrations. Therefore, this species offers perspectives for the treatment of waste-water and the recycling of polluted soils[8].

Bacillus species

The genus *Bacillus* is a member of the family Bacillaceae of which a distinguishing feature is the production of endospores. Members of the genus *Bacillus* are characterized as Gram-positive, rod-shaped, aerobic or facultative anaerobic, endospore-forming bacteria[6,38]. The phylogenetic heterogeneity within the genus *Bacillus* is very large; *Bacillus* species show a large diversity in physiology and their G+C content varies from 32 to 69%. In 1991, Ash *et al.*[2] performed 16S rRNA gene sequencing on 51 *Bacillus* species and found five phylogenetically distinct clusters which has led to the proposal of reclassifying some *Bacillus* species in several (new) genera:

- Group 1: *Bacillus sensu stricto* including the type species *B. subtilis* and 27 other species, some of them were later reclassified in genera such as *Virgibacillus*[13] and *Halobacillus*[37].
- Group 2 consisted of *B. sphaericus* and 5 other species form which three were reclassified as *Sporosarcina*[48].
- Group 3 was reclassified as *Paenibacillus*[3].
- Group 4 was reclassified as *Brevibacillus*[35].
- Group 5 was reclassified as *Geobacillus* [24].

The strains studied in this work are members:

- *B. cereus* is a species belonging to the *B. cereus* group which includes amongst others *B. thuringiensis*, *B. anthracis* and *B. cereus sensu stricto*[29]. This species is known to cause foodborne illnesses[9].
- *B. coagulans* was first described by Hammer[10] in 1915. This species was earlier known as *Lactobacillus sporogenes* but it was reclassified in 1939 as *B. coagulans*. This species indeed shares some characteristics with species of *Lactobacillus*, but was originally classified as *Bacillus* because of its spore-forming capacity[12].

- *B. licheniformis,* a species with high intraspecific diversity[4,5,20].
- *B. sporothermodurans*, a species that was proposed in 1996 for strains that produce highly thermoresistant spores that may survive the ultrahigh temperature treatment (UHT) of milk[26].
- *B. subtilis*, which was devided in 1999 into *B. subtilis* subsp. *subtilis* and subsp. *spizizzenii*[23].

Paenibacillus lactis

The genus *Paenibacillus* was proposed a new genus to reclassify a number of *Bacillus* species[3], belonging to rRNA group 3 of Ash *et al.*[2]. Later, many other named *Bacillus* species were transferred to this genus[27,36]. The members of this genus are facultative anaerobic and produce endospores. Their cells are rods of Gram-positive structure but they usually stain negatively in Gram stain[3]. In this work, a strain of *Paenibacillus lactis* was used. This species name was proposed by Scheldeman *et al.*[30] for 19 closely related *Paenibacillus* strains isolated from raw and heat-treated milk.

Enterococcus faecalis

Enterococci are facultative anaerobic, Gram-positive cocci and belong to the lactic acid bacteria. *Enterococcus faecalis* inhabits the gastrointestinal tracts of humans and other mammals and can cause infections. It was known as *Streptococcus faecalis* until the 1980s when the heterogeneous genus *Streptococcus* was divided into *Streptococcus* sensu stricto and two novel genera *Enterococcus*[31] and *Lactococcus*[32]. The taxonomy of the species *Streptococcus* and *Enterococcus* were summarized by Hardie *et al.*[11] and Köhler[17].

In this chapter, some theoretical considerations about microbiology were discussed. The next chapter presents theoretical aspects of Raman spectroscopy.

References

1. Anderson A.J., Dawes E.A. (1990) Occurrence, metabolism, metabolic role, and industrial uses of bacterial polyhydroxyalkanoates. *Microbiol. Rev.* 54: 450-472.
2. Ash C., Farrow J.A.E., Wallbanks S., Collins M.D. (1991) Phylogenetic heterogeneity of the genus *Bacillus* revealed by comparative-analysis of small-subunit-ribosomal RNA sequences. *Lett. Appl. Microbiol.* 13: 202-206.
3. Ash C., Priest F.G., Collins M.D. (1993) Molecular-identification of ribosomal-RNA group 3 *Bacilli* (Ash, Farrow, Wallbanks and Collins) using a PCR probe test - proposal for the creation of a new genus *Paenibacillus*. *Antonie Van Leeuwenhoek* 64: 253-260.
4. De Clerck E., De Vos P. (2004) Genotypic diversity among *Bacillus licheniformis* strains from various sources. *FEMS Microbiol. Lett.* 231: 91-98.
5. Duncan K.E., Ferguson N., Kimura K., Zhou X., Istock C.A. (1994) Fine-scale genetic and phenotypic structure in natural populations of *Bacillus subtilis* and *Bacillus licheniformis* - Implications for bacterial evolution and speciation. *Evolution* 48: 2002-2025.
6. Garrity, G.M., Staley J.T., Boone D.R., Gastenholtz R.W. (2001) Bergey's manual of systematic bacteriology, 2nd ed. Springer, New York.
7. Gevers D., Cohan F.M., Lawrence J.G., Spratt B.G., Coenye T., Feil E.J., Stackebrandt E., Van de Peer Y., Vandamme P., Thompson F.L., Swings J. (2005) Re-evaluating prokaryotic species. *Nat. Rev. Microbiol.* 3: 733-739.
8. Goris J., De Vos P., Coenye T., Hoste B., Janssens D., Brim H., Diels L., Mergeay M., Kersters K., Vandamme P. (2001) Classification of metal-resistant bacteria from industrial biotopes as *Ralstonia campinensis* sp. nov., *Ralstonia metallidurans* sp. nov. and *Ralstonia basilensis* Steine *et al.* 1998 emend. *Int. J. Syst. Evol. Microbiol.* 51: 1773-1782.
9. Granum P.E., Lund T. (1997) *Bacillus cereus* and its food poisoning toxins. *FEMS Microbiol. Lett.* 157: 223-228.

10. Hammer B.W. (1915) Bacteriological studies on the coagulation of evaporated milk. *Iowa Agric. Exp. Stn. Res. Bull.* 19: 119-131.
11. Hardie J.M., Whiley R.A. (1997) Classification and overview of the genera *Streptococcus* and *Enterococcus*. *J. Appl. Microbiol.* 83: 1S-11S.
12. Hartemink R. (2006) What's in a name - *Bacillus coagulans* or *Lactobacillus sporogenes*? Wageningen University, http://www.food-info.net/uk/ff/sporogenes.htm
13. Heyndrickx M., Lebbe L., Kersters K., De Vos P., Forsyth C., Logan N.A. (1998) *Virgibacillus*: a new genus to accommodate *Bacillus pantothenticus* (Proom and Knight 1950). Emended description of *Virgibacillus pantothenticus*. *Int. J. Syst. Bacteriol.* 48: 99-106.
14. Holt J.G. (1994) Bergey's manual of determinative bacteriology, 9th ed. Williams ad Wilkins, Baltimore.
15. Janssen P., Coopman R., Huys G., Swings J., Bleecker M., Vos P., Zabeau M., Kersters K. (1996) Evaluation of the DNA fingerprinting method AFLP as a new tool in bacterial taxonomy. *Microbiology* 142: 1881-1893.
16. Jurasek L., Marchessault R.H. (2004) Polyhydroxxyalkanoate (PHA) granule formation in *Ralstonia eutropha* cells: a computer simulation. *Appl. Microbiol. Biotechnol.* 64: 611-617.
17. Köhler W. (2007) The present state of species within the genera *Streptococcus* and *Enterococcus*. *Int. J. Med. Microbiol.* 297: 133-150.
18. Madigan M.T., Martinko J.M., Parker J. (2005) Brock biology of microorganisms, 11th ed. Prentice Hall International, Pearson education Inc., Upper Saddle River.
19. Makkar N.S. and Casida L.E. (1987) *Cupriavidus necator* gen. nov., sp. nov.: a non-obligate bacterial predator of bacteria in soil. *Int. J. Syst. Bacteriol.* 37: 323-326.
20. Manachini P.L., Fortina M.G., Levati L., Parini C. (1998) Contribution to phenotypic and genotypic characterisation of *Bacillus licheniformis* and description of new genomovars. *Syst. Appl. Microbiol.* 21: 520-529.

21. Mergeay M., Houba C., Gerits J. (1978) Extrachromosomal-inheritance controlling resistance to cadmium, cobalt, copper and zinc ions - evidence from curing in a pseudomonas. *Arch. Int. Physiol. Bio.* 86: 440-442.
22. Mergeay M., Nies D., Schlegel H.G., Gerits J., Charles P., Van Gijsegem F. (1985) *Alcaligenes eutrophus* CH34 is a facultative chemolithotroph with plasmid-bound resistance to heavy metals. *J. Bacteriol.* 162: 328-334.
23. Nakamura L.K., Roberts M.S., Cohan F.M. (1999) Relationship of *Bacillus subtilis* clades associated with strains 168 and W23: a proposal for *Bacillus subtilis subsp subtilis* subsp. nov. and *Bacillus subtilis subsp spizizenii* subsp. nov. *Int. J. Syst. Bacteriol.* 49: 1211-1215.
24. Nazina T.N., Tourova T.P., Poltaraus A.B., Novikova E.V., Grigoryan A.A., Ivanova A.E., Lysenko A.M., Petrunyaka V.V., Osipov G.A., Belyaev S.S., Ivanov M.V. (2001) Taxonomic study of aerobic thermophilic bacilli: descriptions of *Geobacillus subterraneus* gen. nov., sp. nov. and *Geobacillus uzenensis* sp. nov. from petroleum reservoirs and transfer of *Bacillus stearothermophilus*, *Bacillus themocatenulatus*, *Bacillus thermoleovorans*, *Bacillus kausophilus*, *Bacillus thermodenitrificans* to *Geobacillus* as the new combinations *G. stearothermophilus*, *G. themocatenulatus*, *G. thermoleovorans*, *G. kausophilus*, *G. thermodenitrificans*. *Int. J. Evol. Microbiol.* 51: 433-446.
25. Patnaik P.R. (2006) Dispersion optimization to enhance PHB production in fed-batch cultures of *Ralstonia eutropha*. *Bioresour. Technol.* 97: 1994-2001.
26. Pettersson B., Lembke F., Hammer P., Stackebrandt E., Priest F.G. (1996) *Bacillus sporothermodurans*, a new species producing highly heat-resistant endospores. *Int. J. Syst. Bacteriol.* 46: 759-964.
27. Pettersson B., Rippere K.E., Yousten A.A., Priest F.G. (1999) Transfer of *Bacillus lentimorbus* and *Bacillus popilliae* to the genus *Paenibacillus* with embedded descriptions of *Paenibacillus lentimorbus* comb. nov, and *Paenibacillus popilliae* comb. nov. *Int. J. Syst. Bacteriol.* 54: 531-540.
28. Prescott L.M., Harley J.P., Klein D.A. (2002) Microbiology, 5th ed. McGrac-Hill, Boston.

29. Rasko D.A., Altherr M.R., Han C.S., Ravel J. (2005) Genomics of *Bacillus cereus* group of organisms. *FEMS Microbiol. Rev.* 29: 303-329.
30. Scheldeman P., Gooossens K., Rodriguez-Diaz M., Pil A., Goris J., Herman L., De Vos P., Logan N.A., Heyndrickx M. (2004) *Paenibacillus lactis* sp. nov., isolted from raw and heat-treated milk. *Int. J. Syst. Evol. Microbiol.* 54: 885-891.
31. Schleifer K.H., Kilpperbalz R. (1984) Transfer of *Streptococcus faecalis* and *Streptococcus faecum* to the genus *Enterococcus* nom. rev. as *Enterococcus faecalis* comb. nov. and *Enterococcus faecum* comb. nov. *Int. J. Syst. Bacteriol.* 34: 31-34.
32. Schleifer K.H., Kraus J., Dvorak C., Kilpperbalz R., Collins M.D., Fischer W. (1985) Transfer of *Streptococcus lactis* and related *Streptococci* to the genus *Lactococcus* gen. nov. *Syst. Appl. Microbiol.* 6: 183-195.
33. Setlow P. (2006) Spores of *Bacillus subtilis*: their resistance to and killing by radiation, heat and chemicals. *J. Appl. Microbiol.* 101: 514-525.
34. Shi H., Shiraishi M., Shimizu K. (1997) Metabolic flux analysis of poly(b-hydroxybutyric acid) in *Alcaligenes eutrophus* from various carbon sources. *J. Ferment. Bioneng.* 84: 579-587.
35. Shida O., Takagi H., Kadowaki K., Nakamura L.K., Komagata K. (1996) Proposal for two new genera, *Brevibacillus* gen. nov. and *Aneurinibacillus* gen. nov. *Int. J. Syst. Bacteriol.* 46: 939-946.
36. Shida O., Takagi H., Kadowaki K., Nakamura L.K., Komagata K. (1997) Transfer of *Bacillus alginolyticus*, *Bacillus chondroitinus*, *Bacillus curdlanolyticus*, *Bacillus glucanolyticus*, *Bacillus kobensis*, and *Bacillus thiaminolyticus* to the genus *Paenibacillus* and emended description of the genus *Paenibacillus*. *Int. J. Syst. Bacteriol.* 47: 289-298.
37. Spring S., Ludwig W., Marquez M.C., Ventosa A., Schleifer K.H. (1996) *Halobacillus* gen. nov., with descriptions of *Halobacillus litoralis* sp. nov. and *Halobacillus trueperi* sp. nov., and transfer of *Sporosarcina halophila* to *Halobacillus halophilus* comb. nov. *Int. J. Syst. Bacteriol.* 46: 492-496.

38. Todar K. (2005) Todar's online textbook of bacteriology, The genus *Bacillus*, http://www.textbookofbacteriology.net/Bacillus.html
39. Tortora G.J., Funke B.R., Case C.L. (2007) Microbiology: an introduction, 9th ed. Pearson/Benjamin Cummings, San Francisco.
40. Vandamme P. and Coenye T. (2004) Taxonomy of the genus *Cupriavidus*: a tale of lost and found. *Int. J. Syst. Evol. Microbiol.* 54: 2285-2289.
41. Vandamme P., Pot B., Gillis M., De Vos P., Kersters K., Swings J. (1996) Polyphasic taxonomy, a consensus approach to bacterial systematics. *Microbiol. Rev.* 60: 407-438.
42. Vaneechoutte M., Kämpher P., De Baere T., Falsen E., Verschraegen G. (2004) *Wautersia* gen. nov., a novel genus accomodating the phylogenetic lineage including *Ralstonia eutropha* and related species, and proposal of *Ralstonia [Pseudomonas] syzygii* (Roberts et al. 1990) comb. nov. *Int. J. Syst. Evol. Microbiol.* 54: 317-327.
43. Versalovic J., Schneider M., de Bruijn F.J., and Lupski J.R. (1994) Genomic fingerprinting of bacteria using repetitive sequence based PCR (rep-PCR). *Meth. Mol. Cell. Biol.* 5: 25-40.
44. Versalovic J., Koeuth T., Lupski J.R. (1991) Distribution of repetitive DNA-sequences in Eubacteria and application to fingerprinting of bacterial genomes. Nucleic *Acids Res.* 19: 6823-6831.
45. Wayne L.G., Brenner D.J., Colwell R.R., Grimont P.A.D., Kandler O., Krichevsky M.I., Moore L.H., Moore W.E.C., Murray R.G.E., Stackebrandt E., Starr M.P., Trüper H.G. (1987) Report on the Ad Hoc committee on reconciliation of approaches to bacterial systematics. *Int. J. Syst. Bacteriol.* 37: 463-464.
46. Wigley P, Burton N.F. (2000) Multiple chromosomes in *Burkholderia cepacia* and *B. gladioli* and their distribution in clinical and environmental strains of *B. cepecia*. *J. Appl. Microbiol.* 88: 914-918.
47. Yabuuchi E., Kosako Y., Yano I., Hotta H., Nishiuchi Y. (1995) Transfer of 2 *Burkholderia* and an *Alcaligenes* species to *Ralstonia* gen. nov. - proposal of *Ralstonia picketti* (Raolston, Palleroni and Doudoroff 1973) comb. nov., *Ralstonia*

solanacearum (Smith1896) comb. nov. and *Ralstonia euthropha* (Davis 1969) comb. nov. *Microbiol. Immunol.* 39: 897-904.

48. Yoon J.H., Lee K.C., Weiss N., Kho Y.H., Kang K.H., Park Y.H. (2001) *Sporosarcina aquimarina* sp nov., a bacterium isolated from seawater in Korea, and transfer of *Bacillus globisporus* (Larkin and Stokes 1967), *Bacillus psychrophilus* (Nakamura 1984) and *Bacillus pasteurii* (Chester 1898) to the genus *Sporosarcina* as *Sporosarcina globispora* comb. nov., *Sporosarcina psychrophila* comb. nov and *Sporosarcina pasteurii* comb. nov., and emended description of the genus *Sporosarcina*. *Int. J. Syst. Evol. Microbiol.* 51: 1079-1086.

Chapter 3: Raman spectroscopy

This chapter presents some theoretical aspects of Raman spectroscopy. Not only the Raman effect is explained, but also considerations concerning the Raman band position in a spectrum, Raman band intensity, interferences and instrumental properties are discussed.

3.1 Introduction

Raman scattering or inelastic scattering was theoretically predicted by Smekal[6] in 1923 and observed for the first time using solar irradiation by Raman and Krishnan[5] in 1928. Raman spectroscopy as an analytical technique became important with the invention of lasers in the 1960's and the development of better instrumentation in the 1980's (detector, fiber optics, miniaturizing)[4]. This progress was of large importance as Raman spectroscopy is an inherently weak effect and thus sensitivity is an issue.

3.2 The Raman effect: quantum theory

The Raman effect[2] is a light scattering effect and can be described according to the energy diagram in Fig. 3.1. When monochromatic light interacts with a molecule, it can be excited to a virtual state. This molecule can relax in several ways. When it returns to its original vibrational state, it emits a photon of the same energy (or wavelength) as the incident light, which is called Rayleigh scattering or elastic scattering. Alternatively, the molecule can return to a higher vibrational state by scattering light with a lower energy (or larger wavelength) than the incident light, which is called Stokes scattering, a form of inelastic scattering. Inelastic scattering can also occur when the molecule relaxes towards a lower vibrational state by emitting a photon with higher energy than the incident light and is called anti-Stokes scattering. The shift in energy (or wavelength) between the incident light and the Stokes or anti-Stokes scattering gives information about the vibrational states of a molecule. The intensity of both inelastic scattering processes is proportional to the number of molecules that can undergo these processes. According to the Boltzmann distribution, the number of molecules at thermal equilibrium in a lower vibrational state is

always higher than in a higher vibrational state. Therefore, the Stokes intensity is usually higher than the anti-Stokes intensity and is therefore commonly used in Raman spectroscopy.

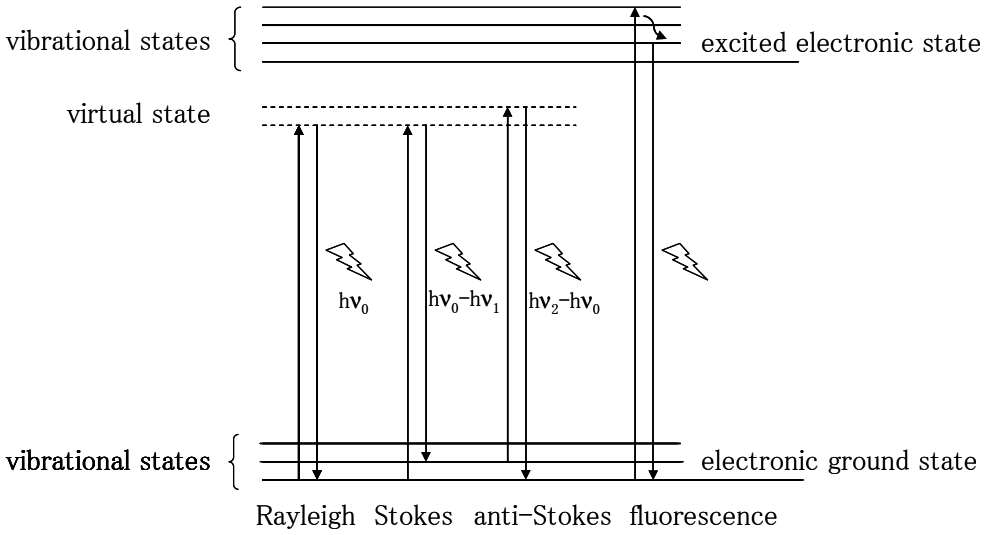


Figure 3.1 Energy diagram that illustrates different types of scattering when a molecule is illuminated with light of energy $h\nu_0$ and fluorescence[2].

3.3 The Raman effect: classical theory and selection rules

The classical theory of the Raman effect is based upon polarizability of molecules, which reflects how easy an electron cloud of a molecule can be distorted by an electric field (e.g. light)[4]. Energy from light can interact with the molecular vibrations and is released as scattered light. The scattered light can have a frequency equal to the incident light (Rayleigh), equal to the incident light minus the vibrational frequency (Stokes) and equal to the incident light plus the vibrational frequency (anti-Stokes). This theory predicts most aspects of Raman scattering correctly and predicts that the Raman intensity is not equal in all directions[4]. Raman scattering is usually observed at 90° or 180° relative to the direction of the incident light and these setups are called right-angle scattering or backscattering geometries, respectively. Because this theory does not take the quantization of vibrational levels into account, it incorrectly predicts that Stokes and anti-Stokes intensities are equal.

When molecules are considered as point masses connected by a spring that oscillates harmonically, the following equation is useful for deducing trends in frequencies of certain molecular vibrations[4]:

$$\omega = \frac{1}{2\pi c}\left(\frac{k}{\mu_r}\right)^{1/2} \tag{3.1}$$

Where k = force constant of the chemical bond

μ_r = reduced mass = $m_1 m_2/(m_1+m_2)$

m_1, m_2 = masses of the vibrating atoms

c = speed of light

w = wavenumber

Symmetry arguments allow determining which vibrations are Raman inactive, which reflect vibrational transitions that are quantum-chemically forbidden[4]. A vibration is Raman inactive when it does not induce a change in polarizability of the molecule[1].

3.4 Raman shift and intensity

A Raman spectrum is a plot of the Raman intensity versus the wavenumber or Raman shift in reciprocal centimetres. The relation between energy, wavelength, wavenumber and Raman shift* is as follows[4]:

$$E = h\nu = \frac{hc}{\lambda} = hc\omega \tag{3.2}$$

$$Raman\ shift(cm^{-1}) = \left(\frac{1}{\lambda_{inc}(nm)} - \frac{1}{\lambda_{Stokes}(nm)}\right)10^{-7} \tag{3.3}$$

Where E = energy of light

h = Planck's constant

ν = frequency of light

c = speed of light

* According to IUPAC, Raman wavenumber should be used instead of Raman shift. However, to stress the fact that a reciprocal difference between the wavelength of the laser light and the Stokes scattering is presented, we use the term Raman shift throughout the whole thesis.

λ = wavelength of light from incident light (inc) or Stokes scattering (Stokes)

ω = wavenumber of light

An expression for the intensity of Raman scattering[4] was derived by Plačzek:

$$I_R = \frac{2^4\ \pi^3}{45\times 3^2\times \mathrm{c}^4}\times\frac{\mathrm{h}\, I_L N(\nu_0-\nu)^4}{\mu_r\,\nu\,(1-e^{-\mathrm{h}\nu/\mathrm{k}T})}\left[45(\alpha_a')^2+7(\gamma_a')^2\right] \qquad (3.4)$$

Where I_L = intensity of the incident light

N = number of molecules that scatter the light

ν = molecular vibrational frequency (s^{-1})

ν_0 = frequency of the incident light (s^{-1})

k = Boltzmann constant

T = absolute temperature

α_a' = mean value for invariant of the polarizability tensor

γ_a' = anisotropy invariant of the polarizability tensor

From this equation, it can be deduced that the intensity of the scattered light is proportional to the number of molecules (basis for quantitative analysis), the intensity of the incident light and $(\nu_0 - \nu)^4$. Higher laser intensity and a higher laser frequency thus enhance the intensity of Raman scattering.

Some molecular factors that affect the intensity of a Raman band are:

- Raman bands of vibrations of polar bonds are generally weak, because their electron clouds are not easily distorted by light.
- The Raman intensity of stretching vibrations is generally stronger than that of bending vibrations.
- Double or triple bonds yield more intense stretching bands.
- The Raman intensity increases with the atomic number of the vibrating atoms.
- Symmetric vibrations generally cause more intense Raman bands.
- Crystalline materials cause stronger Raman bands.

3.5 Fluorescence

As the Raman effect in inherently weak, the suppression and/or subtraction of interferences and noise is important. Fluorescence is one of the most disturbing features in a Raman spectrum[2]. Visible light can excite molecules to a vibrational level in an excited electronic state (Fig. 3.1). After relaxation to a lower vibrational state in the excited electronic state, it can return to a vibrational state of the electronic ground state by emitting fluorescence. Fluorescence can be suppressed by choosing appropriate instrumental properties such as laser wavelength and spectrometer type (see below).

3.6 Advantages and disadvantages

Raman spectroscopy is a non-destructive and fast technique that requires only small sample volumes, that has a high spatial resolution and the sample preparation and water interference is minimal. A disadvantage is the inherent weakness of Raman scattering and the interference of fluorescence. To deal with these issues, the selection of appropriate instrumental properties is important.

Because fluorescence occurs mostly when exciting with visible light, the choice of a laser with a wavelength in near-IR region can be useful to suppress this effect[1]. Using UV excitation, a higher signal intensity can be obtained next to resonance effects, so fluorescence is less of a problem. However, the sensitivity is lower when using near-IR light, while UV radiation can damage the sample. The choice of the detection system is also important. A dispersive spectrometer (CCD detector) is more sensitive than a Fourier-transform spectrometer, but it does not allow high laser wavelengths[3]. The research described in this thesis was performed with a dispersive spectrometer to obtain a high sensitivity. The laser wavelength for this instrument was chosen as high as possible (785 nm, near-IR) to suppress fluorescence.

In this chapter, theoretical considerations concerning the Raman effect and some instrumental properties were discussed. The next chapter discusses the use of dataprocessing, this is the treatment of (raw) data to enhance information extraction.

References

1. Grasselli J.G. (1991) Analytical Raman spectroscopy. Chapter 1 The Raman effect: an introduction. John Wiley and Sons Ltd, New York, 1-19.
2. McCreery R.L. (2000) Raman spectroscopy for chemical analysis (Chemical analysis series vol. 157). Chapter 1 Introduction and scope. John Wiley and Sons, New York, 1-14.
3. McCreery R.L. (2000) Raman spectroscopy for chemical analysis (Chemical analysis series vol. 157). Chapter 5 Instrumentation overview and spectrometer performance. John Wiley and Sons, New York, 73-94.
4. Pelletier M.J. (1999) Analytical applications of Raman spectroscopy. Chapter 1 Introduction to applied Raman spectroscopy. Blackwell Science Ltd, Oxford, 1-52
5. Raman C.V., Krishnan K.S. (1928) A new type of secondary radiation. *Nature* 121: 501-502.
6. Smekal A. (1923) Zur quantentheorie der dispersion. *Naturwiss.* 11: 873-875.

Chapter 4: Data processing

Chapter 3 described the Raman effect and some instrumental properties. In order to be able to extract information from the collected Raman spectra, appropriate (pre)processing is necessary. This chapter presents several approaches that are used in this work.

4.1 Introduction

Bacterial Raman spectra are complex and the differences between them are often very small. It is important that an accurate calibration and further processing is performed, in order to be able to extract information from these spectra. The calibration procedure used for this work is described by Hutsebaut *et al.*[5]. It includes absolute wavelength calibration with a neon lamp, intensity calibration with a tungsten bulb, relative wavelength calibration with seven reference products, corrections for dark noise and signals from optical compounds. In addition, a standardization procedure is performed to eliminate variations between the calibration and the measurement session. Indeed, absolute shifts can be caused by changes in the optical pathway while shifts in laser wavelength can occur for instance when switching the laser on or off. Since several neon and ε-caprolactone spectra are collected during each measurement session, these shifts can be eliminated by shifting the collected spectra to a reference wavelength axis and a reference wavenumber axis.

Further pre-processing steps are:

- Background moddeling with EMSC
- Deleting spikes
- Calculation of average spectra of the four spectra that were recorded around a focus point

Other processing methods that are used in this work are:

- PCA analysis (always with precedent auto scaling)
- Subtraction of the background (if band intensities need to be calculated)

- Calculation of dot products between reference spectra and bacterial spectra
- Use of coefficients from an EMSC model

Some of these methods are discussed in detail below.

4.2 EMSC

Extended multiplicative signal correction (EMSC) was developed by Martens and Stark[8] and is a "soft modelling" method, which implies the use of empirical measurements and as few as possible statistical and causal assumptions. This method can separate physical and chemical effects in light spectroscopy. One of the purposes is to exclude additive (e.g. absorbance effects of chemical interferences) and multiplicative effects (light scattering variation, optical path length variation) from the data. The model used for these purposes is (matrices are printed in bold):

$$\mathbf{z}_i = \mathbf{x}_i b_i + \mathbf{1} a_i + \mathbf{e}_i \qquad (4.1)$$

Where $\mathbf{Z} = \{\mathbf{z}_i = 1, 2, \ldots, I\}$ = the measured spectra of I samples ($\mathbf{z}_i$ column vector)

$\mathbf{X} = \{\mathbf{x}_i = 1, 2, \ldots, I\}$ = the corrected spectra of I samples ($\mathbf{x}_i$ column vector)

$\mathbf{1} = (1, 1, \ldots, 1)'$

a_i = unknown additive effect

b_i = unknown multiplicative effect

$\mathbf{e}_i$ = the residual in the model

Under ideal conditions (Beer's law) the data $\mathbf{x}_i$ can be seen as a sum of contributions of different compounds:

$$\mathbf{x}_i = \mathbf{k}_1 c_{i1} + \mathbf{k}_2 c_{i2} + \ldots + \mathbf{k}_j c_{ij} = \mathbf{K}\mathbf{c}_i \qquad (4.2)$$

Where $\mathbf{K} = \{\mathbf{k}_j, j=1, 2, \ldots, J\}$ = reference spectra of single compounds ($\mathbf{k}_j$ column vector)

$\mathbf{c}_i = \{c_{ij}, j=1, 2, \ldots, J\}$ = the concentration of component j in sample i

The corrected spectra $\mathbf{x}_i$ can be written as a deviation from a mal spectrum $\mathbf{m}$, which is in this study the average spectrum of the dataset. This mal spectrum represents the general shape of a spectrum of the dataset and all spectra will be modelled towards this mal spectrum.

$$\mathbf{x}_i = \mathbf{m} + \mathbf{K}\mathbf{d}_i \tag{4.3}$$

Where $\mathbf{d}_i = \mathbf{c}_i - \mathbf{1}c_0$ = the compounds' concentrations compared to the mal spectrum

Substitution of $\mathbf{x}_i$ in equation 4.1 yields:

$$\mathbf{z}_i = \mathbf{m}b_i + \mathbf{K}\mathbf{d}_i b_i + \mathbf{1}a_i + \mathbf{e}_i \tag{4.4}$$

In the case of conventional multiplicative scattering correction the term $\mathbf{K}\mathbf{d}_i b_i$ is ignored and the parameters a_i and b_i are estimated by the least squares solution:

$$[b_i, a_i] = ([\mathbf{m1}]'[\mathbf{m1}])^{-1}[\mathbf{m1}]'\mathbf{z}_i \tag{4.5}$$

In the case of extended multiplicative scattering correction $\mathbf{d}_i b_i$ is termed $\Delta\mathbf{c}_i$ from which the values $\Delta c_{i,j}$ reflect the difference in the concentration of compound j between the i^{th} sample spectrum and the mal spectrum. The least squares solution is:

$$[b_i, \Delta\mathbf{c}_i, a_i] = ([\mathbf{mK1}]'[\mathbf{mK1}])^{-1}[\mathbf{mK1}]'\mathbf{z}_i \tag{4.6}$$

For both cases, the corrected spectra $\mathbf{x}_i$ can be calculated as follows:

$$\mathbf{x}_i = (\mathbf{m} - \mathbf{1}a_i)/b_i \tag{4.7}$$

In the EMSC toolbox by Martens[7], these procedures are called 'Datacase 103' and 'Datacase 106'. In this study 'Datacase 103' is used to correct a dataset for physical interferences (additive and multiplicative effects (4.5)). When reference spectra of single compounds are included in 'Datacase 106' ($\mathbf{K}\mathbf{d}_i b_i$ term (4.4)), the resulting coefficients $\Delta c_{i,j}$ were used as relative values for the contribution of reference spectrum j in bacterial Raman spectrum i.

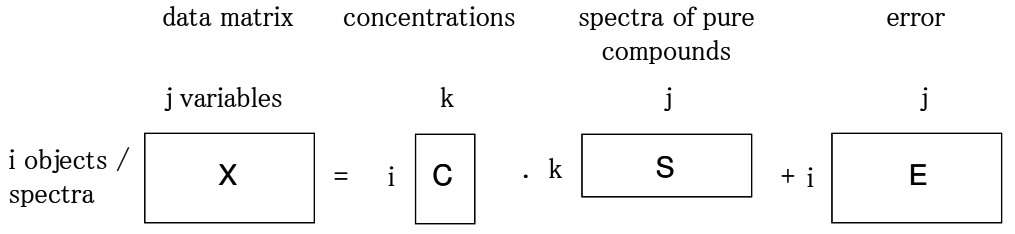
4.3 Modified polynomial fit

Modified polynomial curve fitting was presented in 2003 by Lieber and Mahadevan-Jansen[6] as a method for background subtraction (mainly fluorescence) in biological Raman spectra. When a polynomial curve is fitted to a spectrum (least squares fit), it will not represent the background properly, as it was fitted to a spectrum consisting of background and Raman bands. To eliminate bands from this fit, a modified polynomial fit is generated which includes for each Raman shift the minimal value of the spectrum and the polynomial fit. This first polynomial fit is then again subjected to this procedure, etc. In this work, 150 iterations were performed and the final modified polynomial fit was subtracted from the original spectrum (illustrated in Chapter 9).

4.4 PCA

Principal component analysis (PCA) is a multivariate statistical technique[1,2]. It allows unsupervised pattern recognition which aims to detect similarities and for which no a priori knowledge is used. PCA is a data reducing method and aims to determine the underlying information from multivariate raw data by: (i) interpreting the principal components (scores related to objects and loadings related to variables) and (ii) observing patterns.

The method can be explained as follows. A dataset, for example a set of spectra **X** can be written as a combination of signals from the constituents and some noise.



(4.8)

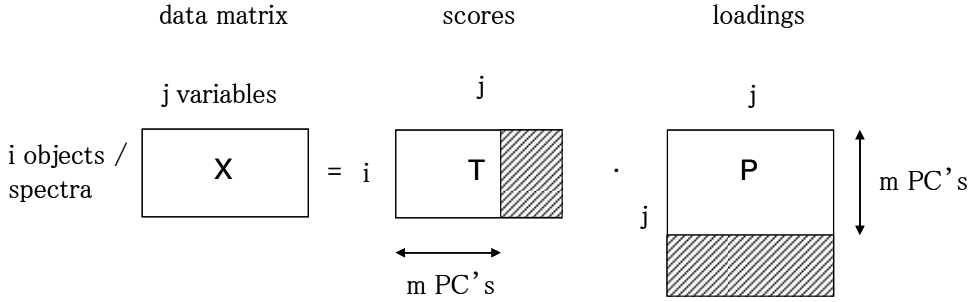
Where **X** = data matrix (objects in rows, variables in columns)

C = concentrations of each constituent in each object

S = spectra of each constituent

E = the error

C and **S** could be predicted, but can never be observed directly and perfectly. In order to observe the most important features and trends in the dataset, it can be reduced by PCA. The data matrix is then decomposed in two matrices:



(4.9)

Where **T** = the scores

P = the loadings

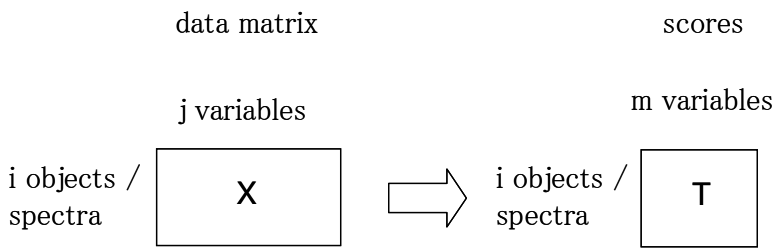
Scores and loadings vectors are orthogonally, which means that the product between any two loadings and scores vectors are zero. If the columns are mean centred (autoscaling, see below), the correlation between any two scores or loadings vectors is also zero. As our original dataset contained j variables, in theory the scores matrix also contains j components. Each component contains a certain percentage of the information (variability) of the dataset, which is expressed by the eigenvalue of the component λ_j:

$$\lambda_j = \frac{S_j}{S_{total}} = \frac{\sum_{i=1}^{I} t_{ij}^2}{\sum_{i=1}^{I}\sum_{j=1}^{J} x_{ij}^2} \tag{4.10}$$

Where S_j = the sum of squares of the jth PC (the jth score vector)

S_{total} = the sum of squares of the entire dataset

The first component has the highest eigenvalue and the eigenvalues further decrease with increasing number of the components. The dataset can now be reduced by retaining m principal components, where m is mostly determined by studying the eigenvalues. In this way, PCA has reduced the data set (i,j) to a scores matrix (i,m):



(4.11)

In order to present our data visually, a scores plot can be constructed of the two principal components that contain the most valuable information. The scores plot shows possible groupings between objects (here spectra). Mostly, the first two principal components are used. Analogously, a loadings plot can be constructed of these two principal components, which shows the relation between the original variables. From the comparison between the scores and loadings plots, it can be deduced which variables (here Raman bands) are characteristic for certain objects and which variables are responsible for the differentiation between them. Indeed a variable with a high loading for PC1, is an important feature in the objects with a high score for PC1 and vice versa. In this way, (groups of) objects/spectra can be related to certain variables/Raman bands, so information about their relative composition can be obtained (as applied in Chapter 6).

In this work, spectra are always autoscaled before performing PCA. Autoscaling is in fact mean centring of the columns (variables/Raman shifts) of the data matrix **x**(i,j)[3]:

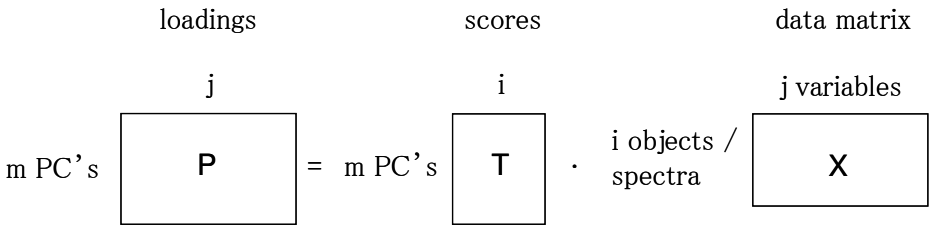
$$auto(x_{ij}) = \frac{x_{ij} - \overline{x_j}}{\sqrt{\sum_{i=1}^{I} (x_{ij} - \overline{x_j})^2 / I}} \qquad (4.12)$$

Where $\overline{x_j}$ = average value of column j

I = number of rows/objects/spectra

In this way, minor differences in certain bands are taken into account to the same extent as larger band differences when performing PCA.

Equation 4.9 can be rewritten as follows:



(4.13)

In this way, it is clear that the m loading vectors in **P** are linear combinations of the original spectra. In chapter 8, loading vectors were used to replace an original dataset. A set of very similar fatty acid spectra caused colinearity problems in the EMSC procedure and was therefore replaced by a limited number (first three principal components) of orthogonal loadings vectors, called mix spectra. After applying the intended EMSC procedure, the information about the mix spectra can be transferred to information about the original objects by studying the scores.

4.5 Calculation of dot products

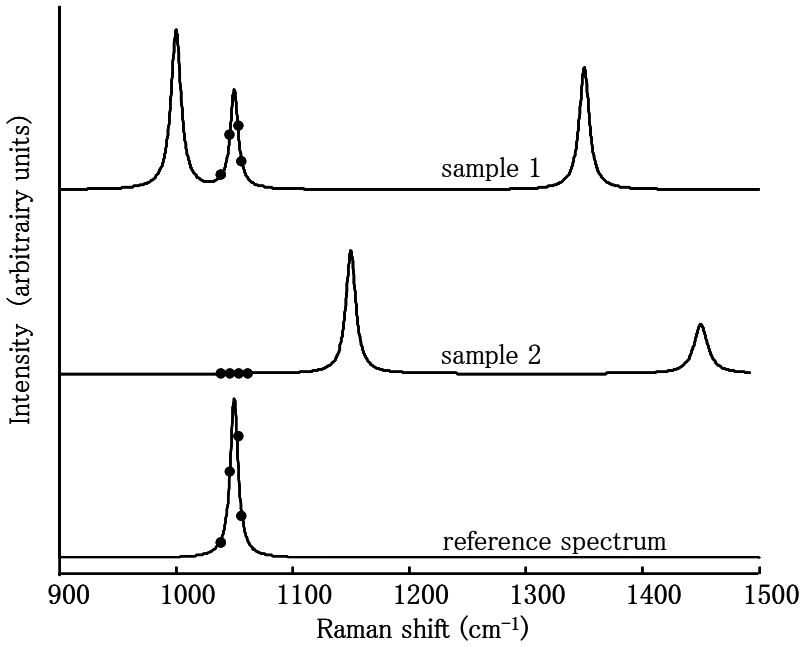
A dot product, also called scalar product or inner product, of two vectors $\mathbf{a} = [a_1, a_2, \ldots, a_n]$ and $\mathbf{b} = [b_1, b_2, \ldots, b_n]$ returns a number. It can be presented in different ways and is defined by[4]:

$$\mathbf{a} \bullet \mathbf{b} = \sum_{i=1}^{n} a_i b_i = a_1 b_1 + a_2 b_2 + \ldots + a_n b_n \tag{4.14}$$

The dot product can also be written as a matrix multiplication by transposing vector **a** to a column vector:

$$\mathbf{a} \bullet \mathbf{b} = \mathbf{a}^T \mathbf{b} \tag{4.15}$$

In this study, dot products were calculated between reference spectra and bacterial Raman spectra. Per pixel (Raman shift) the intensities of the reference spectrum and the bacterial spectrum are multiplied and these values are summed. This is illustrated in Fig. 4.1 with a simplified example that considers a reference spectrum with one band (dots represent pixels). In sample spectrum 2, the intensity equals almost zero at the Raman shifts where the reference spectrum shows a band, so their dot product is approximately zero. In sample spectrum 1, the products between the dotted pixels of each spectrum is significant and the dot product is high. The higher the intensities of the bacterial spectrum are at the Raman shifts where the reference spectrum shows bands, the higher the resulting dot product is. This shows that dot products are values for the correlation between these spectra and therefore resemble the contribution of the reference spectrum to the bacterial spectrum.



Figuur 4.1 Illustration of the dot product between reference and sample Raman spectra

In chapter 2, 3 and 4 some theoretical considerations were presented concerning microbiology, Raman spectroscopy and data processing, respectively. These chapters contain the information necessary to comprehend the results.

References

1. Brereton R.G. (2007) Applied chemometrics for scientists. Chapter 5 Pattern recognition. John Wiley & Sons Ltd., Chichester, 145-190.
2. Brereton R.G. (2003) Chemometrics: Data analysis for the laboratory and chemical plant. Chapter 3 Pattern recognition. John Wiley & Sons Ltd., Chichester, 183-255.
3. Brereton R.G. (2003) Chemometrics: Data analysis for the laboratory and chemical plant. Chapter 6 Evolutionary signals. John Wiley & Sons Ltd., Chichester, 339-398.
4. Brereton R.G. (2003) Chemometrics: Data analysis for the laboratory and chemical plant. Appendix A1. John Wiley & Sons Ltd., Chichester, 409-412.
5. Hutsebaut D., Vandenabeele P., Moens L. (2005) Evaluation of an accurate calibration and spectral standardization procedure for Raman spectroscopy. Analyst 130: 1204-1214.
6. Lieber C.A., Mahadevan-Jansen A. (2003) Automated method for subtraction of fluorescence from biological Raman spectra. Appl. Spectrosc. 57: 1363-1367.
7. Martens H., The EMSC toolbox for MATLAB
   http://www.models.kvl.dk/source/emsctoolbox
8. Martens H., Stark E. (1991) Extended multiplicative scattering correction and spectral interference subtraction: new preprocessing methods for near infrared spectroscopy. J. Pharm. Biomed. Anal. 9: 625-635.

Chapter 5: Reference database of Raman spectra of biological molecules

Joke De Gelder, Kris De Gussem, Peter Vandenabeele, Luc Moens

Journal of Raman spectroscopy, 38 (2007) 1133-1147

In literature, only few reference Raman spectra of sufficient quality of biological molecules are available. The need for such reference spectra is obvious in order to analyze Raman spectra of biological materials. Therefore, a database of Raman spectra of biological molecules was constructed and is presented in this chapter.

5.1 Introduction

Raman spectroscopy is a non-destructive, fast technique of analysis that requires only small sample volumes and nearly no interference of water occurs. The bands in Raman spectra are well resolved compared to similar spectroscopic techniques, such as infrared spectroscopy. Therefore, it finds its application in a lot of research domains in chemistry, such as art analysis[20], analysis of historical objects[4], catalyst research[18], etc. Over the past few years, its possibilities in biology and microbiology have been explored. As mentioned by Carey *et al.*[1], improvements have been made concerning problems such as low sensitivity, fluorescence rejection and data treatment, and thus the application of Raman spectroscopy became possible in these research fields.

Part of the research performed in these domains focuses on the identification of organisms such as bacteria, fungi and yeasts. Several of these studies have been successful as identification at the species[11,12], sometimes even at the strain[7,8,17] level can be obtained. For this purpose, spectra are treated as mathematical data and only little or no attention is given to the interpretation of the spectra. Despite a lot of (bio)chemical information is present in these spectra, it is hard to retrieve it due to the complex nature of Raman spectra of biological materials. Indeed, it is advantageous that the spectra contain contributions of all (complex) biomolecules from whole cells, but it is very hard to analyze them. Several research groups worked on the interpretation of Raman spectra of biological samples. For example, some Raman bands in bacterial spectra have already been assigned to simple biomolecules and groups of molecules, such as nucleic acids and proteins[13]. In mycology, De Gussem *et al.*[3] analysed *Lactarius* spore composition and Edwards *et al.*[5] studied lichens. Despite the progress in this field, there is still a lot of information included

in these spectra that remains unused. Some authors have tried to avoid the complexity of the spectra by enhancing the contribution of specific molecules, by using surface enhanced Raman spectroscopy (SERS) or resonance Raman (RR) spectroscopy[2,23]. This is an interesting approach when only one or a limited number of biomolecules is of interest. However, when an overview of the entire cell composition is needed, one has to rely to general Raman spectroscopy. A first step towards the interpretation of these complex spectra is the knowledge of the Raman band positions caused by the cell components. Lippert *et al.*[10] and Weng *et al.*[22] have already presented Raman spectra of some fatty acids and fats. Surface enhanced Raman spectra of saccharides were recorded by Mrozek *et al.*[16], while Raman spectra of different types of carotenoids were described by Weesie *et al.* [21].
In this paper, we give an overview of Raman spectra of biomolecules. These spectra can serve as reference spectra for the interpretation of Raman spectra from biological materials. This database contains the most important building blocks of biomolecules present in a cell. It provides the basic information necessary to check the presence of biomolecules in biological materials, to monitor metabolism through time or to track changes induced by the environment. Finally, this approach is illustrated by the assignment of several bands in Raman spectra of bacteria and fungal spores.

5.2 Experimental

5.2.1 Raman instrumentation

Raman spectra were recorded with a Kaiser System Hololab 5000R modular Raman microspectrometer. A diode laser of 785 nm (Toptica Photonics AG) was focused through an 100x objective of the microscope (Leica), which resulted in a laser power of approximately 45 mW on the sample. The scattered light is transferred through a confocal 15 μm aperture pinhole and a collection fiber to the spectrograph, where it was detected by a back illuminated deep depletion Pelletier cooled (-70 °C) CCD detector (Andor). The Raman signal was collected in the spectral interval of 150 cm^{-1} until 3500 cm^{-1}, but only the

region from 365 till 1800 cm^{-1} was used. The acquisition time for all spectra was 60 s and the spectral resolution was approximately 4 cm^{-1}.

Spectral calibration was performed in MATLAB as described by Hutsebaut *et al.*[6], which includes (i) absolute wavelength calibration with a neon lamp, (ii) intensity calibration with a tungsten lamp, (iii) relative wavelength calibration with seven reference products, (iv) correction for dark noise of the detector as well as (v) correction for the spectral contribution of optical components.

5.2.2 Reference products

Table 5.1 lists all 61 reference products along with their supplier and technical information. The products were divided into 6 groups: DNA and RNA bases, amino acids, fatty acids and fats, saccharides, primary metabolites and others. For analysis, the reference products were transferred to a CaF_2 slide. For each reference product different focus points (2 to 4) were chosen, where each five Raman spectra of 60 s were collected. These calibrated spectra were treated by extended multiplicative signal correction (EMSC)[15], using the EMSC toolbox developed by H. Martens[14]. In particularly, datacase 103 was applied which represents a correction for physical interferences. Afterwards, the spikes were removed, followed by the calculation of the mean spectrum for each reference product.

Table 5.1 List of the reference products along with their supplier and technical information

Name	Supplier	Specifications
DNA and RNA bases		
adenine	Acros organics	99%
cytosine	Sigma-Aldrich	
guanine	Sigma-Aldrich	98%
thymine	Sigma-Aldrich	
uracil	Sigma-Aldrich	crystalline
Amino acids		
glycine	Sigma-Aldrich	

L-alanine	Sigma-Aldrich	Reagent grade, 98% (TLC)
L-valine	Sigma-Aldrich	Reagent grade, 98% (TLC)
L-serine	Sigma-Aldrich	ReagentPlus®, 99% (TLC)
L-glutamate (monosodium salt hydrate)	Sigma-Aldrich	≥99% (TLC), powder
L-arginine	Sigma-Aldrich	Free base
L-phenylalanine	Sigma-Aldrich	
L-tyrosine	Merck	For biochemistry
L-tryptophan	Sigma-Aldrich	Reagent grade, ≥98%
L-histidine	Sigma-Aldrich	ReagentPlus®, 99% (TLC)
L-proline	Sigma-Aldrich	Biochimika, 99% (NT)
Fatty acids and fats		
lauric acid	Sigma-Aldrich	98%
myristic acid	Sigma-Aldrich	Sigma grade, 99-100%
palmitic acid	Sigma-Aldrich	Free acid, Sigma grade
stearic acid	Sigma-Aldrich	Free acid, grade I, ~99%
12-methyltetradecanoic acid	Sigma-Aldrich	
13-methylmyristic acid	Sigma-Aldrich	
14-methylpentadecanoic acid	Sigma-Aldrich	
14-methylhexadecanoic acid	Sigma-Aldrich	
15-methylpalmitic acid	Sigma-Aldrich	
oleic acid	Sigma-Aldrich	Reagent grade, 99%
cis-vaccenic acid	Sigma-Aldrich	
glycerol	Sigma-Aldrich	*BioChemika Ultra*, anhydrous, ≥99.5% (GC)
triolein	Sigma-Aldrich	(C18:1,-CIS-9) Sigma grade approx 99%
trilinolein	Sigma-Aldrich	(C18:2,-CIS,CIS-9,12) approx 99%
trilinolenin	Sigma-Aldrich	Purum ~98% (GC)
saccharides		
β-D-glucose	MP Biomedicals	99%
lactose	DMV International (Veghel, The Netherlands)	Respitose SV003
cellulose	MP Biomedicals	98%
D(+)-dextrose	MP Biomedicals	anhydrous
D(+)-trehalose	Sigma-Aldrich	Biochimika ~99.5% (HPLC)
amylose	MP Biomedicals	Essential free of amylopectin
amylopectin	MP Biomedicals	
D(+)-mannose	Sigma-Aldrich	For microbiology
D(+)-fucose	Merck	
D(-)-arabinose	Merck	

D(+)-xylose	Merck	
D(-)-fructose	Sigma-Aldrich	99%
D(+)-galactosamine hydrochloride	Sigma-Aldrich	Cell culture tested
N-acetyl-D-glucosamine	Sigma Aldrich	≥ 99%, powder
chitin	MP Biomedicals	Practical grade 90-95%
Primary metabolites		
citric acid	Sigma-Aldrich	Anhydrous, 99%
succinic acid	Merck	
fumarate (sodium salt)	Merck	
malic acid	Merck	
pyruvate (sodium salt)	Sigma-Aldrich	Reagent plus TM, ≥99%
phosphoenolpyruvate	Boehringen Mannheim	
coenzyme A	Sigma-Aldrich	Free acid from yeast, hydrate
acetyl coenzyme A	Sigma-Aldrich	
acetoacetic acid (lithium salt)	Sigma-Aldrich	
D-fructose-6-phosphate (disodium salt)	Sigma-Aldrich	
Others		
β-carotene	Sigma-Aldrich	Type II, synthetic, ≥95% (HPLC), crystalline
ascorbic acid	Certa	Ph. Eur.
riboflavin (sodium phosphate)	Certa	
glutathione (reduced form)	Boehringen Mannheim	

5.3 Results and discussion

Figures 5.1 to 5.12 show Raman spectra of DNA and RNA bases, amino acids, fatty acids and fats, saccharides, primary metabolites and others. The spectra of these groups can clearly be distinguished from each other. The spectra of DNA and RNA bases are dominated by ring breathing vibrations in the region from 600 until 800 cm^{-1} (Fig. 5.1). In spectra of cells, not only the vibrations of these bases but also symmetric >(PO_2^-) stretches of the DNA and RNA backbone will be visible at ca. 1100 cm^{-1}. In general, the presence of proteins can be noticed by amide I and amide III bands at about 1300 and 1655 cm^{-1}. Raman bands of specific amino acids (Fig. 5.2 and 5.3) are discussed below. Raman spectra

of pure fatty acids and fats (Fig. 5.4, 5.5 and 5.6) were collected and can be distinguished from the other groups by typical bands at about 1300 and 1440 cm^{-1}. Spectra of saccharides (Fig. 5.7, 5.8 and 5.9) are characterized by groups of bands in the regions from 1000 until 1200 cm^{-1} and 1300 until 1500 cm^{-1}. The spectra of the primary metabolites (Fig. 5.10 and 5.11) are mutually very different.

5.3.1 DNA and RNA bases

The Raman spectra of the five bases present in DNA and RNA are shown in Figure 5.1. Each spectrum contains an intense band in the region from 600 until 800 cm^{-1} that can be assigned to ring breathing vibrations. The ν(C=O) stretching vibrations cause only for thymine an intense band (1671 cm^{-1}). In the spectra of all five bases almost no similarities can be noticed, so it is easy to distinguish between them. Even T and U do not show similarities while their chemical structure only differs in the substitution of a methyl group at the C_5 position (see appendix). These five bases can be divided into purine bases (adenine and guanine) and pyrimidine bases (cytosine, thymine and uracil), but there are no bands observed that are characteristic for each group, as reported by Socrates *et al.*[19]

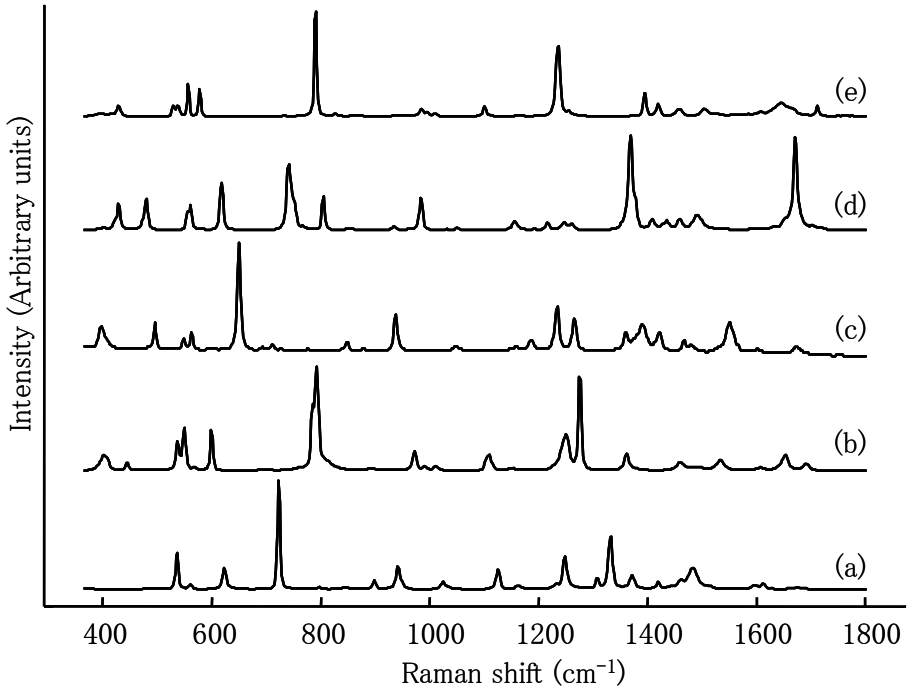


Figure 5.1 Raman spectra of DNA and RNA bases: (a) adenine, (b) cytosine, (c) guanine, (d) thymine, (e) uracil

Table 5.2 Raman bands and intensities for each reference product (very weak bands are not listed)

Product	Raman bands	Fig
adenine	536(m), 560(w), 623(m), 723(s), 898(w), 941(m), 1024(w), 1125(m), 1162(w), 1134(w), 1248(m), 1307(w), 1332(m), 1371(w), 1419(w), 1462(w), 1482(m), 1612(w)	5.1a
cytosine	402(mw), 444(w), 537(m), 548(m), 568(w), 599(m), 792(s), 971(mw), 991(w), 1011(w), 1108(mw), 1250(m), 1275(s), 1362(mw), 1460(w), 1493(w), 1533(w), 1653(mw), 1690(w)	5.1b
guanine	397(m), 496(m), 548(w), 563(m), 650(s), 711(w), 849(w), 937(m), 1048(w), 1158(w), 1187(w), 1234(m), 1266(m), 1360(m), 1391(m), 1422(m), 1468(w), 1479(w), 1550(m), 1602(w), 1674(w)	5.1c
thymine	429(m), 479(m), 561(m), 617(s), 740(s), 767(w,sh), 804(m), 984(m), 1156(w), 1216(w), 1247(w), 1261(w), 1369(vs), 1408(w), 1435(w), 1459(w), 1490(m), 1655(w,sh), 1671(vs), 1702(w)	5.1d
uracil	429(w), 529(w), 537(w), 556(m), 577(m), 790(vs), 984(w), 995(w), 1100(w), 1235(s), 1255(w), 1394(m), 1418(m), 1455(w), 1504(w), 1608(w), 1646(m), 1711(w)	5.1e
glycine	504(s), 560(w), 606(m), 686(m), 894(vs), 932(w), 1049(m), 1129(m), 1156(w), 1323(m), 1338(m), 1346(m), 1398(m), 1441(m), 1499(w), 1589(w), 1627(w), 1634(w), 1674(w)	5.2a
L-alanine	398(ms), 480(w), 532(ms), 652(mw), 773(m), 851(vs), 920(mw), 1021(m), 1112(m), 1146(m), 1236(w), 1305(s), 1359(vs), 1378(mw), 1409(m), 1462(s), 1482(s), 1499(m), 1596(m)	5.2b
L-valine	374(mw), 396(mw), 429(mw), 472(w), 497(w), 542(vs), 664(mw), 715(m), 753(m), 776(s), 824(m), 849(s), 891(w,sh), 902(m), 923(mw,sh), 948(s), 964(m), 1029(mw), 1035(mw), 1066(mw), 1106(w), 1125(m), 1146(mw), 1179(mw), 1191(m), 1272(m), 1321(m,sh), 1330(m), 1343(s,sh), 1351(s), 1398(m), 1427(m), 1452(s), 1508(mw), 1567(w), 1587(w), 1619(mw), 1633(mw), 1660(w)	5.2c
L-serine	515(m), 611(m), 813(m), 853(s), 910(w,sh), 922(w), 968(m), 1008(m), 1090(w), 1134(mw), 1220(mw), 1299(m), 1326(s), 1368(w), 1384(w), 1416(mw), 1426(mw,sh), 1462(m), 1599(w), 1613(w), 1628(w)	5.2d
L-glutamate	382(mw), 429(mw), 477(mw), 499(m), 529(m), 604(mw), 633(w), 663(mw), 740(w), 775(m), 792(mw), 808(mw), 857(vs), 876(s), 925(m), 942(vs), 1003(mw), 1042(mw), 1056(mw), 1075(mw), 1095(m), 1123(mw), 1142(m), 1160(mw), 1192(w), 1258(w), 1282(m), 1293(m), 1317(s), 1341(s), 1353(m), 1401(vs), 1421(s), 1434(s), 1517(w), 1536(w), 1570(w), 1605(mw), 1624(mw), 1641(w), 1683(w)	5.2e
L-arginine	376(w), 410(w), 490(mw), 551(m), 577(m), 613(m), 849(m), 879(mw), 922(m), 982(vs), 1036(mw), 1067(s), 1100(m), 1122(m), 1189(m), 1264(mw), 1298(m), 1330(m), 1377(mw), 1436(s), 1475(m), 1713(w)	5.2f
L-phenylalanine	361(w), 468(m), 527(w), 605(w), 622(m), 747(mw), 832(m), 852(m), 951(w), 1004(vs), 1037(m), 1157(w), 1183(w), 1216(m), 1308(w), 1336(w), 1353(w), 1447(w), 1586(m), 1602(m)	5.3a

Product	Raman bands	Fig
L-tyrosine	379(w), 431(m), 491(w), 527(w), 641(m), 714(w), 797(mw), 828(vs), 845(m), 983(m), 1044(w), 1177(ms), 1200(m), 1214(m), 1247(w), 1265(m), 1326(m)	5.3b
L-tryptophan	393(w), 425(w), 456(w), 498(m), 509(m), 534(m), 548(w), 574(m), 596(m), 626(m), 683(w), 706(m), 741(m), 755(vs), 766(m), 778(m), 802(w), 840(m), 848(m), 865(m), 874(s), 988(w), 1009(vs), 1046(w), 1076(w), 1103(w), 1118(m), 1160(w), 1207(w), 1231(m), 1253(w), 1278(w), 1309(w), 1314(m), 1328(m), 1338(s), 1358(s), 1423(s), 1450(m), 1457(m), 1486(m), 1556(s), 1576(m), 1616(m)	5.3c
L-histidine	404(m), 422(mw,sh), 540(mw), 623(mw), 656(m), 680(w), 731(mw), 784(mw), 804(m), 824(mw), 852(m), 918(m), 929(mw,sh), 963(m), 976(m), 1061(m), 1087(s), 1111(m), 1140(mw), 1174(m), 1224(m), 1250(m), 1271(s), 1317(vs), 1336(m), 1347(m), 1407(m), 1430(m), 1476(mw), 1498(m), 1538(w), 1571(m), 1608(w), 1639(w)	5.3d
L-proline	374(mw), 452(vs), 563(w), 577(w), 642(m), 681(mw), 774(w), 793(mw), 834(vs), 842(vs,sh), 850(s,sh), 877(m), 899(vs), 916(vs), 930(m), 947(m), 951(m,sh), 987(mw,sh), 994(m), 1035(m), 1045(m), 1083(m), 1166(mw), 1175(mw), 1194(mw), 1216(w), 1240(m), 1267(m), 1286(m), 1317(mw), 1333(mw), 1350(mw), 1357(mw), 1378(m), 1389(m), 1410(mw), 1418(mw), 1434(w,sh), 1443(mw,sh), 1454(s), 1479(mw), 1547(w), 1605(w)	5.3e
lauric acid	464(w), 631(w), 664(w), 891(m), 906(m), 1018(w), 1038(w), 1061(s), 1078(m,sh), 1084(m), 1127(vs), 1178(w), 1191(w), 1272(w,sh), 1296(vs), 1369(w), 1408(m), 1446(vs), 1459(s), 1490(s), 1497(w), 1651(w)	5.4a
myristic acid	414(m), 421(w,sh), 670(w), 893(m), 908(m), 986(w), 1039(w), 1063(vs), 1092(m), 1128(vs), 1175(w), 1296(vs), 1371(w), 1412(m), 1422(s), 1438(s), 1454(m), 1507(w), 1627(w), 1650(w)	5.4b
palmitic acid	375(m), 433(w), 670(w), 893(m), 909(m), 1063(vs), 1099(m), 1129(vs), 1174(w), 1296(vs), 1371(w), 1421(s), 1438(vs), 1455(m), 1465(s), 1481(w), 1627(w)	5.4c
stearic acid	818(w), 848(m), 866(m), 890(w), 909(w), 1062(s), 1100(m), 1129(s), 1173(m), 1187(w), 1206(w), 1296(s), 1371(w), 1408(m), 1418(m), 1441(vs), 1463(s), 1474(m), 1498(w), 1648(w)	5.4d
12-methyltetradecanoic acid (15Aiso)	433(w), 669(w), 772(w), 818(m), 840(w), 905(m), 983(w), 990(w), 1018(w), 1040(w), 1063(s), 1098(s), 1130(w,sh), 1141(s), 1171(w), 1265(w), 1296(s), 1302(m), 1344(w), 1371(w), 1380(w), 1412(m), 1440(s), 1447(s), 1459(s), 1472(s), 1505(w), 1649(w)	5.5a
13-methylmyristic acid (15iso)	435(w), 822(m), 907(m), 917(w), 936(w), 1061(s), 1098(m), 1134(s), 1170(w), 1272(w), 1296(vs), 1371(w), 1410(w), 1447(vs), 1469(s), 1492(w), 1649(w)	5.5b

Product	Raman bands	Fig
14-methylpentadecanoic acid (16iso)	418(m), 434(w), 666(w), 815(m), 831(w), 904(m), 918(w), 940(w), 952(w), 1048(w,sh), 1061(s), 1102(m), 1135(s), 1170(w), 1244(w), 1270(w,sh), 1284(m), 1296(s), 1337(w), 1370(w), 1412(m), 1445(s), 1465(s), 1651(w)	5.5c
14-methylhexadecanoic acid (17Aiso)	372(m), 452(w), 667(w), 771(m), 788(w), 815(m), 834(m), 906(m), 969(w), 983(w), 995(w), 1006(w), 1031(m), 1062(vs), 1104(vs), 1140(s), 1170(w), 1186(w), 1253(w), 1296(s), 1344(m), 1370(m), 1411(m), 1441(vs), 1465(w), 1479(m), 1649(m)	5.5d
15-methylpalmitic acid (17iso)	405(m), 416(m), 467(w), 516(w), 666(w), 801(w), 820(s), 846(w), 905(m), 910(m), 930(w), 952(w), 1010(w), 1050(w,sh), 1061(vs), 1103(s), 1135(vs), 1171(m), 1188(w), 1259(w), 1296(vs), 1306(s), 1323(w), 1340(w), 1369(w), 1386(w), 1410(m), 1448(vs), 1464(vs), 1501(w), 1648(w)	5.5e
oleic acid	602(w), 725(w), 845(m,sh), 856(m), 866(m), 890(m), 903(m), 971(m), 1023(m,sh), 1035(m,sh), 1065(m), 1080(m), 1118(m), 1265(m), 1301(s), 1416(m,sh), 1440(vs), 1655(s)	5.6a
vaccenic acid	817(w), 863(m), 889(m), 971(w), 1018(w), 1034(w), 1065(m), 1078(m), 1091(m), 1112(w,sh), 1264(m), 1302(s), 1416(w,sh), 1439(vs), 1655(s)	5.6b
glycerol	392(s), 416(m), 485(s), 548(mw), 675(m), 820(s), 850(vs), 923(m), 976(mw), 1055(vs), 1110(vs), 1257(m, br), 1315(m), 1465(vs)	5.6c
triolein	601(w), 725(w), 852(m), 869(m), 881(m), 971(mw), 1065(m), 1081(m), 1118(mw), 1266(m), 1301(s), 1440(vs), 1655(s), 1743(mw)	5.6d
trilinolein	841(mw), 868(m), 912(mw), 972(m), 1075(m), 1108(mw), 1264(m), 1302(m), 1440(m), 1654(vs), 1743(w)	5.6e
trilinolenin	866(m), 969(mw), 1024(mw), 1081(mw,br), 1160(mw), 1266(m), 1302(m), 1441(m), 1640(sh), 1655(vs), 1741(w)	5.6f
β-D-glucose	405(s), 440(m), 542(s), 650(mw), 772(mw), 841(ms), 914(ms), 1002(mw), 1022(m), 1054(m), 1075(ms), 1120(ms), 1149(m), 1272(m), 1296(ms), 1459(m)	5.7a
lactose	377(vs), 399(m), 421(mw), 446(mw), 477(s), 555(mw), 568(w), 633(mw), 648(w), 677(w), 700(w), 782(w), 851(m), 876(m), 900(mw), 916(ms), 953(mw), 1005(mw), 1021(ms), 1031(m), 1041(m), 1053(ms), 1087(ms), 1120(m), 1142(ms), 1168(w), 1222(w), 1261(m), 1296(mw), 1326(m), 1359(m), 1380(m), 1414(w), 1455(mw), 1470(w)	5.7b
cellulose	380(m), 436(w), 458(m), 493(m), 520(mw), 577(mw), 896(m), 969(w), 998(w), 1046(mw), 1061(m), 1096(vs), 1120(s), 1147(m), 1266(mw), 1337(m), 1379(m), 1413(m), 1461(m)	5.7c
D(+)-dextrose (α-D-glucose)	407(s), 440(m), 541(s), 652(m), 772(m), 841(s), 914(m), 1022(m), 1054(m), 1075(m), 1108(m), 1149(m), 1272(m), 1330(m), 1346(m), 1459(m)	5.7d
D(+)-trehalose	369(s), 407(vs), 430(m), 450(s), 523(vs), 540(s), 580(mw), 604(m), 671(m), 697(m), 732(w), 804(m), 838(vs), 912(vs), 927(w), 957(mw),	5.7e

Product	Raman bands	Fig
	1018(m), 1061(m), 1080(m), 1102(s), 1120(vs), 1149(vs), 1211(w), 1220(w), 1240(mw), 1261(mw), 1273(mw), 1313(mw), 1330(m), 1358(vs), 1371(m), 1386(m), 1411(m), 1455(m), 1469(mw)	
amylose	407(m), 442(m), 481(vs), 579(m), 712(w), 758(w), 854(m), 901(m), 936(ms), 944(m), 1045(m), 1083(ms), 1123(ms), 1208(w), 1262(mw), 1267(m), 1339(m), 1379(m), 1460(m)	5.7f
amylopectin	409(m), 439(m), 477(vs), 515(mw,sh), 578(mw), 612(w), 717(w), 762(m), 769(w), 866(m), 905(sh), 940(m), 1051(m), 1082(m), 1109(m), 1127(m), 1264(m), 1339(m), 1378(m), 1460(m)	5.7g
D(+)-mannose	397(s), 426(m), 447(s), 473(ms), 495(s), 530(s), 576(m), 604(m), 621(m), 683(m), 830(s), 861(s), 882(s), 914(m), 934(m), 958(m), 1033(m), 1087(s), 1106(s), 1139(ms), 1239(m), 1263(m), 1352(m), 1369(m), 1459(m), 1481(m)	5.8a
D(+)-fucose	383(w), 440(vs) 529(m), 669(m), 700(w), 772(mw), 815(m), 879(m), 964(mw), 1032(mw), 1076(mw), 1088(mw), 1113(m), 1129(m), 1155(m), 1256(m), 1273(m), 1331(m), 1452(m)	5.8b
D(-)-arabinose	406(w), 427(w), 512(mw), 580(m), 613(m), 698(m), 843(vs), 896(mw), 936(mw), 994(m), 1052(m), 1095(m), 1138(mw), 1259(m), 1311(mw), 1356(w), 1375(w), 1476(m)	5.8c
D(+)-xylose	410(m), 428(s), 504(m), 526(s), 565(m), 608(m), 757(mw), 903(vs), 930(m), 1017(mw), 1086(s), 1115(s), 1148(mw), 1315(mw), 1340(m), 1374(m), 1398(m), 1468(mw)	5.8d
D(-)-fructose	401(w,sh), 420(s), 463(m), 526(m), 595(m), 626(vs), 818(s), 872(s), 925(w), 978(m), 1048(m), 1060(w), 1082(m), 1143(w), 1176(w), 1250(w), 1265(s), 1340(w), 1399(w), 1455(w), 1471(m)	5.8e
D(+)-galactosamine	378(m), 421(m), 462(m), 530(vs), 606(mw), 657(w), 704(w), 776(w), 822(w), 872(vs), 887(m), 938(m), 948(m), 999(mw), 1020(m), 1038(mw), 1062(m), 1094(m), 1121(m), 1143(m), 1153(m), 1238(m), 1267(vs), 1318(w), 1336(m), 1369(mw), 1386(m), 1452(mw), 1468(w), 1514(mw), 1588(w), 1616(w)	5.9a
N-acetyl-D-glucosamine	416(m), 463(w), 481(m), 513(s), 527(m), 554(m), 578(s), 630(mw), 698(w), 790(m), 865(m), 916(mw), 929(m), 973(vs), 1002(m), 1020(m), 1039(m), 1091(m), 1126(s), 1151(m), 1206(w), 1257(m), 1279(m), 1292(mw), 1321(m), 1330(m), 1361(m), 1380(m), 1432(m), 1473(m), 1552(m), 1627(m)	5.9b
chitin	366(m), 396(m), 460(m), 498(m), 711(mw), 895(m), 955(m), 1059(m), 1107(s), 1149(m), 1205(mw), 1262(m), 1328(m), 1371(m), 1414(m), 1449(mw), 1626(m),1656(m)	5.9c
citric acid	380(s), 420(m), 489(w), 513(m), 554(m), 597(m), 641(w), 685(m), 784(vs), 881(w), 903(m), 942(s), 1053(m), 1082(m), 1145(w), 1170(w), 1211(m), 1345(w), 1390(m), 1426(w,sh), 1434(m), 1465(w), 1633(w), 1691(s), 1732(m)	5.10a

Product	Raman bands	Fig
succinic acid	386(s), 582(w), 684(m), 936(vs), 1086(m), 1230(w), 1294(m), 1369(m), 1420(s), 1432(s), 1457(m,sh), 1655(m)	5.10 b
fumarate	768(m), 913(m), 982(m), 1293(m), 1430(vs), 1563(w), 1593(w), 1629(m), 1640(m), 1657(s)	5.10 c
malic acid	400(mw), 440(mw), 535(mw), 611(s), 660(w), 749(s), 882(mw), 911(mw), 964(vs), 1033(m), 1098(m), 1176(w), 1186(w), 1222(w), 1261(w), 1277(w), 1309(w), 1351(m), 1378(mw), 1423(m), 1447(mw), 1633(mw), 1675(mw)	5.10 d
pyruvate	393(w), 435(mw), 547(w), 636(m), 834(s), 980(mw), 1197(w), 1371(m), 1408(vs), 1430(w,sh), 1628(w), 1658(mw), 1700(vs)	5.11 a
Phosphoenol-pyruvate	391(w), 406(w), 431(w), 451(mw), 465(w), 535(w), 548(mw), 741(w), 756(mw), 787(vs), 846(m), 863(mw), 891(w), 966(mw), 973(mw), 1034(s), 1085(w), 1097(w), 1185(w), 1234(w), 1266(m), 1286(w), 1318(w), 1355(w), 1373(mw), 1397(mw), 1446(m), 1472(w), 1631(mw)	5.11 b
coenzyme A	525(m), 559(m), 633(w), 663(m), 722(vs), 787(m,sh), 803(m), 879(m), 927(w), 949(w), 1027(m), 1053(m), 1100(m), 1157(w), 1203(m), 1248(m), 1325(s), 1407(s), 1420(s,sh), 1461(m), 1508(m), 1557(m), 1608(w), 1654(w)	5.11 c
acetyl coenzyme A	496(m), 531(m), 564(m), 631(s), 683(w,sh), 724(vs), 799(m), 882(m), 926(w), 948(w), 1026(m), 1045(m,sh), 1094(m), 1121(m), 1207(m), 1243(m), 1306(s), 1333(s), 1375(m), 1408(m), 1462(m), 1508(m), 1559(w), 1576(m), 1658(w)	5.11 d
acetoacetate	401(m), 486(m), 498(m), 519(m), 551(m), 582(m), 640(s), 719(w), 761(s), 830(vs), 903(m), 928(s), 1005(s), 1025(m), 1087(w), 1158(m), 1171(m), 1191(m), 1219(w), 1283(m), 1326(m), 1362(m), 1399(vs), 1422(vs), 1444(vs), 1580(m), 1600(w), 1626(w), 1653(w), 1703(vs), 1715(s,sh)	5.11 e
D-fructose-6-phosphate	390(m), 534(m), 617(m), 716(w), 834(m), 973(s), 1071(m), 1127(m), 1229(w,sh), 1266(m), 1347(m), 1376(m,sh), 1458(m), 1724(w)	5.11 f
β-carotene	1008(m), 1156(vs), 1190(m), 1211(m), 1270(w), 1280(w,sh), 1353(w), 1394(w), 1448(w), 1515(vs)	5.12 a
ascorbic acid	448(m), 566(s), 588(m), 629(s), 696(m), 710(m), 820(s), 1025(m,sh), 1130(vs), 1256(s), 1296(m), 1319(s), 1498(m), 1653(vs), 1667(vs)	5.12 b
riboflavin	422(w), 451(w), 502(m), 531(w), 603(w), 677(w), 742(m), 785(m), 1158(m), 1184(m), 1226(s), 1253(w), 1345(vs), 1402(m), 1464(m), 1496(m), 1534(m), 1576(m), 1621(w), 1658(w), 1704(w)	5.12 c
gluthatione	400(s), 446(w), 523(w), 550(m), 625(vs), 643(m), 660(s), 679(vs), 722(w), 776(s), 811(m), 828(m), 867(m), 885(s), 917(m), 931(s), 953(m), 972(m), 988(m), 1015(m), 1041(m), 1074(w), 1117(w), 1143(w), 1169(m), 1224(m), 1235(m), 1255(w), 1280(s), 1309(m), 1334(m), 1368(m), 1403(m), 1415(m), 1443(m), 1455(m), 1536(w), 1629(m), 1660(w), 1703(w)	5.12 d

Abbreviations: w, weak; mw, medium weak; m, medium; ms, medium strong; s, strong; vs, very strong; sh, shoulder; br, broad. The reference spectra can be requested at http://www.analchem.ugent.be/Raman

5.3.2 Amino acids

Raman spectra of eleven amino acids with a variety of functional groups on the R side chains were recorded. In Figure 5.2, Raman spectra of six amino acids without a ring structure on the R side chain are shown. Lin-Vien *et al.*[9] describe that several Raman bands for carboxylic acids and NH_2 groups are rather weak. This can explain why the amino acids, which have these two functionalities in common, show little similarities in their Raman spectra.

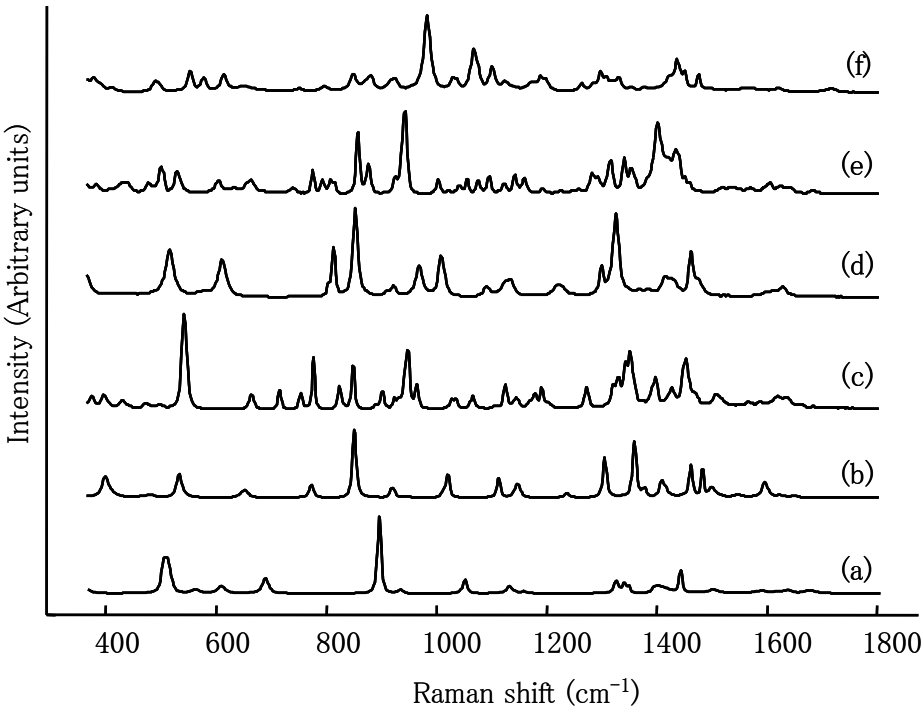


Figure 5.2 Raman spectra of amino acids with a non-cyclic R side chain: (a) glycine, (b) L-alanine, (c) L-valine, (d) L-serine, (e) L-glutamate, (f) L-arginine

In the Raman spectra of glycine (5.2a) and L-alanine (5.2b), which have small and simple R side chains, an intense band in the region between 850 and 900 cm^{-1} appears which can be ascribed to the ν(CNC) symmetric stretch of the amino group[9]. In amino acids with a more extended or complex R side chains, the influence of this R group on the position of the ν(C-N) stretch band can be large and therefore it is hard to assign a band to this vibration. For all amino acid spectra in Figure 5.2, $\delta(CH_2)$ or $\delta(CH_3)$ deformations are observed in the region 1420-1500 cm^{-1}. For L-valine (5.2c), L-serine (5.2d) and L-glutamate (5.2e) some bands are observed originating from functional groups of the R side chain. In the

spectrum of L-valine (5.2c), $\delta(CC_3)$ deformation and tertiary δ(CH) deformation of the isopropyl group give rise to bands at respectively 542 and 1351 cm^{-1}.[9] For L-serine (5.2d), the bands 813 and 853 cm^{-1} are located in the region where in phase ν(C-C-O) stretches of primary alcohol functions as well as ν(C-N) stretches of amines are expected. The symmetric $\nu(CO_2^-)$ stretch of L-glutamate (5.2e) can be observed at 1401 cm^{-1}.

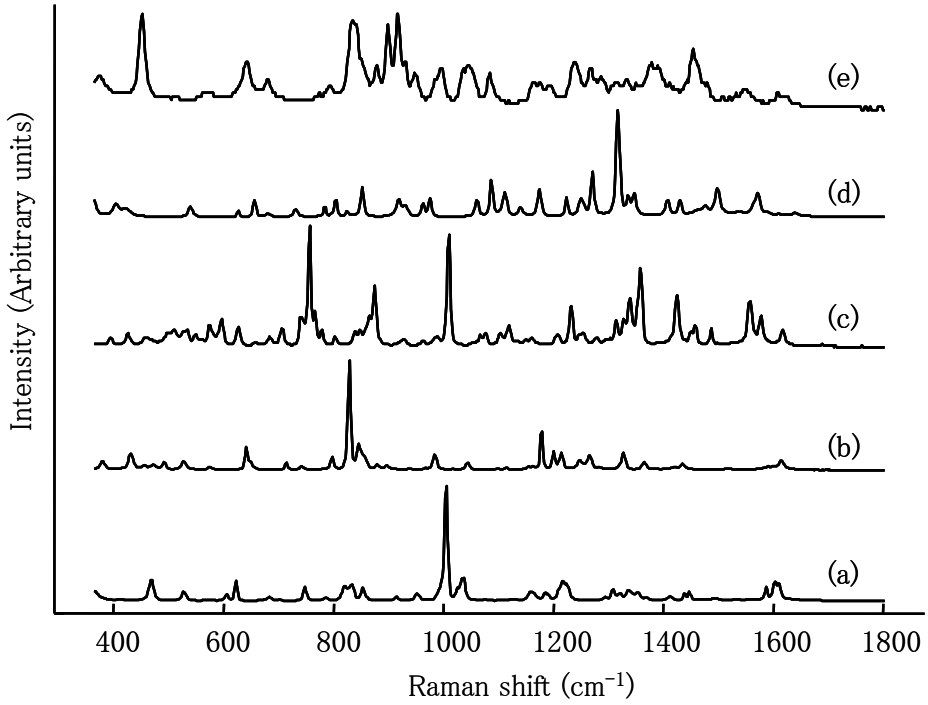


Figure 5.3 Raman spectra of amino acids with a cyclic R side chain: (a) L-phenylalanine, (b) L-tyrosine, (c) L-tryptophan, (d) L-histidine, (e) L-proline

Raman spectra of five amino acids of which the R side chain contains a ring structure are shown in Figure 5.3. For L-phenylalanine (5.3a), L-tyrosine (3b) and L-tryptophan (3c), that all contain a benzene structure, the most intense Raman bands can be attributed to the ring vibrations. Indeed, the intense bands at 1004 and 1009 cm^{-1} present in the spectra of respectively L-phenylalanine and L-tryptophan can be assigned to the trigonal ring breathing of the benzene ring. As L-tyrosine contains a parasubstituted benzene ring, its ring breathing vibration is located at 828 cm^{-1}. The Raman bands at 1358 and 1423 cm^{-1} of L-tryptophan, are described as typical for indole rings by Socrates *et al.*[19] For L-histidine and L-proline, it is hard to attribute the Raman bands to specific functional groups. Instead, quantum chemical calculations would be necessary in order to perform accurate band assignments.

5.3.3 Fatty acids and fats

The spectra of four unbranched saturated fatty acids are shown in Figure 5.4. Characteristic for these fatty acids are the three Raman bands between 1050 and 1150 cm^{-1}, the band at 1296 cm^{-1} and a group of bands in the region between 1400 and 1500 cm^{-1} due to ν(C–C) stretching vibrations, $\delta(CH_2)$ twist vibrations and $\delta(CH_3)$ or $\delta(CH_2)$ deformations, respectively. As the fatty acid chain elongates with two carbon atoms, several effects in the Raman spectra can be observed. Two shifts occur with increasing chain length: (i) the ν(C–C) stretching vibration band at 1084 cm^{-1} in the spectrum of lauric acid (5.4a) shifts over 1092 cm^{-1} (myristic acid, 5.4b) and 1099 cm^{-1} (palmitic acid, 5.4c), to 1100 cm^{-1} for stearic acid (5.4d) and (ii) the chain expansion vibration shifts from 464 cm^{-1} for lauric acid (5.4a), over 414 cm^{-1} (myristic acid, 5.4b), to 375 cm^{-1} for palmitic acid (5.4c). The relative intensity of the bands at approximately 891 and 908 cm^{-1}, due to $\rho(CH_2)$ rocking vibrations, changes with the chain length.

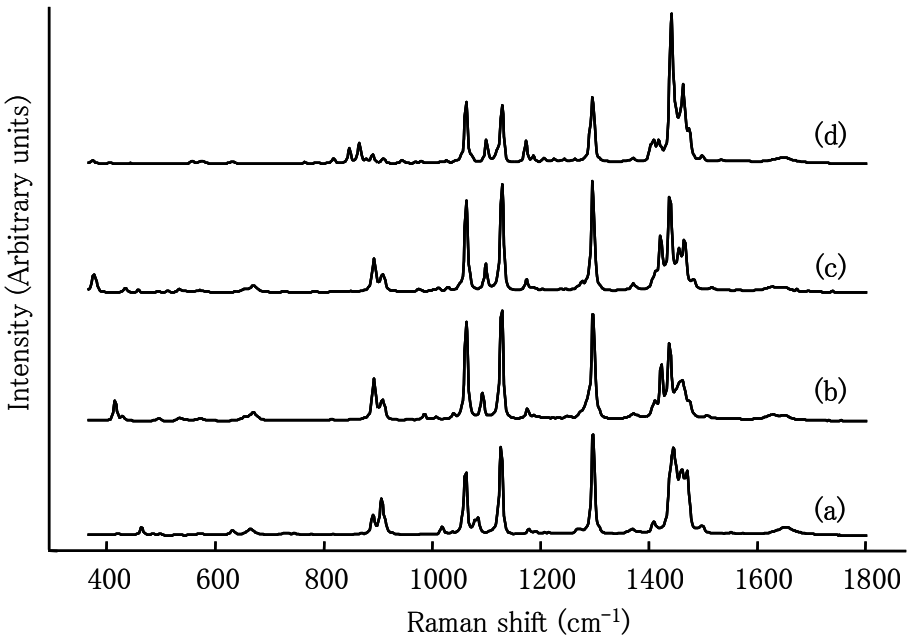


Figure 5.4 Raman spectra of saturated lineair fatty acids: (a) lauric acid, (b) myristic acid, (c) palmitic acid, (d) stearic acid

In Figure 5.5, spectra of saturated branched fatty acids are shown (full names together with their abbreviations are given in Table 5.2). Although the general profile of these Raman spectra is similar to the spectra of the linear fatty acids, there are some clear differences.

Comparison of the Raman spectra of palmitic acid (5.4c) and 16iso (5.5c) shows that the $\delta(CH_2)$ and $\delta(CH_3)$ deformations result in only two intense bands in the region from 1400 until 1500 cm^{-1} for branched fatty acids, while the Raman spectra of the linear fatty acids show four to five smaller bands in that region. The $\rho(CH_2)$ rocking doublet at 893 and 909 cm^{-1} in palmitic acid (5.4c) is decomposed into several bands in the spectrum of 16iso (5.5c) and some additional bands appear around 800 cm^{-1}. In the region of the chain expansion vibrations (below 470 cm^{-1}) more bands appear in the spectra of the branched fatty acids.

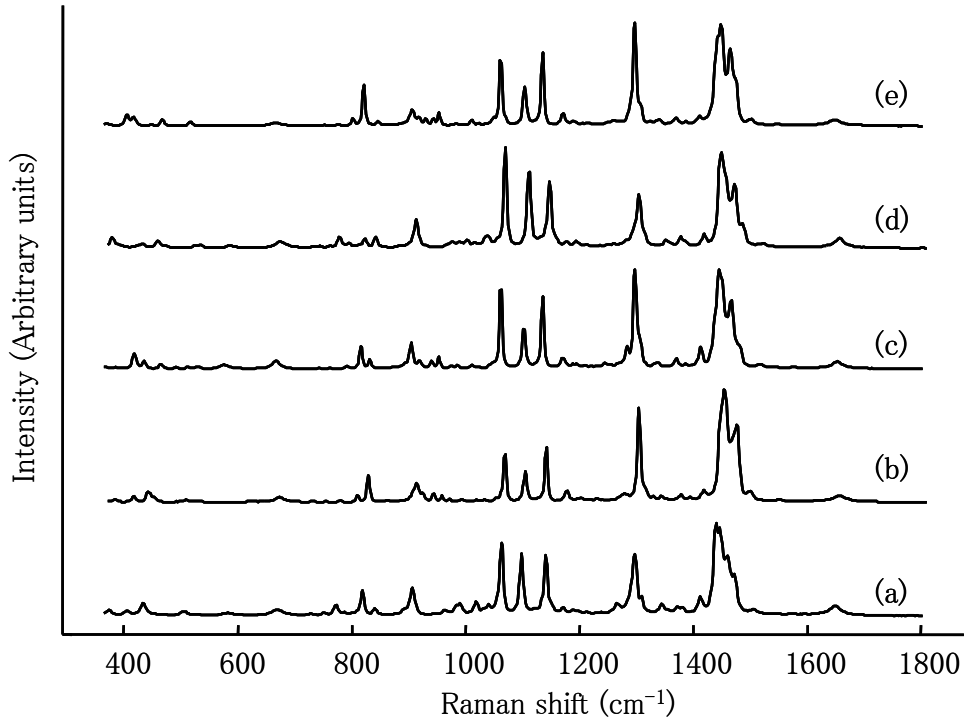


Figure 5.5 Raman spectra of saturated branched fatty acids: (a) 12-methyltetradecanoic acid (15Aiso), (b) 13-methylmyristic acid (15iso), (c) 14-methylpentadecanoic acid (16iso), (d) 14-methylhexadecanoic acid (17Aiso), (e) 15-methylpalmitic acid (17iso)

The spectra of the iso fatty acids (fatty acids branched at the one but last carbon atom) clearly show some differences compared to the spectra of the Aiso acids (fatty acids branched at the third carbon atom, counted from the end of the carbon chain). Firstly, in the spectra of the Aiso acids, the relative intensity of the band at 1296 cm^{-1} is lower than in the spectra of the iso acids. Secondly, for the Aiso acids the three bands in the region 1050 to 1150 cm^{-1} have about the same intensity, but for the iso acids the band at

approximately 1090 cm^{-1} is smaller compared to the two surrounding bands. And thirdly, in the region of the $\rho(CH_2)$ rocking vibrations the Aiso fatty acids show only one band at approximately 906 cm^{-1}, while the iso fatty acids show three to four extra bands between 910 and 970 cm^{-1}.

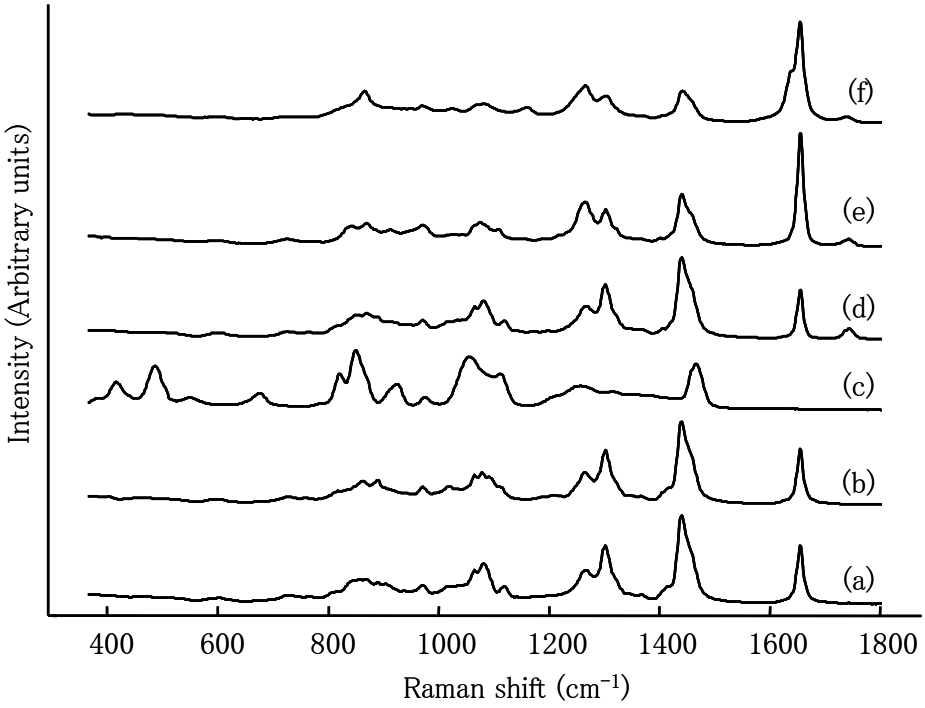


Figure 5.6 Raman spectra of the unsaturated fatty acids (a) oleic acid and (b) cis-vaccenic acid, of (c) glycerol and of the fats (d) triolein, (e) trilinolein and (f) trilinolenin

The Raman spectra of two mono unsaturated fatty acids are shown in Figure 5.6: oleic acid (5.6a) and cis-vaccenic acid (5.6b). These acids are fluids and so their Raman spectra show broader bands compared to the saturated fatty acids, although all bands are located around the same wavenumber. For the unsaturated fatty acids (5.6a,b) an additional band at 1655 cm^{-1} appears due to $\nu(C=C)$ stretching vibrations. These acids differ only in the position of the double bond (see appendix), so their Raman spectra are highly similar. Only in the fingerprint region (600-1200 cm^{-1}) some minor differences can be observed. The spectrum of triolein (glycerol esterificated three times with oleic acid) shows a very similar Raman pattern compared to oleic acid. Indeed, no contribution of the glycerol spectrum (5.6c) is seen in the three fats (5.6d,e,f): the glycerol spectrum shows clear bands at 416 and 485 cm^{-1}, a region where no bands are present in the spectra of the fats. The most important difference between the fatty acids (5.6a,b) and the fats (5.6d,f) is the

presence of a weak band at about 1743 cm^{-1} that can be attributed to ν(C=O) stretching vibrations from the ester bonds between glycerol and the fatty acids. Triolein, trilinolein and trilinolenin consist of glycerol esterified with identical fatty acids containing respectively one, two and three double bonds (see appendix). Consequently the intensity of the ν(C=C) stretching band at 1655 cm^{-1} increases compared to the other Raman bands in the spectra 5.6d, e and f. Next to some minor changes in the fingerprint region, it is noteworthy that the relative intensity of the bands at 1265 and 1301 cm^{-1}, due to respectively δ(=CH) deformations and $\delta(CH_2)_n$ deformations, changes as the fats contain more double bonds.

5.3.4 Saccharides

Figure 5.7 shows the spectra of β-D-glucose and D(+)-dextrose (α-D-glucose) and some of their dimers and polymers. The spectra of β-D-glucose (5.7a) and D(+)-dextrose (5.7d) are very similar, but can be distinguished from one another in the regions from 1000 to 1200 and from 1300 to 1500 cm^{-1} that contain bands of ν(C-O) and ν(C-C) stretches respectively $\delta(CH_2)$ and $\delta(CH_2OH)$ deformations[22]. The spectra of D(+)-dextrose (5.7d) and its dimer D(+)-trehalose (5.7e) are also very similar, although a clear shift of an intense band can be observed from 541 to 523 cm^{-1}. This band is associated with exocyclic deformations, while the band at approximately 405 cm^{-1} can be assigned to endocyclic deformations[16]. In the spectra of β-D-glucose (5.7a) and lactose (5.7b) the endocyclic and exocyclic deformation bands both show large shifts from 405 to 377 cm^{-1} and from 542 to 477 cm^{-1} respectively. In the regions of δ(COH), δ(CCH) and δ(OCH) side group deformations (800-950 cm^{-1}), of ν(C-O) and ν(C-C) stretches (950-1200 cm^{-1}) and of $\delta(CH_2)$ and $\delta(CH_2OH)$ deformations (1250-1500 cm^{-1}), the spectrum of lactose shows more bands than the spectrum of β-D-glucose (5.7a). This is probably due to the fact that lactose consists of β-D-glucose and another monosaccharide, viz. β-D-galactose. In the spectra of cellulose (5.7c), amylose (5.7f) and amylopectin (5.7g), polymers of β-D-glucose and D(+)-dextrose, the Raman bands are much broader than those in the spectra of the monomers. Despite the broadening of the Raman bands in the spectra of amylose (5.7f) and

amylopectine (5.7g), they resemble the spectrum of their monomer D(+)-dextrose (5.7d) in the region from 800-1500 cm^{-1}. In the region below 700 cm^{-1}, intense bands for endo and exocyclic deformations were observed in the spectrum of D(+)-dextrose (5.7d), but only one intense band is present in the spectra of amylose (5.7f) and amylopectin (5.7g). The branching in the structure of amylopectin does not induce many changes in its Raman spectrum compared to that of amylose. The spectrum of cellulose (5.7c) does not resemble the spectrum of its monomer β-D-glucose (5.7a).

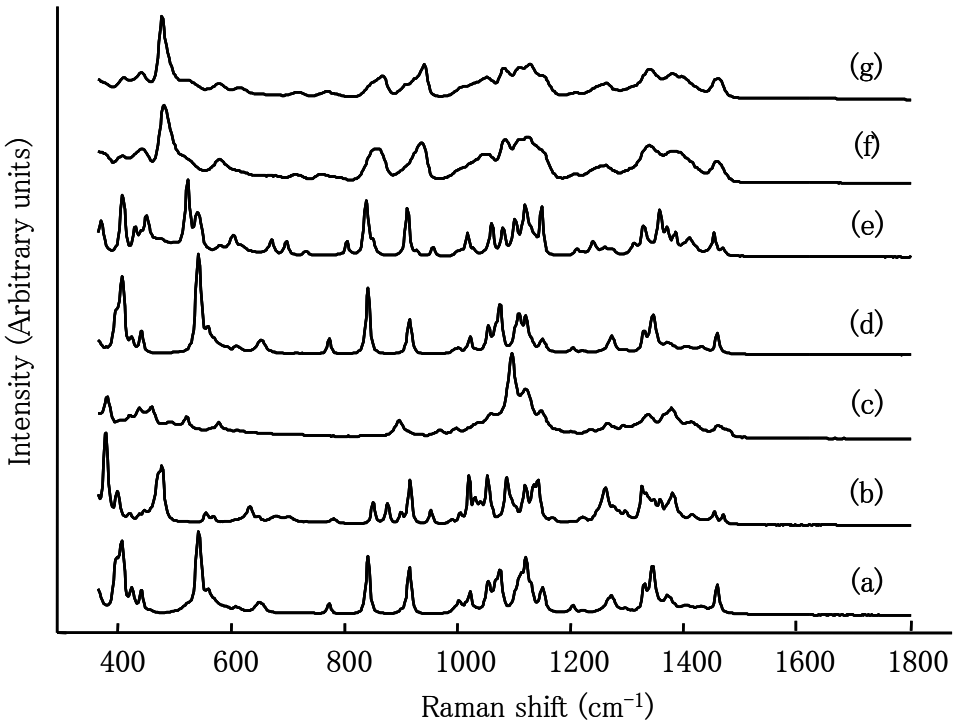


Figure 5.7 Raman spectra of saccharides: (a) β-D-glucose, (b) lactose, (c) cellulose, (d) D(+)-dextrose, (e) D(+)-trehalose, (f) amylose, (g) amylopectine

The spectra of five other monosaccharides are shown in Figure 5.8. Except for D(+)-mannose, they can also appear in a furanose structure instead of a pyranose structure. For D(+)-mannose, we noticed sample inhomogenity and so the average spectrum contains many small bands (5.8a) and is hard to compare with the spectra of the other monosaccharides. D(+)-fucose (5.8b), D(-)-arabinose (5.8c) and D(+)-xylose (5.8d) differ only in a modification of the CH_2OH group, but their Raman spectra are clearly different from one another and from the spectrum of β-D-glucose.

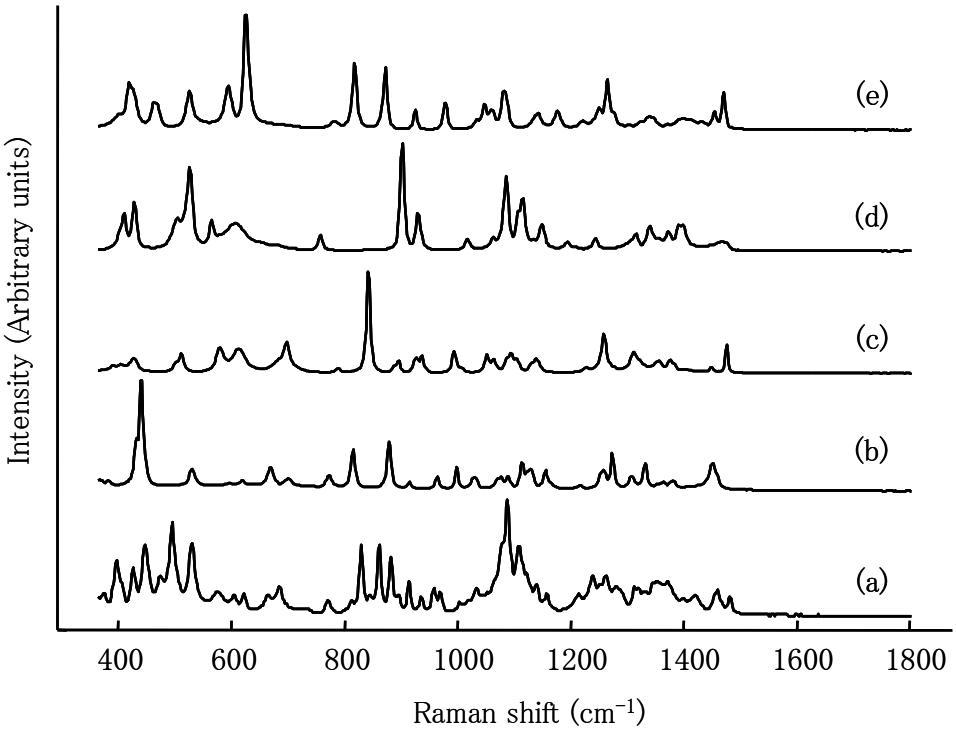


Figure 5.8 Raman spectra of saccharides: (a) D(+)-mannose, (b) D(+)-fucose, (c)-D(-)-arabinose, (d) D(+)-xylose, (e) D(-)-fructose

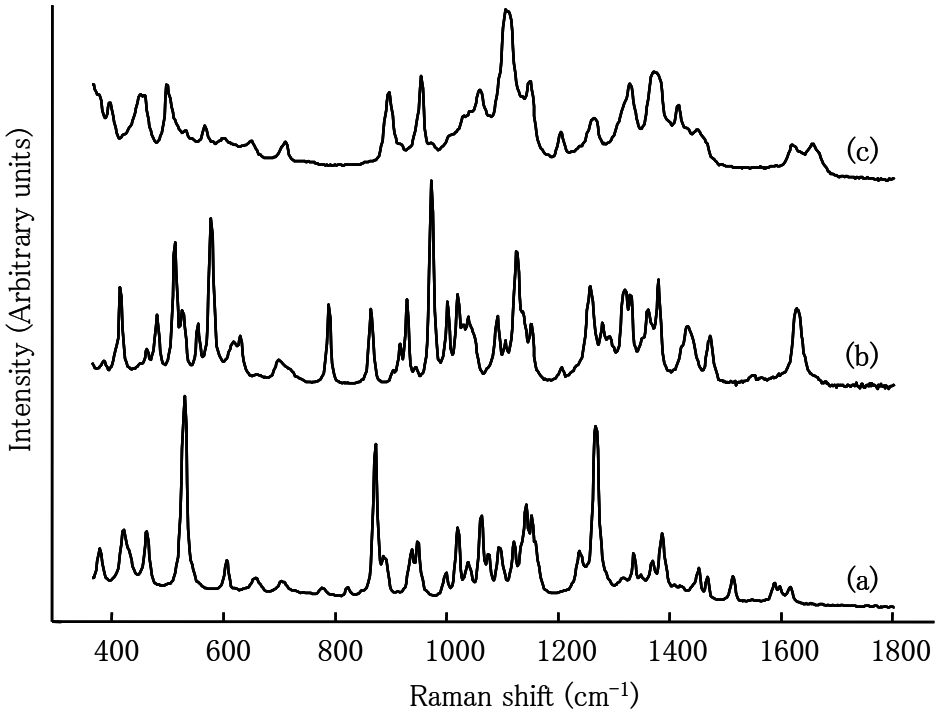


Figure 5.9 Raman spectra of nitrogen containing saccharides: (a) D(+)-galactosamine, (b) N-acetyl-D-glucosamine, (c) chitin

Figure 5.9 presents spectra of three nitrogen-containing saccharides. For D(+)-galactosamine (5.9a) many shifts can be observed compared to the spectrum of β-D-glucose (5.7a), which are caused by the amino group and the change of a stereocenter (see appendix). The most prominent band, compared to other monosaccharides, is the intense band at 1267 cm^{-1}, which is situated in the region of $\rho(NH_2)$ rocking and $\delta(NH_2)$ twisting vibrations[19]. For N-acetyl-D-glucosamine (5.9b), the presence of the amide group induces a much more complex Raman spectrum compared to β-D-glucose (5.7a). The band at 1627 cm^{-1} is typical for the $\nu(C=O)$ stretch of an acetyl amide[9]. The spectrum of chitin (5.9c) shows broader bands than the spectrum of its monomer N-acetyl-D-glucosamine (5.9b).

5.3.5 Primary metabolites

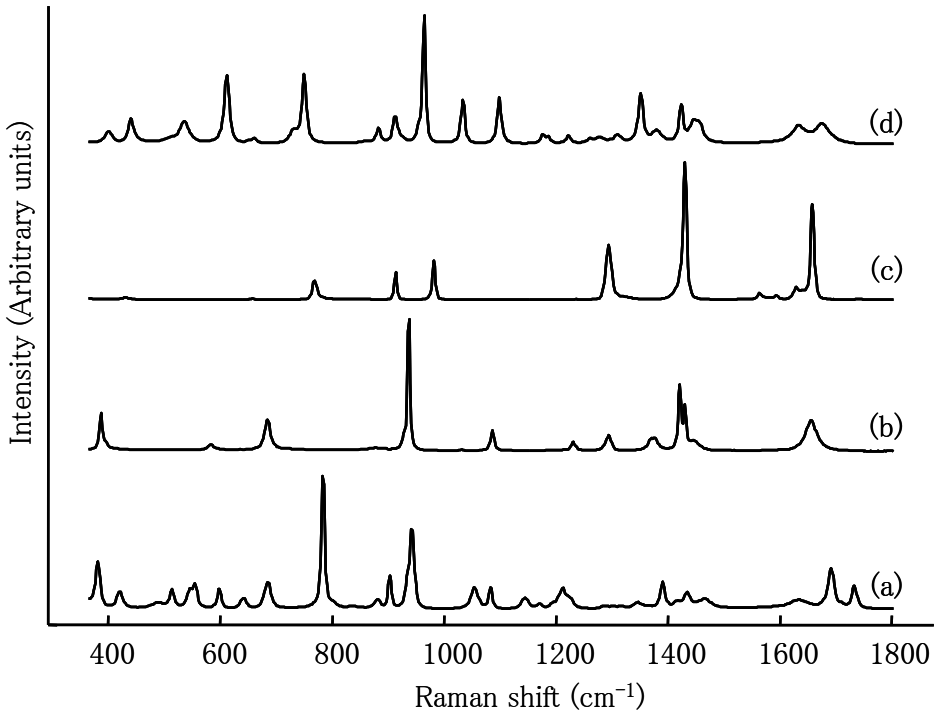


Figure 5.10 Raman spectra of components of the citric acid cycle: (a) citric acid, (b) succinic acid, (c) fumarate, (d) malic acid

Figures 5.10 and 5.11 show Raman spectra of primary metabolites. Figure 5.10 contains the spectra of four intermediates from the citric acid cycle: citric acid (5.10a), succinic acid (5.10b), fumarate (5.10c) and malic acid (5.10d). In the spectra of the acids (5.10a, c, d), a

strong band is present in the region between 930 and 970 cm^{-1} that can be assigned to the δ(OH--O) out of plane wagging vibration of intermolecular hydrogen bonds. The spectrum of citric acid shows an intense band at 784 cm^{-1}, which is probably caused by a $\delta(C_3CO)$ deformation. For fumarate (disodium salt, the acid was not commercially available) the symmetric $\nu(CO_2^-)$ stretch results in a strong band at 1430 cm^{-1}, while the ν(C=C) bond that is conjugated with the carbonyl groups gives rise to a ν(C=C) stretch band at 1657 cm^{-1}. Although succinic acid, fumarate and malic acid are subsequent intermediates in the citric acid cycle and have readily similar chemical structures (see appendix), their Raman spectra can be easily distinguished from each other. From comparison with biological materials, we suspect that Raman bands in the spectra of the pure acids better resemble the bands in biological spectra than Raman bands in the spectra of the corresponding salts (such as disodium fumarate).

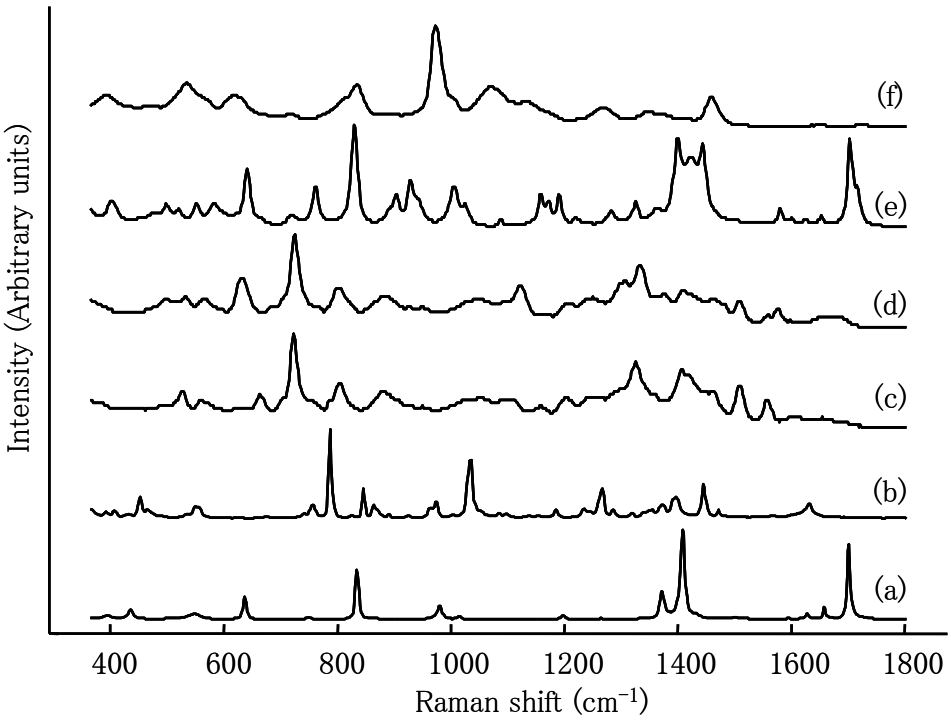


Figure 5.11 Raman spectra of primary metabolites related to the citric acid cycle: (a) pyruvate, (b) phosphoenolpyruvate, (c) coenzyme A, (d) acetyl coenzyme A, (e) acetoacetate, (f) D-fructose-6-phosphate

Figure 5.11 present Raman spectra of six metabolites that are related to the citric acid cycle. The spectra of pyruvate (5.11a) and phosphoenolpyruvate (5.11b) are very different, because the Raman bands of different functional groups dominate them. For

pyruvate (5.11a), the most intense bands at 1408 and 1700 cm^{-1} can be attributed to the symmetric and asymmetric $\nu(CO_2^-)$ stretch respectively. The symmetric and asymmetric ν(POC) stretches give rise to the intense bands at 787 and 1034 cm^{-1} in the spectrum of phosphoenolpyruvate (5.11b). The spectra of coenzyme A (5.11c) and acetyl coenzyme A (11d) are similar and both contain a band at approximately 723 cm^{-1}, which is the most intense band in the spectrum of adenine (5.1a). The ν(C-S) stretching bands of coenzyme A (5.11c) shifts from 663 cm^{-1} to 631 cm^{-1} in the spectrum of acetyl coenzyme A (5.11d), due to the conversion of the thiol to a thioester. For acetoacetate (5.11e), analogous to pyruvate (5.11a), the bands for symmetric and asymmetric $\nu(CO_2^-)$ stretches are observed in the region between 1395 and 1450 cm^{-1} and between 1700 and 1720 cm^{-1}. In the spectrum of D-fructose-6-phosphate (5.11f), the bands are broader than in the spectrum of D(-)-fructose (5.8e). The asymmetric ν(POC) stretch of the phosphate group (973 cm^{-1}) causes the most intense band as this functional group is highly polarisable.

5.3.6 Others

Figure 5.12 presents Raman spectra of four other products that are of importance for living organisms. The spectrum of β-carotene (5.12a) is characterized by two intense bands at 1156 and 1515 cm^{-1}, that can be assigned to a combination of a ν(C-C) stretch and ν(C-H) deformation on the one hand and ν(C=C) stretch on the other hand. The spectrum of ascorbic acid (5.12b) shows a broad band at 1653-1667 cm^{-1} due to the ν(C=O) stretch of the lacton. In the spectrum of riboflavin (5.12c) a strong band can be observed at 1226 cm^{-1} that can be assigned to an aromatic vibration of the polycyclic structure. In the spectrum of glutathione (5.12d) intense bands are observed that are related to the presence of the sulphur atom: (i) ν(C-S) stretching causes intense bands in the region 600-700 cm^{-1} and (ii) the band at 400 cm^{-1} can be assigned to δ(C-S) deformations. Furthermore, the ν(C=O) stretching of amide and carboxylic groups gives rise to a broad band at 1630 cm^{-1}.

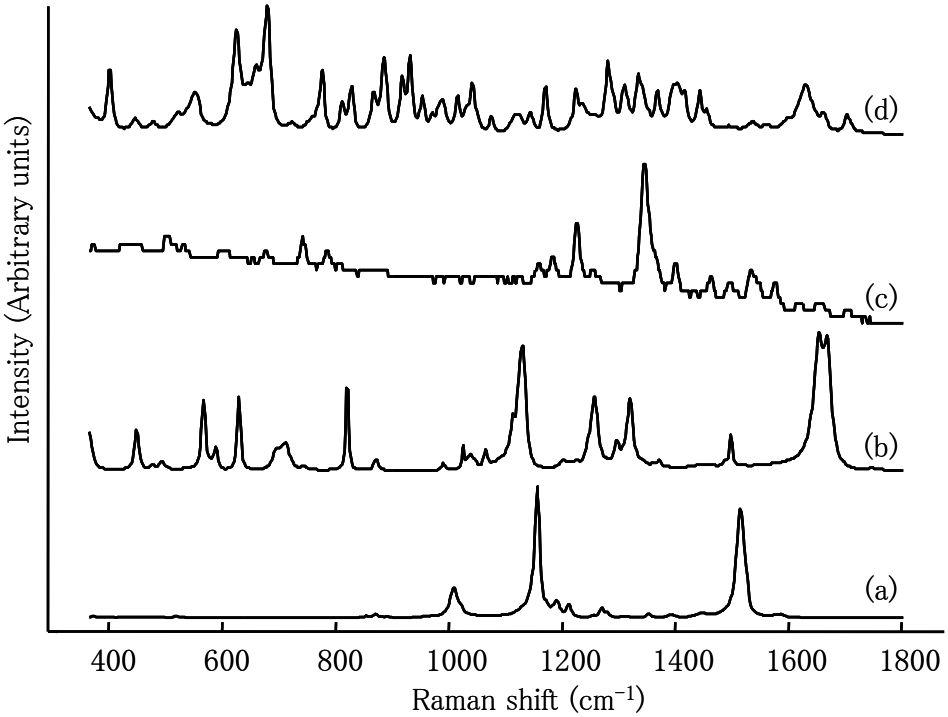


Figure 5.12 Raman spectra of (a) β-carotene, (b) ascorbic acid, (c) riboflavin, (d) glutathione

5.3.7 Examples of Raman spectra of biological materials

All these reference spectra can be used for the interpretation of Raman spectra of biological materials. To illustrate the applicability of these reference spectra, we selected two Raman spectra of biological materials. In Figure 5.13, Raman spectra of fungal spores (5.13a) and bacteria (5.13b) are shown. There is a high similarity between the spore spectra of *Amanita phalloides* (5.13a) and the spectra of fatty acids and fats (Fig. 5.4, 5.5, 5.6). Indeed, the typical bands for these molecules that were mentioned above appear in the region from 1000 until 1700 cm^{-1}. Therefore, we can conclude that fatty acids and fats are the main components of these spores. The bands at 970, 1065, 1081, 1121, 1302, 1440 and 1656 cm^{-1} can be attributed to triolein and trilinolenin. Oleic acid bands can be observed at 1036, 1302 and 1440 cm^{-1}. The amyloidic ornamentation, consisting of amylopectine, is observed in this spectrum trough the Raman bands at 477 and 864 cm^{-1}. In the spectrum of *Bacillus valismortis* LMG 18725 (5.13b) clear bands of fatty acids (1450(br), 1660(br) cm^{-1}), phenylalanine (1003 cm^{-1}), adenine (723 cm^{-1}) and carotenes (1155 and 1520 cm^{-1}) can be observed.

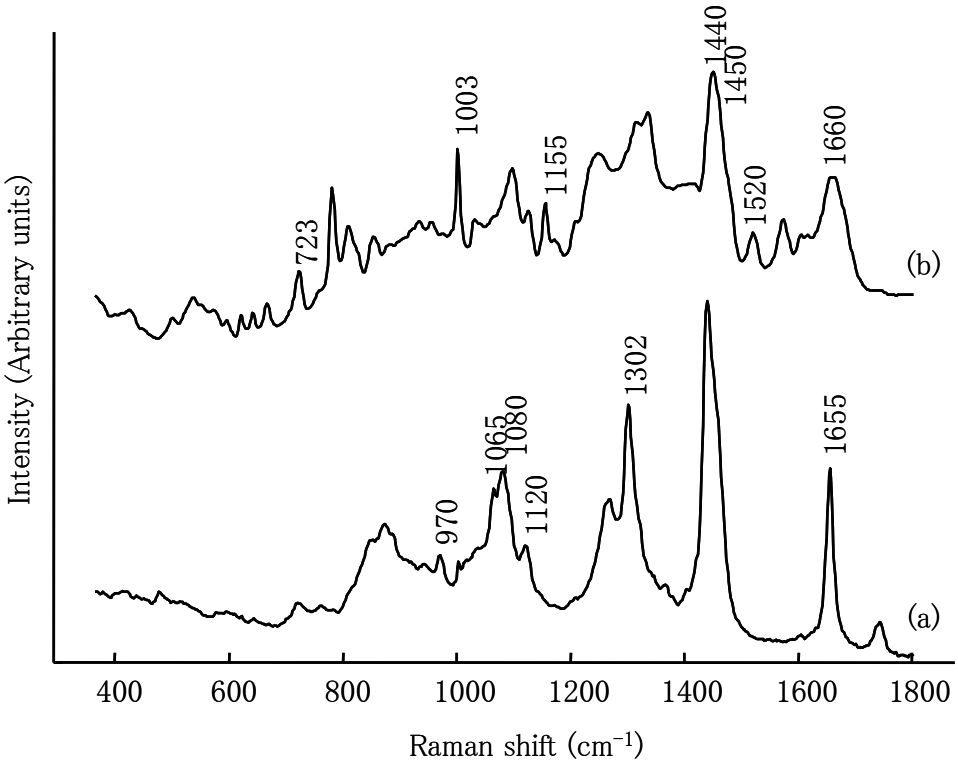


Figure 5.13 Raman spectra of biological materials: (a) spores of *Amanita phalloides* and (b) *Bacillus vallismortis* LMG 18725

5.4 Conclusions

We presented Raman spectra of six groups of biomolecules. Most of these spectra can be easily distinguished from each other. The spectra of the DNA and RNA bases show almost no similarities, so it is easy to distinguish between them and no characteristic bands are observed for either purine or pyrimidine bases. The spectra of all bases are characterized by intense ring breathing vibrations in the region from 600 until 800 cm^{-1}. The Raman bands of the functional groups present in the R side chain dominate the Raman spectra of amino acids. Only structurally very similar molecules, such as fatty acids mutually and saccharides mutually, give rise to similar Raman spectra. The spectra of fatty acids can be recognized by three intense bands in the region from 1000 until 1200 cm^{-1} and bands at about 1296 and 1440 cm^{-1}. For fats the most prominent bands are those at approximately 1265, 1302, 1440 and 1655 cm^{-1}. Examples of Raman spectra of bacteria and fungal spores

show that these reference spectra can be used to analyze complex biological Raman spectra.

This database was after publication of the paper (De Gelder et al., J. Raman Spectrosc. 38 (2007) 1138-1147) further extended with reference spectra of biomolecules of interest for specific studies. These include spectra of calcium-dipicolonate, cysteïne, poly-3-hydroxybutyrate and β-3-hydroxybutyrate, presented in Chapter 6 and 9. It must be noted that the spectra of the reference products can differ slightly from their signals in biological samples. Indeed, a different physicochemical state and the physiological environment can induce shifts in Raman band positions and/or intensities. A strongly pronounced example of these shifts is illustrated in Chapter 9 where a ν(C=O) stretching band of poly-3-hydroxybutyrate (PHB) is located at 1725 cm^{-1} in the reference spectrum and at 1734 cm^{-1} in the bacterial spectra. The physiological conditions of biomolecules in a cell can not be fully imitated in vitro and an attempt to do so would lead to reference spectra including contributions of biomolecules other than the one of interest. The reference spectra in this database should thus be considered as approaches of the biomolecules' signals in bacterial Raman spectra. They are useful to evaluate the possible contribution of certain biomolecules to a complex bacterial Raman spectrum.

The next chapter starts with an illustration of the use of these reference spectra to extract information about the presence of specific cell compounds from bacterial Raman spectra. It focuses on compounds that clearly contribute (observed by visual inspection) to bacterial Raman spectra. Processing methods are illustrated that allow information extraction of such compounds from Raman spectra of sporulating bacteria and endospore solutions.

References

1. Carey P.R. (1999) Raman spectroscopy, the sleeping giant in structural biology, awakes. *J. Biol. Chem.* 274: 26625-26628.
2. Chumanov D., Picorel R., de Zarate I.O., Cotton T.M., Seibert M. (2000) Resonance raman and surface-enhanced resonance raman spectra of LH2 antenna complex from *Rhodobacter sphaeroides* and *Ectothiorhodospira* sp excited in the Q(x) and Q(y) transitions. *Photochem. Photobiol.* 71: 589-595.
3. De Gussem K., Vandenabeele P., Verbeken A., Moens L. (2005) Raman spectroscopic study of *Lactarius* spores (Russulales, Fungi). *Spectrochim. Acta A* 61: 2896-2908.
4. Edwards H.G.M. (2004) Probing history with Raman spectroscopy. *Analyst* 129: 870-879.
5. Edwards H.G.M., Newton E.M., Wynn-Williams D.D., Lewis-Smith R.I. (2003) Non-destructive analysis of pigments and other organic compounds in lichens using Fourier-transform Raman spectroscopy: a study of Antarctic epilithic lichens. *Spectrochim. Acta A* 59: 2301-2309.
6. Hutsebaut D., Vandenabeele P., Moens L. (2005) Evaluation of an accurate calibration and spectral standardization procedure for Raman spectroscopy. *Analyst* 130: 1204-1214
7. Hutsebaut D., Vandroemme J., Heyrman J., Dawyndt P., Vandenabeele P., Moens L., De Vos P. (2006) Raman microspectroscopy as an identification tool within the phylogenetically homogeneous '*Bacillus subtilis*'-group. *Syst. Appl. Microbiol.* 29: 650-660.
8. Ibelings M.S., Maquelin K., Endtz H.P., Bruining H.A., Puppels G.J. (2005) Rapid identification of *Candida* spp. in peritonitis patients by Raman spectroscopy. *Clin. Microbiol. Infec.* 11: 353-358.
9. Lin-Vien D., Colthup N.B., Fateley W.G., Grasselli J.G. (1991) The handbook of infrared and Raman characteristic frequencies of organic molecules. Academic press, San Diego.

10. Lippert J.J.L., Peticolas W.L. (1972) Raman active vibration in long-chain fatty-acids and phospholipids sonicates. *Biochimica et Biophisica Acta* 282: 8.
11. Lopez-Diez E.C., Goodacre R. (2004) Characterization of microorganisms using UV resonance Raman spectroscopy and chemometrics. *Anal. Chem.* 76: 585-591.
12. Maquelin K., Choo-Smith L.P., Endtz H.P., Bruining H.A., Puppels G.J. (2002) Rapid identification of *Candida* species by confocal Raman micro spectroscopy. *J. Clin. Microbiol.* 40: 594-600.
13. Maquelin K., Krischner C., Choo-Smith L.P., van den Braak N., Endtz H.Ph., Naumann D., Puppels G.J. (2002) Identification of medically relevant microorganisms by vibrational spectroscopy. *J. Microbiol. Meth.* 51: 255-271.
14. Martens H., The EMSC toolbox for MATLAB http://www.models.kvl.dk/source/emsctoolbox
15. Martens H., Nielsen J.P., Engelsen S.B. (2003) Light scattering and light absorbance separated by extended multiplicative signal correction. Application to near-infrared transmission analysis of powder mixtures. *Anal. Chem.* 75: 394-404.
16. Mrozek M.F., Weaver M.J. (2002) Detection and identification of aqueous saccharides by using surface-enhanced Raman spectroscopy. *Anal. Chem.* 74: 4069-4075.
17. Oust A., Moretro T., Naterstad K., Sockalingum G.D., Adt I., Manfait M., Kohler A. (2006) Fourier transform infrared and Raman spectroscopy for characterization of *Listeria monocytogenes* strains. *App. Environ. Microb.* 72: 228-232.
18. Stair P.C. (2001) Advances in Raman spectroscopy methods for catalysis research. *Curr. Opin. Solid St. M.* 5: 365-369.
19. Socrates G. (2001) Infrared and Raman characteristic group frequencies. Third edition, John Wiley & Sons, Chichester.
20. Vandenabeele P., Edwards H.G.M., Moens L. (2007) A decade of Raman spectroscopy in art and archaeology. *Chem. Rev.* 107: 675-686.
21. Weesie R.J., Merlin J.C., Lugtenburg J., Britton G., Jansen F.J.H.M., Cornard J.P. (1999) Semiempirical and Raman spectroscopic studies of carotenoids. *Biospectroscopy* 5: 19-33.

22. Weng Y.M., Weng R.H., Tzeng C.Y., Chen W. (2003) Structural analysis of triacylglycerols and edible oils by near-infrared Fourier transform Raman spectroscopy. *Appl. Spectrosc.* 57: 413-18.
23. Zeiri L., Bronk B.V., Shabtai Y., Eichler J., Efrima S. (2004) Surface-enhanced Raman spectroscopy as a tool for probing specific biochemical components in bacteria. *Appl. Spectrosc.* 58: 33-40.

Chapter 6: Raman spectroscopic study of bacterial endospores

Joke De Gelder, Patsy Scheldeman, Karen Leus, Marc Heyndrickx,
Peter Vandenabeele, Luc Moens, Paul De Vos

Analytical and Bioanalytical Chemistry, 389 (2007) 2143-2151

In the previous chapter we established a database with reference spectra of biomolecules. Now we will examine to which extent reference spectra can be utilized to study sporulating bacteria and bacterial endospore solutions with Raman spectroscopy. In general, an overview is given of approaches to extract information when the biomolecules of interest contribute highly to the bacterial or endospore Raman spectra. These approaches use information from reference spectra and include the calculation of difference spectra and the use of principal component analysis (PCA).

6.1 Introduction

Upon depletion of essential nutrients, some Gram-positive bacteria can form endospores with a complex structure that ensures the maintenance of a dormant state. When conditions become favorable, they may germinate to vegetative cells. Detection of spores is very important, because of their role in e.g. food spoilage and foodborne disease. Resistance of endospores to extreme conditions can be influenced by multiple features such as water content of the core, the presence of small acid-soluble proteins (SASPs) that can bind to DNA, spore coat proteins and core mineral ions and 2,6-pyridinedicarboxylic acid (dipicolinic acid; DPA), as reported by Setlow[26]. This review describes that resistance to each specific condition, such as wet heat, dry heat, ultraviolet light and chemicals, is determined by a combination of the above mentioned spore characteristics. However, the way these components contribute to these forms of resistance is not always clear. Another study[23] indicates that DPA may contribute to a spore's resistance by protecting spore DNA from damage. Therefore, not only the detection of spores is important, but also the study of their composition can be of interest in relation to their resistance.

Raman spectroscopy is a fast and non-destructive method of analysis that has proven to be successful in microbiology for the identification of bacterial species[11,12] or strains[10,18]. As Raman spectra contain contributions of all Raman active molecules present in the sample, this technique is also useful for the detection of certain cell components[5,15,16]. Several research groups have already investigated the possibility to detect endospores by focusing

on the Raman bands of the calcium complex of DPA (CaDPA)[2,3,4,6,7,17,19]. Bell *et al.*[2] reported that surface enhanced Raman spectroscopy (SERS) is suitable for the quantitative analysis of pure CaDPA. It has been demonstrated that it is possible to detect DPA in bacterial samples by using Raman spectroscopy[6], as well as by using surface enhanced Raman spectroscopy (SERS)[4], coherent anti-Stokes Raman spectroscopy[19] and resonance Raman spectroscopy[7]. Even determination of concentrations of CaDPA in single spores with Raman spectroscopy[3] and SERS[17] has been reported.

In this study, we focus on the analysis of the spore's composition rather than on detection of spores. Changes in Raman spectra during growth and sporulation of *Bacillus licheniformis* LMG 7634 were studied. Raman spectra of spore suspensions of several strains cultured under different conditions were compared. Different growth media and incubation temperatures were applied and certain samples were pre-treated with hydrogen peroxide. The latter is known to be a heat resistance inducing stress factor and resistance against it is influenced by spore coat layers[20] and the binding of SASPs to DNA[24]. A sublethal hydrogen peroxide treatment of spores resulted in a heat resistance induction effect for a subpopulation of *B. sporothermodurans* spores, termed HRS[22]. The possibility of discriminating these spore suspensions from each other by their Raman spectra was studied, as well as possible assignments of discriminating bands to biomolecules in the spores.

6.2 Experimental

6.2.1 Strains and culturing conditions

Bacillus licheniformis LMG 7634 was precultured overnight on Tryptone Soya Agar (TSA) (Oxoid) at 37 °C. A single colony was picked to make other streaks on TSA that was incubated for 1, 2, 3, 4, 5 or 6 days. This was repeated at least 5 times for each incubation time to include variation in microbial growth.

Table 6.1 lists all the strains and culturing conditions used to prepare purified spore suspensions of other spore-forming species.

Table 6.1: Strains and applied culturing conditions from which spore suspensions were prepared, along with the abbreviations used in the text.

Sample name	Species	Strain	Peroxide treatment	Growth medium	Temperature (°C)
MB372 perox	*Bacillus sporothermodurans* HRS	MB 372	10% H_2O_2	BHI	37
MB1313 perox		MB 1313	10% H_2O_2	BHI	37
MB385 perox	*Bacillus sporothermodurans*	MB 385	10% H_2O_2	BHI	37
MB385 CCY 30°C			No	CCY[28]	30
MB385 CCY 37°C			No	CCY	37
MB385 CCY 42°C			No	CCY	42
MB385 milk			No	milk agar	37
MB385 ut			No	Sporulation medium[22]	37
MB1317 perox		MB 1317	10% H_2O_2	BHI	37
MB1632 perox	*Bacillus cereus*	MB 1632	15% H_2O_2	BHI	37
MB1632 ut			No	Sporulation medium	37
MB1632 CCY			No	CCY	37
MB1634 perox		MB 1634	15% H_2O_2	BHI	37
MB1634 ut			No	Sporulation medium	37
MB1928 perox	*Paenibacillus lactis*	MB 1928	10% H_2O_2	BHI	37
MB1928 ut			No	Sporulation medium	37
subt CCY	*Bacillus subtilis*	unknown	No	CCY	37

ut = untreated
HRS = highly heat resistant spores
BHI = brain heart infusion medium (Oxoid)

6.2.2 Sample preparation

For *B. licheniformis* LMG 7634, part of a single colony was transferred to a CaF_2 plate with a 1 µL smear loop. The smears were dried for 3 min on silica.

The preparation of the purified spore suspensions was described by Scheldeman *et al.*[21]. The suspensions were centrifuged at 4 °C and 11000 rpm for 5 min. Part of the pellet was transferred with a 1 µL smear loop to a CaF_2 plate. A total of five smears per spore suspension were made and dried on silica for 5 min.

6.2.3 Products

L-cysteine and 2,6-pyridinedicarboxylic acid (dipicolinic acid; DPA) were purchased from Sigma-Aldrich (Belgium). Calcium dipicolinate (CaDPA) was synthesized from DPA as described in Ghiamati *et al.*[7].

6.2.4 Raman spectroscopy

The Raman spectra were recorded with a System Hololab 5000R modular Raman microspectrometer (Kaiser). The 785 nm laser light from a diode laser (Toptica Photonics AG) was focused through a 100x objective of the microscope (Leica) to obtain a power of 50 to 60 mW at the sample. The scattered light was transferred to the spectrograph by a confocal aperture collection fiber (15 µm numerical aperture) where it was detected by a back illuminated deep depletion Pelletier cooled (-70 °C) CCD detector (Andor). The Raman signal was collected in the spectral interval between 150 and 3500 cm^{-1}, but only the region between 365 and 1780 cm^{-1} was used, while the spectral resolution was approximately 4 cm^{-1}. On each smear of bacteria or spores on the CaF_2 plate, a central focus point was fixed around which four Raman spectra of 60 s were collected.

MATLAB (The Mathworks, Natick, MA, USA) was used for data preprocessing. Calibration was performed as described by Hutsebaut *et al.*[9], including (i) absolute wavelength calibration with a neon lamp, (ii) intensity calibration with a tungsten bulb, (iii) relative

wavelength calibration with seven reference products, (iv) correction for dark noise and (v) correction for the contribution of optical components. Extended multiplicative signal correction (EMSC) was applied[14] using the EMSC toolbox developed by Martens[13]. More specifically, the Raman spectra were treated with DataCase 103 which includes a correction for physical interferences. Spikes were omitted from the spectra and the average of each set of four spectra collected around one focus point was calculated. The plotted spectra are normalized. For the calculation of band intensities, a background subtraction was performed. Autoscaling and principal component analysis (PCA) were performed to study the possibility of discriminating different spore suspensions or *B. licheniformis* samples of different incubation times. From a comparison of a principal component (PC) and its corresponding factor in the loading plot, changes in Raman bands could be related to certain samples, and thus to certain incubation times. For the spore suspensions, the first 20 PCs were subjected to a cluster analysis in SPSS (Chicago, IL, USA) using squared Euclidean distance and Ward algorithm.

6.3 Results and discussion

6.3.1 Monitoring sporulation of Bacillus licheniformis LMG 7634

For the Raman measurements, *B. licheniformis* LMG 7634 was cultured on TSA for 1 to 6 days, because microscopic study showed that after 6 days of culturing the maximum spore yield was reached. Raman spectra were recorded from at least six independent cultures per incubation time. In this way physiological variation due to differences in microbial growth was covered. Per incubation period, an average value of the recorded spectra was calculated (Fig. 6.1). These spectra are further refered to as ‘mean spectra’. The most remarkable feature in these spectra is the band at 1018 cm^{-1} that rises with increasing incubation time. In order to visualize additional differences between sporulating and non-sporulating cells, a difference spectrum was calculated between the mean spectra obtained after 6 days and 1 day of incubation, respectively (Fig. 6.2). Comparison of this difference spectrum and a reference spectrum of CaDPA shows that this compound can contribute in

a bacterial spectrum to bands at approximately 661, 822, 1018, 1396, 1448 and 1575 cm^{-1} (Fig. 6.2). Band positions of CaDPA can differ by several wavenumber units between the reference spectrum and the bacterial spectra, because of a different physicochemical state of CaDPA in bacteria and in the reference product. CaDPA is one of the unique components from the spore core and represents about 5 – 10% of the dry weight of *Bacillus* spores[25]. The mean spectra also show bands of tyrosine, phenylalanine, fats and bases of DNA/RNA, but some of these may overlap with bands of CaDPA, when present. Band assignments are summarized in Table 6.2.

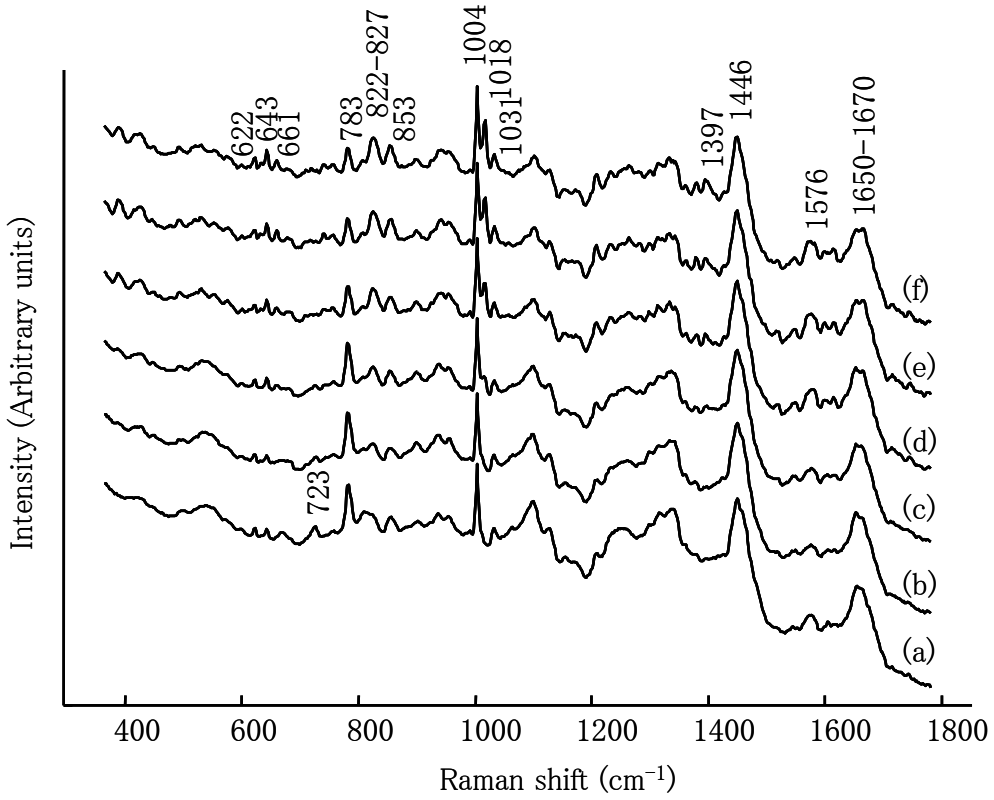


Figure 6.1 Mean Raman spectra of *B. licheniformis* LMG 7634 (TSA, 37°C) per incubation time: (a) 1 day, (b) 2, (c) 3, (d) 4, (e) 5 and (f) 6 days.

To study the evolution of Raman bands over all applied incubation times in more detail, autoscaling and subsequent PCA were performed on the pre-processed Raman spectra (Fig. 6.3). From the score plot, it can be derived that longer incubation times have higher scores for the first PC, which is indicated by the arrow in Figure 6.3a. Despite this observation, no distinct groups could be delineated according to incubation time, because of the microbial variation between different cultures at the same incubation time.

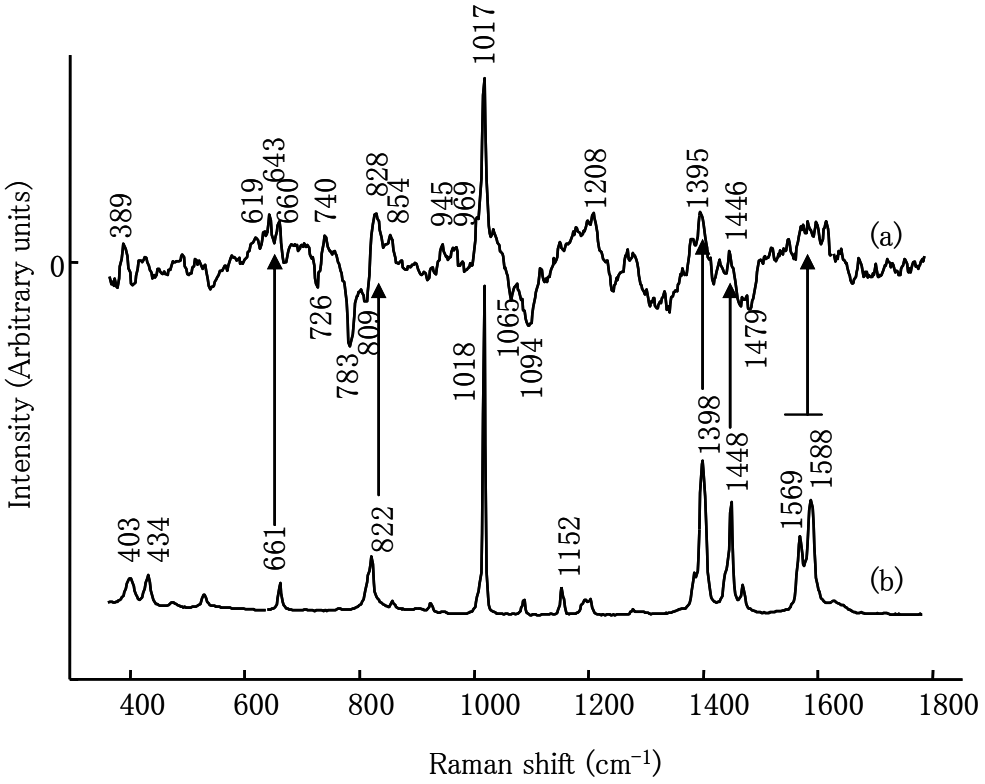


Figure 6.2 (a) Difference spectrum between the mean Raman spectra of *B. licheniformis* LMG 7634 cultured for 6 days and 1 day and (b) reference Raman spectrum of calcium dipicolinate.

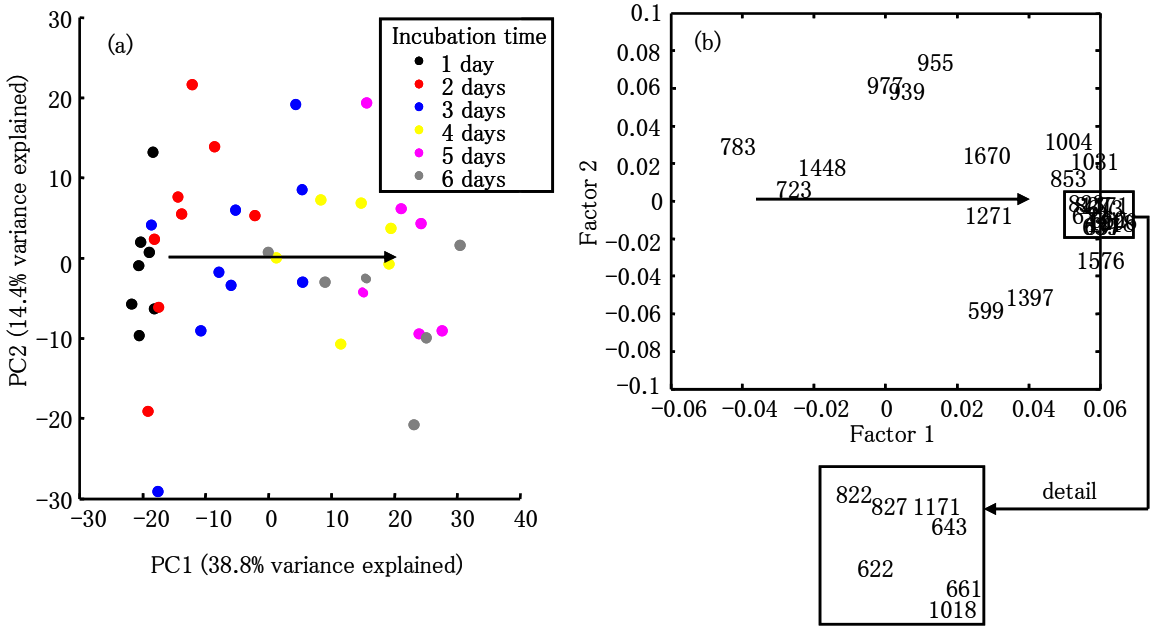


Figure 6.3 PCA analysis performed on the Raman spectra of *B. licheniformis* LMG 7634: (a) score and (b) loading plot of the first two principal components. The arrow indicates the trend in incubation time.

Table 6.2 Band assignments for Raman spectra of endospores (Fig. 6.5) and endospore forming bacteria (Fig. 6.1).

Raman shift (cm^{-1})*	Band assignments
527	ν(S–S) stretching (cysteine in spore coat)+
622	Phe
638	ν(C–S) stretching (cysteine in spore coat)+
643	Tyr
661	CaDPA
723	A, CoA, acetyl–CoA
783	C, U, citric acid
822–827	CaDPA and Tyr
853	Tyr
1004	Phe
1018	CaDPA
1031	ν(PO_2^-) symmetrical stretching
1250–1300	Amide III, δ(CH_2) deformation
1397	CaDPA
1448	CaDPA, δ($C–H_2$) deformation (e.g. from fats)
1576	CaDPA, G, A
1616	Tyr
1650–1680	Amide I

Apart from CaDPA and cysteine, band assignments are based on literature[5,15].
*No relative band intensities are given because they vary widely between the different bacteria/spore spectra.
+Only visible in some of the spore spectra, not in bacterial spectra.

A loading plot was constructed with Raman shifts of those bands that are clearly visible in the spectra (Fig. 6.3b). Comparison of the score and loading plots provides information about the Raman bands that are responsible for the differences between the spectra over the six incubation periods. Raman shifts that have a high score for factor 1 in the loading plot contribute highly in the spectra that have a high score for PC1 in the score plot, and vice versa (indicated in Figure 6.3 by the arrows in the score and loading plots that point in the same direction). Thus, the Raman bands with a high score for factor 1 are more intense in the spectra of cultures that were incubated longer. Therefore, it can be concluded from the loading plot that the Raman bands of CaDPA (661, 822, 1018, 1397 and 1576 cm^{-1}), phenylalanine (622, 1004 cm^{-1}) and tyrosine (643, 827, 853 cm^{-1}) contribute more in the spectra of cultures that were incubated longer. These findings show that Raman spectroscopy can be applied to monitor sporulation by CaDPA production, as

shown in Figure 6.4, where band intensities of the most intense band of CaDPA (1018 cm^{-1}) are plotted as a function of incubation time. The variation of the band intensity for each incubation time is due to heterogeneity of the spore production process within one culture as well as between different cultures. This variation is larger at longer incubation times. The increase in band intensities of phenylalanine and tyrosine might be explained by the production of small acid-soluble proteins (SASPs) that protect DNA in the spore core.

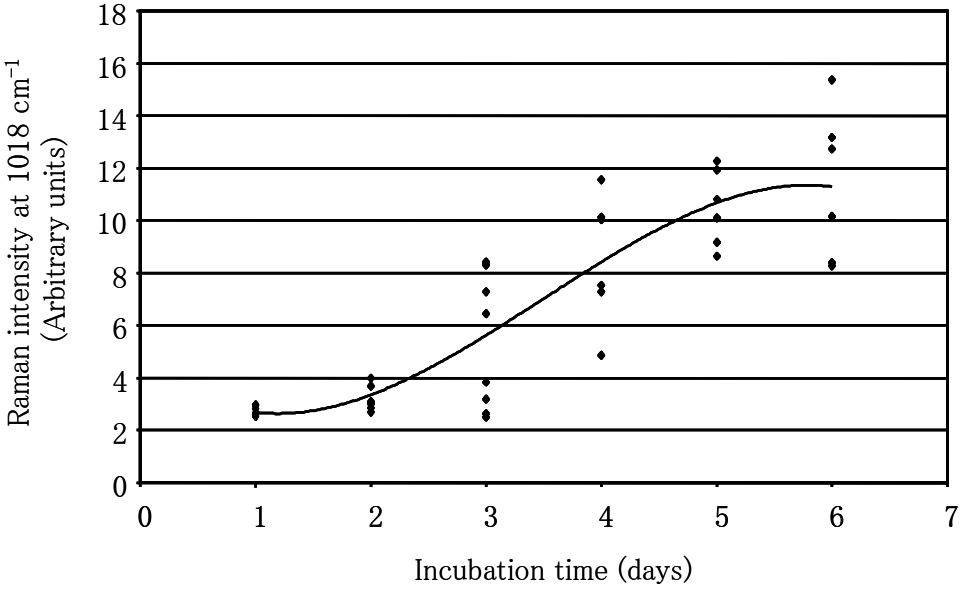


Figure 6.4 Band intensities of the most intense band of CaDPA (at 1018 cm^{-1}) versus incubation time for Raman spectra of *B. licheniformis* LMG 7634.

Analogously, Raman bands that have a strongly negative score for factor 1 are more prominent in Raman spectra of cultures with short incubation periods. They include bands at 723 cm^{-1} (adenine (A), coenzyme A (CoA), acetyl-CoA) and 783 cm^{-1} (guanine (G), uracil (U), citric acid). The observation that the band at 1448 cm^{-1} also has a negative score for factor 1 indicates that this band changes mainly owing to $\delta(C-H_2)$ deformation vibrations from e.g. fats and to a lesser extent owing to CaDPA production. The bands showing a negative score for factor 1 thus reflect the ongoing metabolism during exponential growth, which slows down at longer incubation times when spores are formed.

6.3.2 Discrimination between endospores from different species, strains or cultivation conditions

The strains, from which 17 endospore suspensions were obtained, are listed in Table 6.1 along with their culturing conditions. Raman spectra were recorded from five bacterial smears that were prepared for each spore suspension. The average spectra for each spore suspension were calculated and are shown in Figure 6.5a-d.

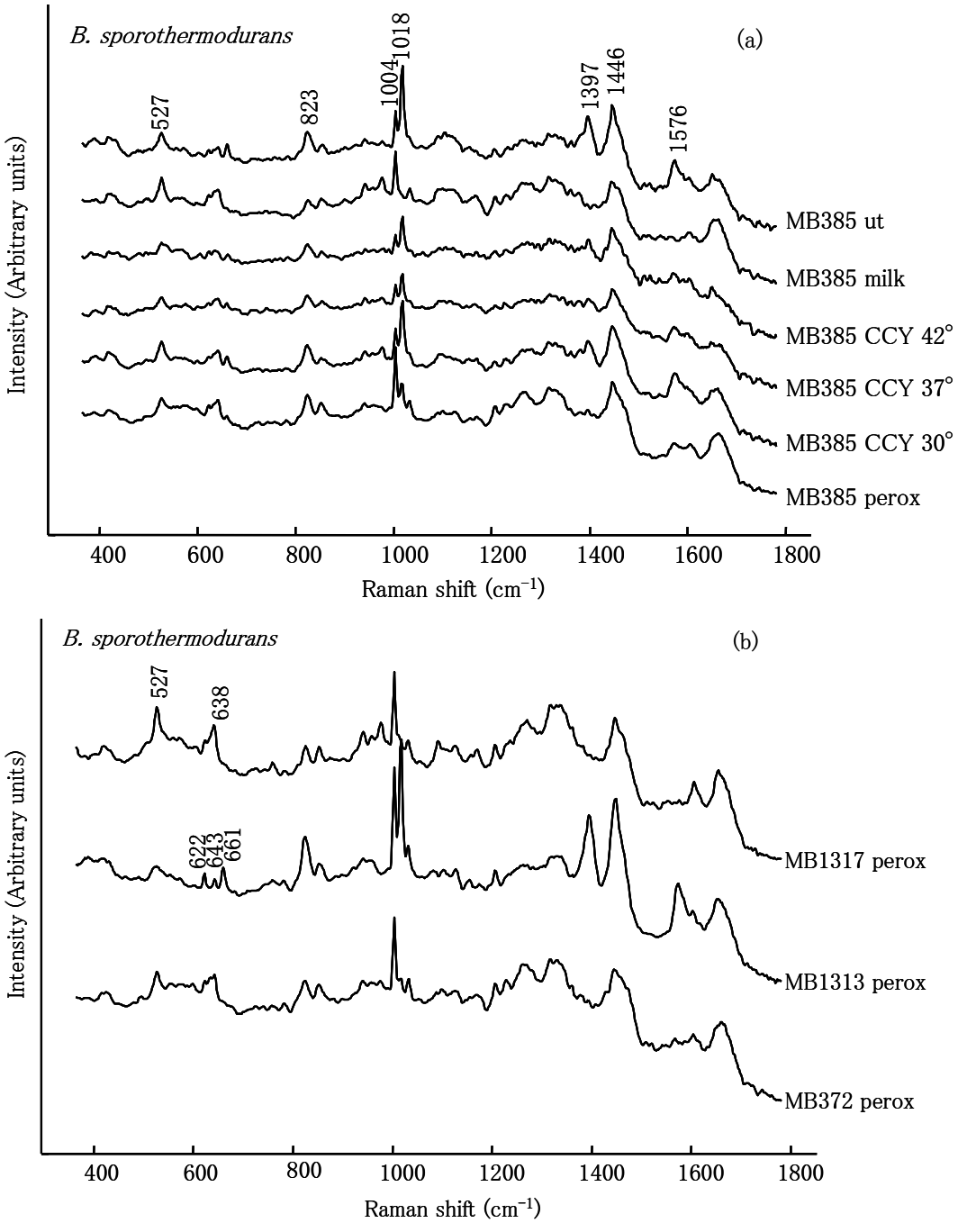


Figure 6.5 Mean Raman spectra of spore suspensions of (a) *B. sporothermodurans* MB385, (b) other *B. sporothermodurans* strains. Specifications for the labels in these spectra are listed in Table 6.1

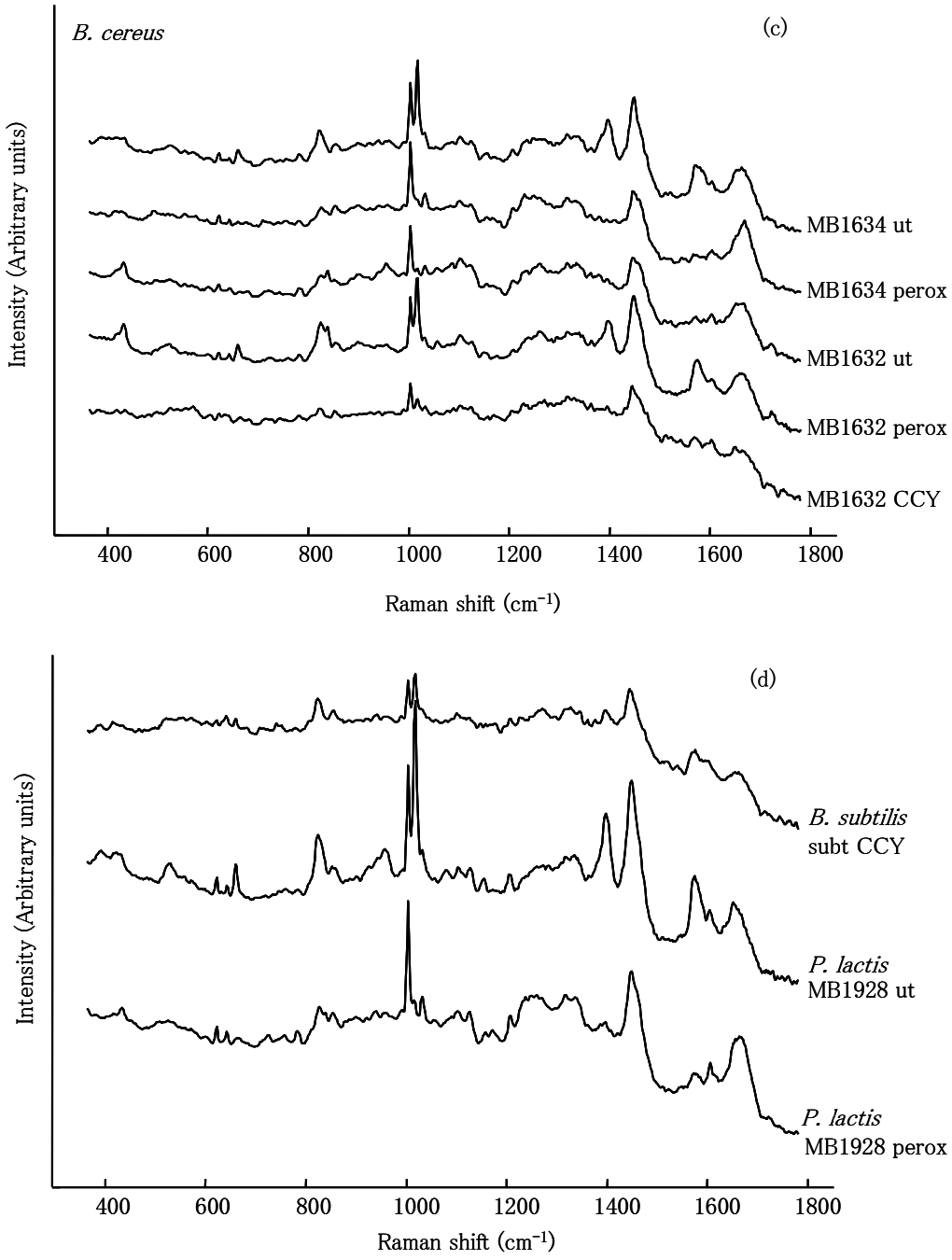


Figure 6.5 Mean Raman spectra of spore suspensions of (c) *B. cereus*, (d) *P. lactis* and *B. subtilis*. Specifications for the labels in these spectra are listed in Table 6.1.

These spectra show a large variation in the intensity of the CaDPA band at 1018 cm^{-1}. When CaDPA is present in large amounts, there is also a remarkable contribution in the bacterial spectra at other Raman band positions such as 661, 822, 1397, 1448 and 1576 cm^{-1}. From Table 6.2, it is clear that the bands at 822 and 1448 cm^{-1}, may also have a contribution from or overlap with bands of respectively tyrosine (827 cm^{-1}) and other CH_2

deformations from e.g. fats (1448 cm^{-1}). The observation that the bands at 822 and 1448 cm^{-1} in bacterial spectra rise together with the other bands of CaDPA, confirms the assumption that the bands at 822 and 1448 cm^{-1} contain a significant contribution from CaDPA. Other prominent bands are those at 622 and 1004 cm^{-1} from phenylalanine, and at 643, 827 and 853 cm^{-1} from tyrosine. In some spectra, for example that of MB1317 perox (Fig. 6.5b), two bands appear at 527 and 638 cm^{-1}. In these regions, sulphur related vibrations can occur such as δ(C–S) bending, ν(S–S) stretching and ν(C–S) stretching[27]. According to Aronson and Fitz-James[1], cysteine residues can be incorporated in spore coat proteins during late sporulation by disulfide interchange between cystine or cysteine and proteins with intermolecular disulfide bonds. In the reference spectrum of cysteine (data not shown) the most intense bands are located at 443 and 639 cm^{-1} and can be assigned to δ(C–S–H) out of plane bending and ν(C–S) stretching[27]. By incorporating cysteine in proteins, the S–H bond and thus the corresponding Raman band at 443 cm^{-1} disappears and is replaced by an S–S bond which causes a stretching band at 527 cm^{-1}. In the fingerprint region of the spore spectra, bands of tyrosine, phenylalanine, CaDPA and incorporated cysteine may (partly) overlap. Therefore, Table 6.2 gives an overview of the possible contributions of several biomolecules in Raman spectra of endospore forming bacteria and spore suspensions.

In order to study the similarity/dissimilarity of spore spectra between different strains and/or culturing conditions, PCA was performed on the pre-processed data. The first 20 PCs were selected for cluster analysis, the dendrogram of which is given in Figure 6.6. For some samples, e.g. MB1632 ut, all five Raman spectra (obtained from different smears of one spore suspension) are clustered in the same subgroup, so the composition of the spores in these suspensions is regarded as homogenous. For the samples MB385 perox and MB1632 perox, one of the five Raman spectra is considered as an outlier (marked in Figure 6.6 with an asterisk). This suggests that these samples show a higher heterogeneity in the composition of the spores within the particular spore suspensions. This is in concordance with a recent study by Huang *et al.*[8] which reports heterogeneity in CaDPA concentrations amongst single spores of a population. The dendrogram shows that, on the basis of Raman spectra, the spore suspensions can be divided in two main groups (A and B).

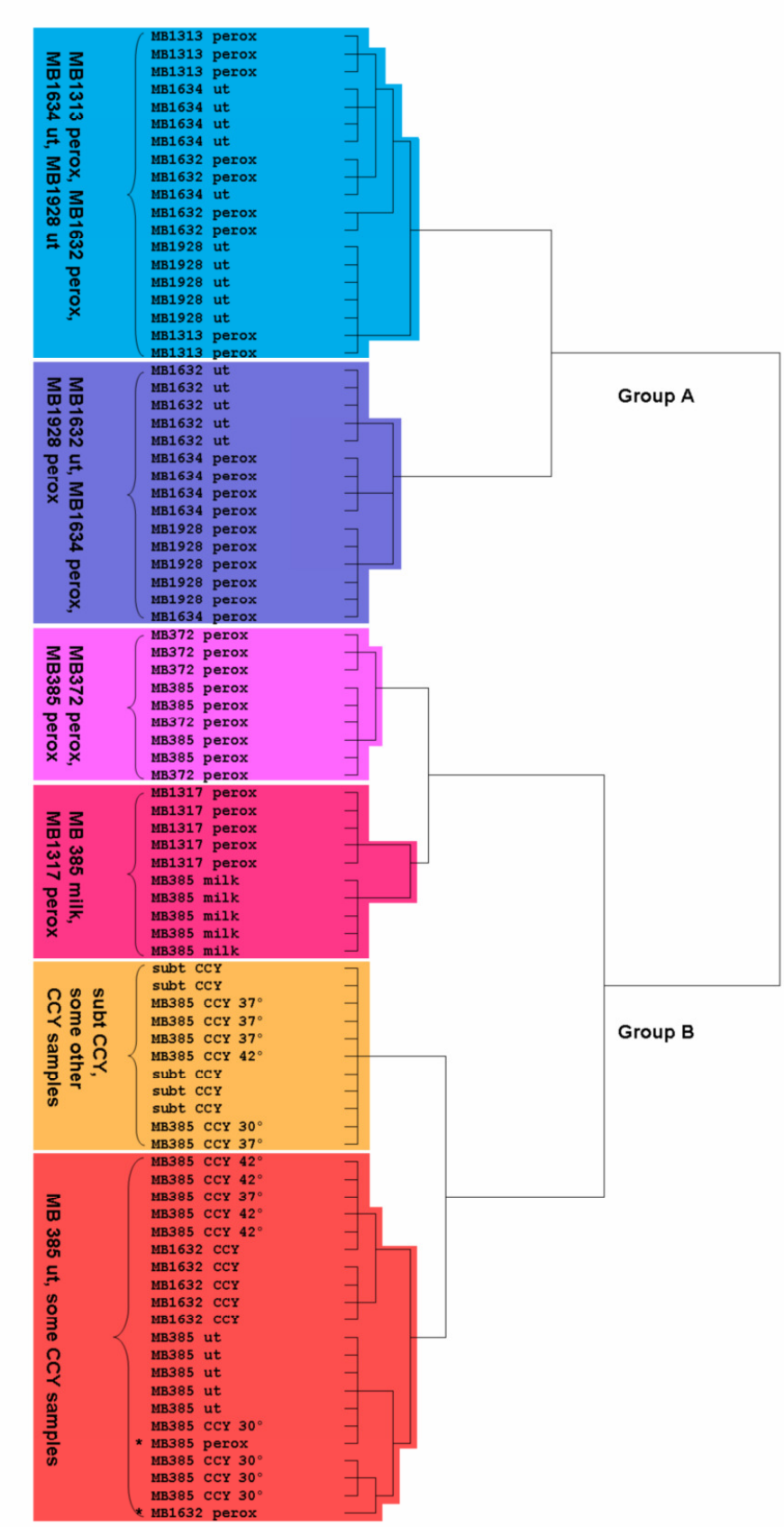


Figure 6.6 Dendrogram obtained from cluster analysis of the first 20 principal components of the PCA analysis performed on the Raman spectra of spore suspensions. Specifications for the labels in these spectra are listed in Table 6.1.

The key factor for this separation is a combination of the strain differences and the growth medium (casein hydrolysate and yeast extract (CCY) or not). Indeed, all samples of the strains *B. sporothermodurans* MB1313, *B. cereus* MB1632 and MB1634 and *Paenibacillus lactis* MB1928 are gathered in group A (bluish boxes), except for the sample MB1632 CCY which is located in group B (reddish boxes) together with the other CCY samples plus all samples of the *B. sporothermodurans* strains MB372, MB385 and MB1317 and the *B. subtilis* strain. The subdivision of group A (two bluish boxes in Figure 6.6) cannot be linked to certain strains or culturing conditions. In group B (four reddish boxes in Figure 6.6), one subdivision is formed containing the spore samples of strains cultured on CCY and the untreated sample of MB385. The other subdivision within group B consists of three *B. sporothermodurans* strains that were peroxide stressed and of MB385 sporulated on milk agar.

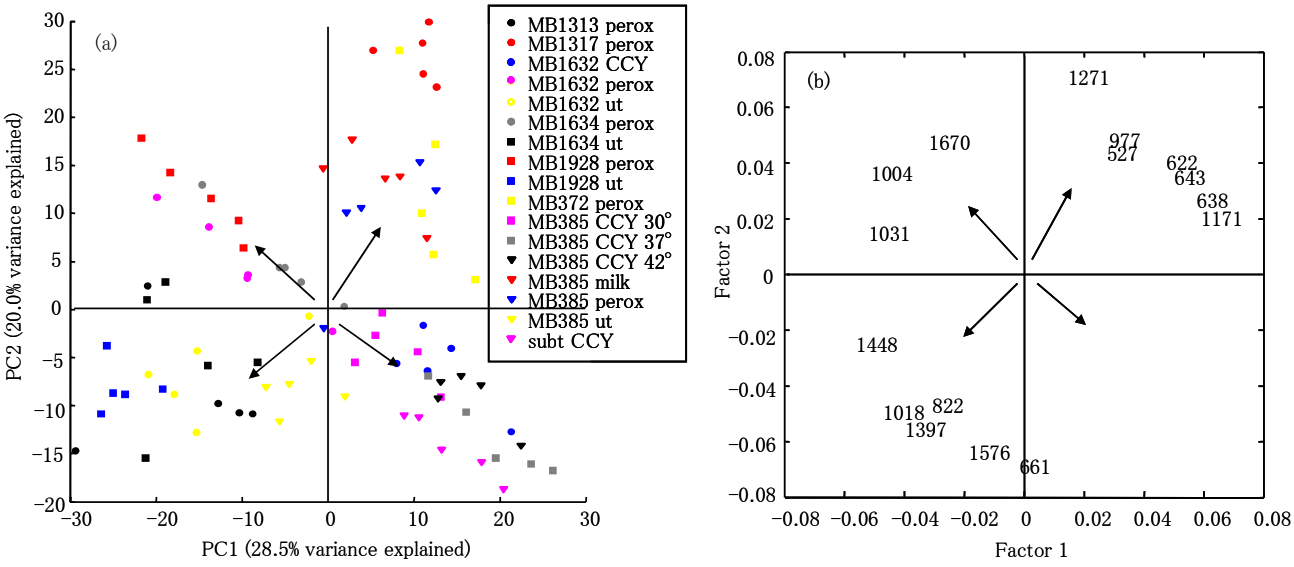


Figure 6.7 (a) Score and (b) loading plot of the first 2 principal components of the PCA analysis performed on the Raman spectra of spore suspensions.

From this cluster analysis (Fig. 6.6), it is clear that according to the Raman spectra, the spore suspensions can be divided in groups. In order to search for markers that form the basis of this division, we compare the score plot of the first two PC's with the corresponding loading plot (Fig. 6.7a,b). The Raman shifts plotted in the loading plot, are those of the most prominent bands in the Raman spectra. According to the four quadrants

in the score and loading plot (indicated by arrows in Figure 6.7a,b), the composition of four groups of samples is summarized in Table 6.3.

Table 6.3 Summary of the information about the composition of the spore suspensions: i) groupings were based on the four quadrants in score plot of the PCA analysis (indicated by arrows in Fig. 6.7a), ii) comparison of the score and loading plots (Fig. 6.7a and b) gave information about the presence of certain Raman bands, which was confirmed and/or completed by studying the mean spectra (Fig. 6.5).

Group	Quadrant: PC1/2 or factor1/2	Samples in score plot	Raman shifts (cm^{-1}) in loading plots	Biomolecule information
I	+/+	mb372 perox mb385 perox mb385 milk mb1317 perox	527,638 622 643 977 1171 1271	Prominent bands of cysteine incorporated in spore coat proteins. Fingerprint region shows many bands. Weak CaDPA bands.
II	+/−	mb385 CCY 30° C mb385 CCY 37° C mb385 CCY 42° C mb1632 CCY subt CCY		No prominent bands. Spectra show little detail and have a background that falls back at 1400 cm^{-1}. Bands of phenylalanine and CaDPA are moderate present in the spectra.
III	−/−	mb385 ut mb1313 perox mb1632 perox mb1634 ut mb1928 ut	661,822,1018, 1397,1448, 1576	Bands of CaDPA are present in the spectra in high intensity.
IV	−/+	mb1632 ut mb1634 perox mb1928 perox	1004 1031 1670	No prominent bands in the fingerprint region. Weak CaDPA bands.

From this table, it is clear that Raman spectra of spores cultured on CCY medium (group II) are highly similar, even for different strains and species. They show no prominent bands in the fingerprint region and have a moderate intensity for phenylalanine and CaDPA (Fig. 6.5a,c). The spectra show higher noise than spectra derived from spores that were not harvested from CCY medium and their background is rather flat up to 1400 cm^{-1} from which point on it drops. These effects on noise and background could be caused by a thicker or denser spore coat. Group III is characterized by a large amount of CaDPA, while

the spectra of group I and IV show only weak bands of CaDPA (Fig. 6.5). This observation of large variations in CaDPA content amongst spores of different species and strains was confirmed by Huang *et al.*[8]. In addition, from these groupings, it can be concluded that there is no unambiguous correlation between a peroxide treatment and the amount of CaDPA present in the spores of different species or strains. Nevertheless, compared to untreated samples, there is always a remarkable increase or decrease of the band intensity at 1018 cm^{-1} when a peroxide treatment was applied. Indeed, the Raman spectra of the samples MB385 ut, MB1313 perox, MB1632 perox, MB1634 ut and MB1928 ut (Fig. 6.5) show high intensities at 1018 cm^{-1} (CaDPA), while the spectra of MB372 perox, MB385 perox, MB1317 perox, MB1632 ut, MB1634 perox and MB1928 perox (Fig. 6.5) show lower intensities for CaDPA. The spectra of group I (samples MB372 perox, MB385 perox, MB385 milk and MB1317 perox) are characterised by prominent bands at 527 and 638 cm^{-1} that can be assigned to ν(S-S) and ν(C-S) stretching vibrations of cysteine residues incorporated in spore coat proteins. The bands of cysteine are only visible in the spectra of some of the spore suspensions of *B. sporothermodurans* (Fig. 6.5a,b). We stress that the division in Table 6.3 reflects similarities between Raman spectra of the spore suspensions. Thus, this division is based on the spore components that contribute significantly in the Raman spectra, rather then on the over-all spore composition.

6.4 Conclusions

In this work, Raman spectra of endospores and endospore-forming bacteria were studied. The spectra of *B. licheniformis* LMG 7634 recorded at incubation times of 1 to 6 days, show trends in band changes as a function of incubation time. Not only indications of a lower metabolism, but also band increases due to endospore formation could be observed. The most important changes during sporulation include the rise of bands at ca. 661, 822, 1018, 1397, 1448 and 1576 cm^{-1}, which can be assigned to CaDPA, a unique endospore core component. Sporulation of bacteria can easily be monitored by the band intensity at 1018 cm^{-1}. Bands from phenylalanine (622 and 1004 cm^{-1}) and tyrosine (643, 827 and

853 cm^{-1}) also increase as a function of the incubation time, which could be explained by the formation of SASPs during sporulation.

Seventeen spore suspensions were obtained from different species, strains and culturing conditions. Their Raman spectra vary widely in intensity of bands from biomolecules such as phenylalanine, tyrosine and calcium dipicolinate. Cluster analysis shows that even the spore composition within a suspension, obtained from a certain strain cultured under a certain condition, can be heterogeneous. Two bands at 527 and 638 cm^{-1} occur in some spectra of *B. sporothermodurans* spore suspensions and were assigned to cysteine residues incorporated in spore coat proteins. These variations suggest that important differences in spore composition are possible. A recent study by Huang *et al.*[8] confirms that Raman spectroscopy can detect CaDPA variations within a population and between different species and strains, but no information about other spore components could be deduced from their spectra. PCA shows that the Raman spectra of spores are not only influenced by the strain or species, but also by the culturing medium. Indeed, Raman spectra of spore suspensions obtained from cultures on CCY medium form a distinct group in the PCA score plot. These spectra have similar intensities at 1018 cm^{-1} (CaDPA) and show little detail in the fingerprint region. Peroxide treatment seems to induce large changes in CaDPA concentrations compared to untreated samples of the same strains, but no consistent increases or decreases of CaDPA are observed between these culturing conditions for the studied strains. In general, Raman spectroscopy is suitable for the study of the composition of endospores, especially for CaDPA and amino acids such as tyrosine, phenylalanine and cysteine.

This study shows that bacterial Raman spectra, combined with reference spectra, can be used to study differences in cell or endospore content between different populations (different species, strains or culturing conditions). Indeed, biomolecules that show a high and varying signal in bacterial Raman spectra cause clear bands in their difference spectra. Studying band positions in loading plots and comparing them to score plots of a principal component analysis (PCA), leads to conclusions about the samples' composition and the

degree of chemical similarity between different samples. This demonstrates that Raman spectroscopy is suitable for studying the bacterial cell content. However, bacterial Raman spectra reflect mainly the presence of good Raman scatterers and/or molecules that are present in the cell in high concentrations. Therefore, the presence or change in concentration of minor components or molecules with weak Raman activity can be masked. Well distinguished bands in bacterial Raman spectra are often related to specific chemical bonds or structures, for example aromatic rings (purine and pyrimidine bases of DNA and RNA, and CaDPA) or organosulphur bonds (cysteine).

Although this chapter demonstrates that the use of difference spectra and PCA are suitable to study Raman spectra of sporulating bacteria or endospores, these approaches are insufficient to study compounds with minor contributions to overall bacterial Raman spectra. In addition, a lot of signals from specific compounds overlap, resulting in bacterial Raman spectra where most bands contain contributions of more than one biomolecule. To approach this phenomenon, more complex processing methods should be applied, in order to extract information about cell compounds with a minor contribution or from overlapping bands. In the next chapter, a few methods are illustrated in order to study Cupriavidus metalidurans LMG 1195 in several stages of its growth.

References

1. Aronson A.I., Fitz-James P. (1976) Structure and morphogenesis of the bacterial spore coat. *Bacteriol. Rev.* 40: 360-402.
2. Bell S.E.J., Mackle J.N., Sirimuthu N.M.S. (2005) Quantitative surface-enhanced Raman spectroscopy of dipicolinic acid - towards rapid anthrax endospore detection. *Analyst* 130: 545-549.
3. Chan J.W., Esposito A.P., Talley C.E., Hollars C.W., Lane S.M., Huser T. (2004) Reagentless identification of single bacterial spores in aqueous solution by confocal laser tweezers Raman spectroscopy. *Anal. Chem.* 76: 599-603.
4. Daniels J.K., Caldwell T.P., Christensen K.A., Chumanov G. (2006) Monitoring the kinetics of *Bacillus subtilis* endospore germination via surface-enhanced Raman scattering spectroscopy. *Anal. Chem.* 78: 1724-1729.
5. De Gelder J., De Gussem K., Vandenabeele P., De Vos P., Moens L. (2007) Methods for extracting biochemical information from bacterial Raman spectra: An explorative study on *Cupriavidus metallidurans*. *Anal. Chim. Acta* 585: 234-240.
6. Farquharson S., Grigely L., Khitrov V., Smith W., Sperry J.F., Fenerty G. (2004) Detecting *Bacillus cereus* spores on a mail sorting system using Raman spectroscopy. *J. Raman Spectrosc.* 35: 82-86.
7. Ghiamati E., Manoharan R., Nelson W.H., Sperry J.F. (1992) UV resonance Raman spectra of *Bacillus* spores. *Appl. Spectrosc.* 46: 357-364.
8. Huang S-s., Chen D., Pelczar P.L., Vepachedu V.R., Setlow P., Li Y-q. (2007) Levels of Ca^{2+}-dipicolinic acid in individual *Bacillus* spores determined using microfluidic raman tweezers. *J. Bacteriol.* 189: 4681-4687.
9. Hutsebaut D., Vandenabeele P., Moens L. (2005) Evaluation of an accurate calibration and spectral standardization procedure for Raman spectroscopy. *Analyst* 130: 1204-1214.
10. Hutsebaut D., Vandroemme J., Heyrman J., Dawyndt P., Vandenabeele P., Moens L., De Vos P. (2006) Raman microspectroscopy as an identification tool within the

phylogenetically homogeneous '*Bacillus subtilis*'-group. *Syst. Appl. Microbiol.* 29: 650-660.

11. Kirschner C., Maquelin K., Pina P., Thi N.A.N., Choo-Smith L.P., Sockalingum G.D., Sandt C., Ami D., Orsini F., Doglia S.M., Allouch P., Mainfait M., Puppels G.J., Naumann D. (2001) Classification and identification of *enterococci*: a comparative phenotypic, genotypic, and vibrational spectroscopic study. *J. Clin. Microbiol.* 39: 1763-1770.
12. Lopez-Diez E.C., Goodacre R. (2004) Characterization of microorganisms using UV resonance Raman spectroscopy and chemometrics. *Anal Chem* 76: 585-591.
13. Martens H., The EMSC toolbox for MATLAB http://www.models.kvl.dk/source/emsctoolbox
14. Martens H., Nielsen J.P., Engelsen S.B. (2003) Light scattering and light absorbance separated by extended multiplicative signal correction. Application to near-infrared transmission analysis of powder mixtures. *Anal. Chem.* 75: 394-404.
15. Maquelin K., Krischner C., Choo-Smith L.P., van den Braak N., Endtz H.Ph., Naumann D., Puppels G.J. (2002) Identification of medically relevant microorganisms by vibrational spectroscopy. *J. Microbiol. Meth.* 51: 255-271.
16. Naumann D., Keller S., Helm D., Schultz C., Schrader B. (1995) FT-IR spectroscopy and FT-Raman spectroscopy are powerful analytical tools for the noninvasive characterization of intact microbial cells. *J. Mol. Struct.* 347: 399-504.
17. Nelson W.H., Dasari R., Feld M., Sperry J.F. (2004) Intensities of calcium dipicolinate and *Bacillus subtilis* spore Raman spectra excited with 244 nm light. *Appl. Spectrosc.* 58: 1408-1412.
18. Oust A., Moretro T., Naterstad K., Sockalingum G.D., Adt I., Manfait M., Kohler A. (2006) Fourier transform infrared and Raman spectroscopy for characterization of *Listeria monocytogenes* strains. *App. Environ. Microb.* 72: 228-232.
19. Petrov G.I., Yakovlev V.V., Sokolov A.V., Scully M.O. (2005) Detection of *Bacillus subtilis* spores in water by means of broadband coherent anti-Stokes Raman spectroscopy. *Opt. Express* 13: 9537-9542.

20. Riesenman P.J., Nicholson W.L., (2000) Role of the spore coat layers in *Bacillus subtilis* spore resistance to hydrogen peroxide, artificial UV-C, UV-B, and solar UV radiation. *Appl. Environ. Microb.* 66: 620-626.
21. Scheldeman P. (2004) PhD dissertation: Occurrence and resistance of potentially highly heat resistant spore forming bacteria in milk products and at dairy farms. Ghent University, Department of Biochemistry, Physiology and Microbiology.
22. Scheldeman P., Herman L., Foster S., Heyndrickx M. (2006) *Bacillus sporothermodurans* and other highly heat-resistant spore formers in milk. *J. Appl. Microbiol.* 101: 542-555.
23. Setlow B., Atluri S., Kitchel R., Koziol-Dube K., Setlow P. (2006) Role of dipicolinic acid in resistance and stability of spores of *Bacillus subtilis* with or without DNA protective α/β-type small acid-soluble proteins. *J. Bacteriol.* 188: 3740-3747.
24. Setlow B., Setlow P. (1993) Binding of small, acid-soluble spore proteins to DNA plays a significant role in the resistance of *Bacillus subtilis* spores to hydrogen peroxide. *Appl. Environ. Microbiol.* 59: 3418-3423.
25. Setlow P. (1995) Mechanisms for the prevention of damage to DNA in spores of *Bacillus species*. *Annu. Rev. Microbiol.* 49: 29-54.
26. Setlow P. (2006) Spores of *Bacillus subtilis*: their resistance to and killing by radiation, heat and chemicals. *J. Appl. Microbiol.* 101: 514-525.
27. Socrates G. (2001) Infrared and Raman characteristic group frequencies. Third edition, John Wiley and Sons Ltd, Chichester.
28. Stewart G.S.A.B., Johnstone K., Hagelberg E., Ellar D.J. (1981) Commitment of bacterial spores to germinate - a measure of the trigger reaction. *Biochem. J.* 198: 101-106.

Chapter 7: Methods for extracting biochemical information from bacterial Raman spectra: An explorative study on *Cupriavidus metallidurans* LMG 1195

Joke De Gelder, Kris De Gussem, Peter Vandenabeele, Paul De Vos, Luc Moens

Analytica Chimica Acta, 585 (2007) 234-240

Comparison of reference spectra from our database (Chapter 5) with bacterial Raman spectra shows that many bands in bacterial spectra may contain contributions of several biomolecules. Methods like the calculation of difference spectra and PCA (applied in Chapter 6) are not able to provide information about the separate biomolecules that contribute to these composite bacterial Raman bands. Therefore, more enhanced approaches are required to extract information on individual biomolecules from the bacterial Raman spectra. These approaches are discussed in Chapter 7 and do not focus on a single characteristic band. Instead, they take the whole reference spectrum of a biomolecule into account to evaluate the biomolecule's presence in the cell. Next to two methods that estimate the contribution of reference spectra in bacterial Raman spectra (namely the calculation of dot products and EMSC coefficients), the potential of 2D correlation spectroscopy is illustrated to obtain information about changes in Raman bands (and thus biomolecules) following a perturbation. This is illustrated by studying changes in bacterial cell composition, due to different incubation times.

7.1 Introduction

Over the years, Raman spectroscopy has been developed and applied for microbiological applications. As Raman spectroscopy is a whole cell non-destructive fingerprinting technique, an important advantage is that biochemical information of all Raman-active cell components is present in bacterial spectra. There is also minor interference of water and only small samples are required. Several research groups have already worked out an identification procedure for bacteria[9,10]. Even closely related species within the *Enterococcus* group[8] and within the *Bacillus subtilis* group[6] could be distinguished by Raman analysis. For these purposes, Raman spectra are considered as mathematical data on which chemometric techniques, such as principal component analysis (PCA), linear discriminant analysis (LDA) and hierarchical cluster analysis (HCA), are applied, allowing discrimination at the species and even at the strain level[6,7,10,17].

Next to identification, it would also be very interesting to retrieve the biochemical information incorporated in the bacterial Raman spectra. Several Raman bands of bacterial spectra have already been assigned to common biomolecules such as phenylalanine (Phe), tyrosine (Tyr) and the pyrimidine and purine bases of DNA and RNA and certain regions are attributed to groups of molecules such as fatty acids and proteins[4,10,15]. Unfortunately, a lot of biochemical information in bacterial Raman spectra remains unassigned. Therefore, our aim is to take the first step in performing a more detailed interpretation of bacterial Raman spectra in order to extract more biochemical information about the cell's metabolism. We do not determine absolute concentrations of biomolecules, but focus on relative changes in certain biomolecules during five different stages of growth of *Cupriavidus metallidurans* LMG 1195. We focused on the primary metabolites of this species and their evolution during the various phases of growth. Several approaches are presented to retrieve this information from bacterial Raman spectra: (i) the interpretation of the average spectra per growth stage and the difference spectra between these average spectra, (ii) the use of dot product values between the bacterial Raman spectra and the reference spectra of biomolecules, (iii) use of coefficients obtained from an extended multiplicative signal correction (EMSC) to map out the production and/or consummation of biomolecules and (iv) the employment of two dimensional (2D) correlation spectroscopy to track which changes in Raman bands are related.

EMSC is a method developed to remove physical and chemical interferences from spectra[14]. In contrast with other chemometric techniques which model physical and chemical interferences separately, EMSC models them at once. It uses empirical knowledge such as spectra of the analytes and/or interferents. Application of this technique on biological material was described by De Decker *et al.*[2].

2D correlation spectroscopy[16] is a chemometric technique, which is especially skilled for modeling spectral changes related to an external factor, such as time. This chemometric technique results in 2D synchronous spectra that indicate which spectral channels are correlated to each other, while the complementary asynchronous spectra are used to determine the order in which spectral changes occur. As a result, 2D correlation analysis

has been used in many different research areas, including spectroscopic analysis of complex biomaterials[1,3,16].

7.2 Experimental

7.2.1 Strain and medium

Cupriavidus metallidurans LMG 1195 (formerly known as *Ralstonia metallidurans*, *Ralstonia eutropha*, *Alcaligenes eutrophus*) was cultured in a defined medium with sodium gluconate as carbon source (Table 7.1). The strain was cultured on Petri dishes to check purity, but liquid medium was used for the Raman experiments. An overnight grown liquid culture was diluted with medium to OD (590 nm) 0.25. Then 2 mL of this suspension was added to 250 mL Erlenmeyer flasks each containing 50 mL of the fluid medium. These Erlenmeyer flasks (OD (590 nm) approximately 0.01) were shaken at 28 °C for 16, 24, 48, 72 or 96 hours respectively.

7.2.2 Sample preparation

Out of each Erlenmeyer flask minimal 3 samples of 2 mL were pipetted into Eppendorf cups. The growth experiments were conducted at random for about 10 times during a period of two months in order to include biological and instrumental variations. In order to eliminate spectral interference of the medium, the samples were washed. Therefore, samples in the Eppendorf cups were centrifuged at 15000 g for 2 minutes. The supernatant was removed and 1 mL of physiological water was added. The Eppendorf cups were vortexed and the suspension was pipetted up and down three times before centrifugation (15000 g for 2 minutes). The same procedure was repeated once, with 10 minutes of centrifugation. After removing the supernatant, a dense pellet was obtained, which was transferred to a CaF_2 plate with a 1 μL smear loop. The smears were dried for 20 minutes with silica. This sample preparation allows us to obtain high quality Raman spectra from fluid cultures.

Table 7.1 Medium composition

Per liter bidistillated water:	
Tris/HCl	6.06 g (50 mmol L^{-1})
NaCl	4.68 g (80 mmol L^{-1})
KCl	1.49 g (20 mmol L^{-1})
NH_4Cl	1.07 g (20 mmol L^{-1})
Na_2SO_4	0.43 g (3 mmol L^{-1})
$MgCl_2.6H_2O$	0.20 g (1 mmol L^{-1})
$CaCl_2.2H_2O$	0.03 g (0,2 mmol L^{-1})
$Na_2HPO_4.2H_2O$	40 mg or 4ml from a 1% solution
Fe(III)NH_4citrate	10 mL from a solution of 0.48 g L^{-1}
SI 7 trace elements solution	1 mL
As a carbon source 0,2% (weight/volume) sodium gluconate is added. The solution is brought to pH 7 with concentrated HCl. For the preparation of solid medium 2% agar is added.	
Trace solution: per liter bidistillated water:	
25% HCl	1,3 ml (10 mmol L^{-1})
$ZnSO_4.7H_2O$	144 mg (0.5 mmol L^{-1})
$MnCl_2.4H_2O$	100 mg (0.5 mmol L^{-1})
H_3BO_3	62 mg (1 mmol L^{-1})
$CoCl_2.6H_2O$	190 mg (0.8 mmol L^{-1})
$CuCl_2.2H_2O$	17 mg (0.1 mmol L^{-1})
$NiCl_2.6H_2O$	24 mg (0.1 mmol L^{-1})
$Na_2MoO_4.2H_2O$	36 mg (0.15 mmol L^{-1})

7.2.3 Raman spectroscopy

All Raman spectra were recorded with a Kaiser System Hololab 5000R modular Raman microspectrometer. The 100x objective of the microscope (Leica) was used and the samples were excited using 45 – 50 mW (at the sample) of 785 nm laser light from a diode laser (Toptica Photonics AG). The scattered light is transferred to the spectrograph by a confocal aperture collection fiber (15 µm N.A.) where it was detected by a back illuminated deep depletion Pelletier cooled (-70 °C) CCD detector (Andor). The Raman signal was collected in the spectral interval 150 – 3500 cm^{-1}, but only the region between 365 and 1800 cm^{-1} was used, while the spectral resolution was approximately 4 cm^{-1}.

To record bacterial Raman spectra, on each smear of the CaF_2 plate a central focus point was chosen and around this center, 4 spectra of 60 s were collected. The smears were measured at random in order to eliminate systematic errors.

For the interpretation of bacterial Raman spectra we considered it useful to construct a database of Raman spectra of biomolecules (Chapter 5). We purchased some pure amino acids, purine and pyrimidine bases, fatty acids and components of the main metabolic pathways of *C. metallidurans* (Sigma-Aldrich, Belgium). At least three sets of 4 Raman spectra around 3 different focus points of each product were collected.

7.2.4 Data preprocessing

MATLAB was used for data preprocessing. The calibration was performed as described by Hutsebaut *et al.*[5], using (i) a neon lamp for absolute wavelength calibration, (ii) a tungsten bulb operating at 6.500A for intensity calibration, (iii) seven reference products for relative wavelength calibration, (iv) dark spectra and (v) optics spectra.

For subsequent interpretation of the spectra and spectral changes, extended multiplicative signal correction (EMSC) was applied[14] using the modeled EMSC toolbox developed by Martens[13]. In particularly, the bacterial Raman spectra used for the calculation of the dot-products were subjected to EMSC correction for the elimination of physical interference (datacase 103). In this procedure, multiplicative effects as well as baselines (which are brought to an equal level for all spectra) are modeled. For the calculation of mean spectra and difference spectra, and for the interpretation of trends in chemical composition using EMSC, datacase 106 was used which includes a physical EMSC correction as well as it uses an input from reference spectra of chemical compounds (referred to as 'Good spectra')[13]. In this procedure the average spectrum of the bacterial Raman data set was used as mal spectrum $\mathbf{m}$, while reference spectra from the biomolecule database were used as 'Good Spectra' input. The spectra of the biomolecule database were, for each reference product separately, subjected to EMSC using datacase 103. These spectra were averaged out to 1 reference spectrum k_i per product. In the EMSC procedure (datacase 106), the bacterial spectra were modeled according to the equation $\mathbf{z}_i = \mathbf{m}\, b_i + \mathbf{k}_1\, \Delta c_{i,1} + \cdots + \mathbf{k}_j\, \Delta c_{i,j} + \mathbf{1}\, a_i + \mathbf{e}_i$

where $\mathbf{z}_i$ represent the measured spectra, $\mathbf{m}$ the mal spectrum (here the average spectrum of the dataset), a_i the unknown additive effect, b_i the unknown multiplicative effect and $\Delta c_{i,j}$ values for the difference in concentration of component j between the bacterial spectrum i and the mal spectrum[14]. The coefficients $\Delta c_{i,j}$ were estimation by ordinary least squares (OLS) and used as a value for the contribution of the reference spectra in the bacterial Raman spectra. The coefficients were imported in SPSS where ANalysis Of VAriance (ANOVA) was performed to identify significant differences at the 95% confidential level.
Spikes were removed from the spectra and the average of each set of 4 spectra collected around 1 focus point was taken. Principal component analysis was used to show that the spectra recorded in the different growth stages can be distinguished from each other. For two-dimensional correlation spectroscopy in-house routines were implemented in MATLAB, in which the Hilbert transform algorithm for equidistant data-intervals was used.

7.3 Results and discussion

Raman spectra of *C. metallidurans* LMG 1195 were collected at five points of the growth curve, corresponding to 16, 24, 48, 72 and 96 hours of incubation. Despites the standardized protocols for inoculation and incubation, there are still considerable OD variations within the groups of the spectra collected after 16 and 24 hours of incubation (both exponential growth) due to biological variation. In the PCA plot (not shown), this variation caused an overlap between 16 and 24 hours spectra. The spectra collected after 16 and 24 hours of incubation were therefore rearranged into groups according to OD (590nm) values of lower than 1.0, respectively 1.0 and above 1.0. These groups are more representative for the different growth stages and were further referred to as stage 1 and 2. Stage 3, 4, and 5 still correspond to the data obtained after 48, 72 and 96 hours of incubation, respectively. The PCA plot (Fig. 7.1a) now clearly shows separate groups and a direction of growth can be seen which reaches a maximum in PC2 when stationary growth begins (stage 3).

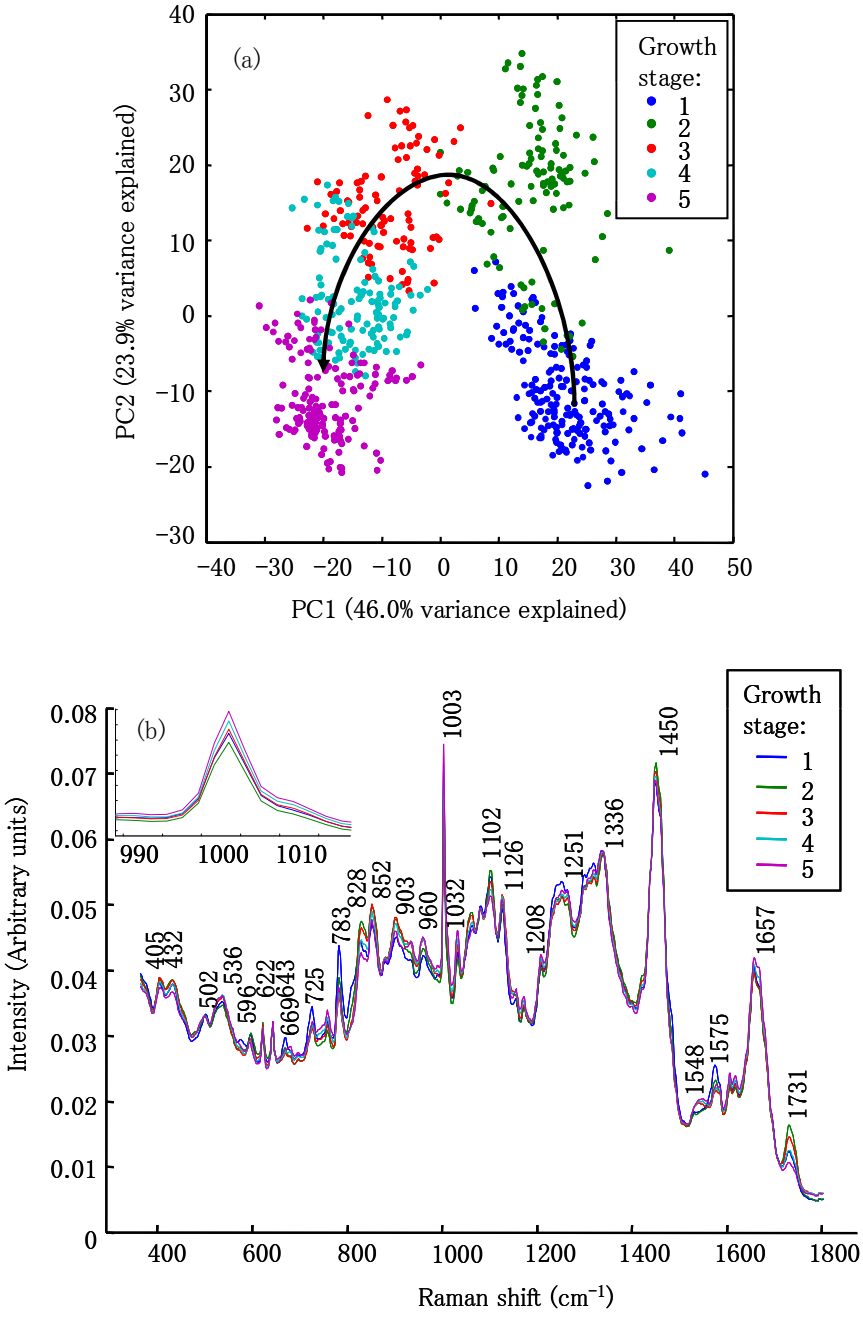


Figure 7.1 (a) PCA plot colored by growth stage (arrow indicates growth) and (b) average spectra per growth stage. The Phe ring breathing band is showed in detail in the inset.

To study the differences between the five observed growth stages the average spectra per growth stage were calculated (Fig. 7.1b). Although these spectra are very similar, some trends as a function of growth stage can already been seen at 725, 783, 828, 852, 1032, 1102 and 1575 cm^{-1}.* In literature[12], some of these bands have been assigned to

* The values of the Raman shifts mentioned for specific metabolic products can vary 1 cm^{-1} between the average spectra, difference spectra and 2D correlation spectra.

biomolecules: adenine (A) for 725 cm^{-1}, cytosine (C) and uracil (U) for 783 cm^{-1}, tyrosine (Tyr) for 828 and 852 cm^{-1}. Unfortunately, only major differences can be observed. In order to see more details, difference spectra between each couple of two successive growth stages were calculated. In the difference spectrum of stages 1 and 2 for example (Fig. 7.2), some features appear at 808, 1140 and 1482 cm^{-1} that were not visible as differences in the average spectra or that were present in the average spectra as shoulders to more intense bands.

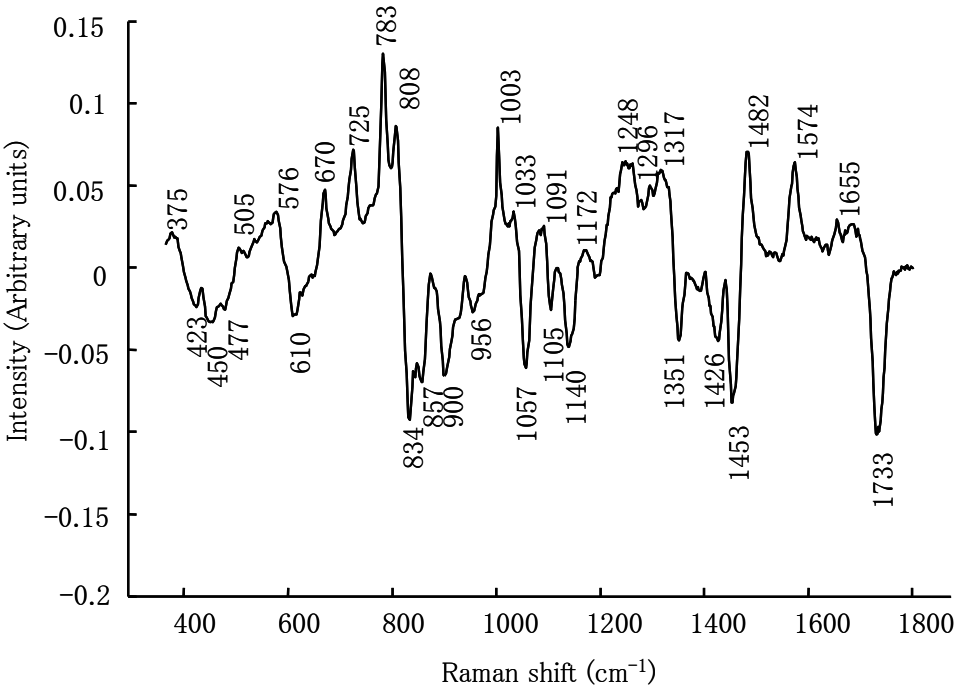


Figure 7.2 Difference spectrum between the average spectra of growth stage 1 and 2

However, only intense bands of biomolecules that are present in sufficient amounts give rise to visible bands in the average spectra and to visible features in difference spectra. Moreover, the most intense band(s) of some biomolecules may overlap. These two effects complicate the interpretation of average and difference spectra. From our reference database (Chapter 5), we found for example that the Raman band at 725 cm^{-1}, which was in literature assigned to A, could also be caused by coenzyme A (CoA) and acetyl coenzyme A (acetyl-CoA) respectively. The same remark can be made for the band at 783 cm^{-1} which was in literature[12] assigned to C and U, but could also be caused by citric acid according to our database. As little other intense bands of these biomolecules are clearly visible in the average and difference spectra and because these bands can also include contributions of

different biomolecules, it is hard to determine which biomolecule has the largest contribution for a certain Raman band. Moreover, opposite changes over time in the concentration profile of biomolecules with the same main Raman bands can average out, which makes it impossible to study the evolution in time for each biomolecule separately. Therefore, more advanced mathematical approaches are necessary that do not focus on just a few Raman bands, but that take the whole reference spectrum of the biomolecules into account.

Table 7.2 Trends in dot product values and EMSC coefficients (obtained by application of EMSC procedure by Martens *et al.*, datacase 106: reference database an input of 'Good spectra'[5]) for each reference spectrum of the database. >> and << indicate significant differences at the 95% confidential level that were tracked with ANOVA.

Reference product	dot product	EMSC coefficients
Fatty acids		
Myristic acid (fat14)	1>2>3>4>5	1<<2<<3>4>>5
Palmitic acid (fat16)	1>>2>3>4>5	1>>2>3>>4>>5
Stearic acid (fat18)	1>>2<3<4<<5	1>2>>3<<4<<5
Vaccenic acid (vac)	1>>2>3<<4<<5	1>>2>>3>4<<5
Metabolites		
D-fructose-6-phosphate (Df6P)	1<<2>>3>>4>>5	1<<2<3>>4>>5
Malic acid (mal)	1<<2<3>>4>>5	1<<2<3<<4<<5
Li acetoacetate (acac)	1<<2<<3<<4<<5	1<<2<3>4<5
Succinic acid (succ)	1>>2<<3<<4<<5	1>>2<<3<<4<<5
Citric acid (cit)	1>>2>>3>4>5	1>>2>>3>>4<5
Acetyl coenzyme A (acCoA)	1>>2>>3<4<<5	1>>2>>3>>4>>5
Coenzyme A (CoA)	1>>2>>3>4<<5	1>>2<<3<<4<<5
Phosphoenolpyruvate (PEP)	1>>2>3<<4<<5	1>>2<<3<<4<<5
Amino acids		
Tryptophan (Tryp)	1>2<3<4<5	1>>2<3<<4<<5
Glycine (Gly)	1>>2>3<4<<5	1<<2<<3<4>5
Phenylalanine (Phe)	1>>2<<3<<4<<5	1>>2<<3<<4<<5
Tyrosine (Tyr)	1<<2<<3>>4>5	1<<2>>3>>4>>5
DNA/RNA bases		
Adenine (A)	1>>2>>3<4<<5	1<<2<<3<4<5
Cytosine (C)	1>2>>3>4>5	1>>2>>3<4<5
Thymine (T)	1>2<<3<<4<5	1>>2>>3<<4<<5
Uracil (U)	1<<2>>3>4>5	1>>2>>3<4<<5

A possible approach that could be used to determine the contribution of reference spectra in bacterial Raman spectra is the calculation of the dot product, that was used in a vector correction routine by Maquelin *et al.*[11] to estimate and subtract the contribution of water and medium from bacterial Raman spectra. Here, we use the dot product of a reference spectrum from our database (listed in Table 7.2) and a bacterial Raman spectrum as a number for the contribution of this reference spectrum in the bacterial Raman spectrum. These dot products were calculated between all reference spectra of the database and all bacterial Raman spectra of our dataset. The whole bacterial Raman dataset was first treated with a pure physical EMSC (datacase 103, no input of reference spectra) and all replicate spectra of each reference product separately, were also treated with EMSC (datacase 103). ANOVA was applied on these values (factor: growth stage), which reveals for each reference spectrum whether there are significant differences in contribution between the different stages of growth (Table 7.2). The advantage of this approach is that all bands of the Raman spectrum of the reference product are taken into account, in contrary to the interpretation of the average and difference spectra, which is only based on one or a few intense bands of the reference spectra. Nevertheless, one has to keep in mind that overfitting can occur due to contributions of structural very similar biomolecules. For example, the trends in dot product values between the different stages of growth for myristic acid (fat14) and palmitic acid (fat16) on one hand, and A, CoA and acetyl-CoA on the other hand, are similar. The dot product values for these biomolecules probably enclose the contributions of not only a given, but also of several very similar products (that give rise to very similar Raman spectra). To distinguish contributions of structural similar products, a more refined technique has to be used, namely the use of coefficients from the EMSC procedure.

To address the issue of overfitting, the original bacterial dataset was also treated with an EMSC procedure (datacase 106, physical correction and input of 'Good Spectra') where the reference database of biomolecules was taken into account. Reference spectra that systematically gave rise to negative dot product values with our bacterial data set were not used as input in this EMSC treatment. These biomolecules are weak Raman scatterers or

are not present in sufficient amounts in the cell to be detected by Raman spectroscopy and hence negative values are a result of sample and spectral variation. The EMSC coefficients ($\Delta c_{i,j}$: value for the difference in concentration of component j between the bacterial spectrum i and the mal spectrum) were imported in SPSS and an ANOVA test was conducted (factor growth stage). The results of this test are summarized in Table 7.2, where the relation between the average EMSC coefficients per stage of growth is given by < or > , and by >> or << for a significant difference at 95% confidential level. Most differences were significant and the mean plots (Fig. 7.3) do not show random variations but rather clear indications as a function of growth, which are acceptable from the microbiological point of view.

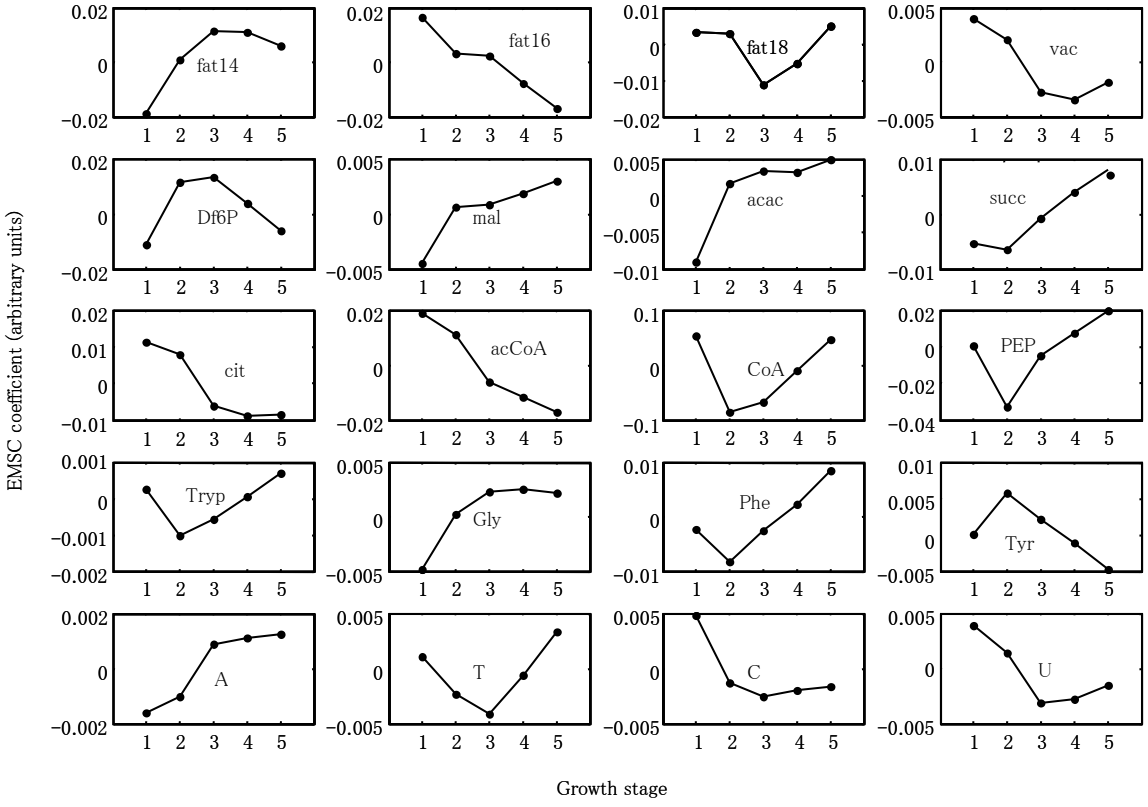


Figure 7.3 Mean plots of the EMSC coefficients (obtained by application of EMSC toolbox by Martens, datacase 106: reference database an input of 'Good spectra'[14]) as a function of the growth stage

From comparison of the dot product values and the EMSC coefficients, a similar evolution in function of growth was seen for certain biomolecules. In contrary, for some of the structural very similar biomolecules, the EMSC coefficients show different trends, where

the dot product values show similar tendencies. For example, the evolution of the dot product values as a function of growth is the same for A, CoA and acetyl-CoA, but EMSC coefficients show different evolutions for these three structural similar biomolecules. This confirms the expectation that simultaneous fitting of the reference spectra in the EMSC procedure provides an advantage over separate calculation of dot products. Less overfitting occurs in the EMSC procedure and the contributions of A, CoA an acetyl-CoA in the bacterial Raman spectra can now be separated from each other. Some other biomolecules also show different trends for dot product values and EMSC coefficients, even if they are not structurally very similar to other biomolecules from the database. This is also due to overfitting caused by resemblance of intense bands in certain reference spectra when calculating dot products. For example, both malic acid and phosphoenolpyruvate (PEP) cause an intense Raman band at 1033 cm^{-1} and PEP and citric acid have a band at 783 cm^{-1} in common, despite the fact that these molecules are structurally not very similar. EMSC coefficients show different trends for PEP and malic acid than the dot products, because all reference spectra are fitted together in the EMSC procedure, in contrary to the calculation of dot products for each reference spectrum separately. In this way the use of EMSC coefficients reduces overfitting and consequently confusion of reference spectra and the results are more reliable. For biomolecules with a very high, visible contribution in the bacterial Raman spectra, the trends in dot product values and EMSC coefficients are similar and resemble the observed visible evolution in the average and difference spectra. For example, dot product values and EMSC coefficients show for Phe a decrease from 16 h incubation to 24 h incubation and an increase from 24 h to 96 h incubation, while for Tyr an opposite trend is observed. These increases and decreases can be confirmed by studying the characteristics bands of these biomolecules (1003 cm^{-1} for Phe and 828-832 cm^{-1} for Tyr) in the average spectra (Phe band in the inset of Figure 7.1b) and the consequent difference spectra. Indeed, Phe shows a positive band at 1003 cm^{-1} in the difference spectra between growth stages 1 and 2 (Fig. 7.2) and a negative band at 1003 cm^{-1} in the difference spectra between the other growth stages (data not shown). This consistency supports that the trends predicted by the EMSC coefficients are reliable. Furthermore, the evolution in the Raman band at 783 cm^{-1} resembles the trend

of citric acid, C and U predicted by the EMSC coefficients. Probably both citric acid and the DNA/RNA bases C and U contribute to the Raman band at 783 cm^{-1} in the bacterial spectra.

From Figure 7.3 a remark can be made concerning the metabolic pathway of *C. metallidurans*. Apparently, the formation of D-fructose-6-phosphate (Df6P) is ceased in the stationary phase. Subsequently, there is no conversion to acetyl-CoA and citric acid, which are still metabolized in the citric acid cycle. The levels of citric acid and acetyl-CoA decrease and the level of the intermediates of the citric acid cycle (malic acid and succinic acid) remain constant or increase.

An important remark here is that the reference spectra were collected from pure, mostly powdery products. The physiologic form of the product may be different in the cell and therefore it is possible that the Raman bands of the reference spectra are not exactly identical to those of the same molecules retrieved from the bacterial spectra. Nevertheless, it is assumed that the changes are minimal because of the high consistency between the results of the different presented processing techniques. Moreover, the predicted trends are not random variations but acceptable evolutions as a function of growth.

In order to confirm or complete the results, 2D correlation spectroscopy was applied (2D spectrum not shown). The bacterial Raman spectra were collected after five fixed incubation times and some of them were redivided into groups according to OD values, so five classes were obtained which reflect different growth stages. Because incubation time and OD are not linearly related to the spectral changes, the 2D correlation algorithm for constant intervals was used to treat the five stages of growth. Because of the complex nature of the bacterial Raman spectra and because most spectral changes as a function of growth were not linear, no conclusions were drawn from the asynchronous 2D spectra. The symmetry axis signal of the 2D synchronous spectrum (Fig. 7.4), referred to as power spectrum, shows the Raman bands that vary most importantly during growth. Most of these bands were also found in the difference spectra.

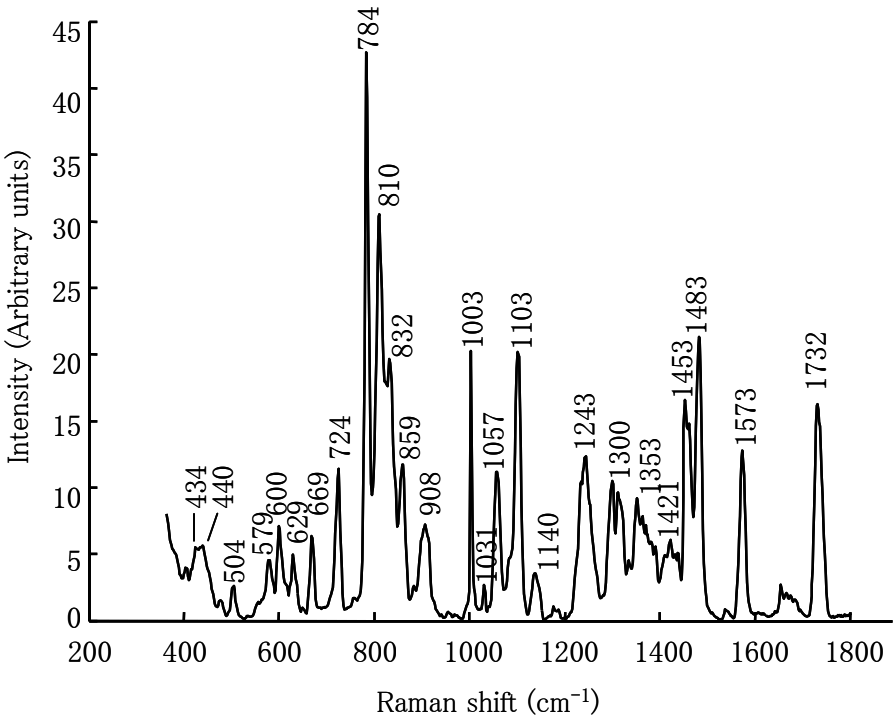


Figure 7.4 2D synchronous power spectrum

The signals in the 2D synchronous spectrum on a vertical line at each wavenumber, referred to as slice spectra, reflect which wavenumbers are changing together during the experiment. All synchronous slice spectra for each wavenumber of the bands observed in the power spectrum, were plotted and subsequently a hierarchical cluster analysis was performed (Pearson correlation, average linkage; data not shown). Three clearly separate groups of slice spectra are formed, from which three representative slice spectra at 1003, 832 and 784 cm^{-1} are shown in Figure 7.5. These wavenumbers can be related to respectively Phe, Tyr and a combination of citric acid, C and U.

The slice spectrum of 784 cm^{-1} provides further proof for the contribution of citric acid, C and U at that wavenumber. Indeed, the changes in the Raman band at 784 cm^{-1} seem to be positively correlated to bands that appear in the reference spectra of these biomolecules: 579, 1099 and 1234 cm^{-1} for U, 554 (sh), 559, 1254 and 1654 cm^{-1} for C and 1050, 1084 (sh), 1392 (sh) and 1732 cm^{-1} for citric acid. Not all bands of these biomolecules can be found in this slice spectrum, which is acceptable because changes in opposite directions of different biomolecules with a band at the same wavenumber can counteract each other. The slice spectrum also contains a positive band at 724 cm^{-1} which was earlier associated with

acetyl-CoA, CoA and A. Moreover, the most intense bands of the reference spectrum of acetyl-CoA can also be found in this slice spectrum at 629 and 1335 cm^{-1}. For A the same argumentation can be made for the bands at 535 and 1335 cm^{-1}, while for CoA the resemblance between bands in the reference spectrum and the slice spectrum for 784 cm^{-1} is minimal. The strong correlation between the bands at 724 and 784 cm^{-1} further supports the suggestion of the contribution of acetyl-CoA and citric acid in the Raman bands at respectively 724 and 784 cm^{-1}. Indeed, a similar trend of these biomolecules in function of growth can be expected, regarding that citric acid succeeds acetyl-CoA in the metabolic pathway of *C. metallidurans*. This similarity in trends was also predicted by the EMSC coefficients.

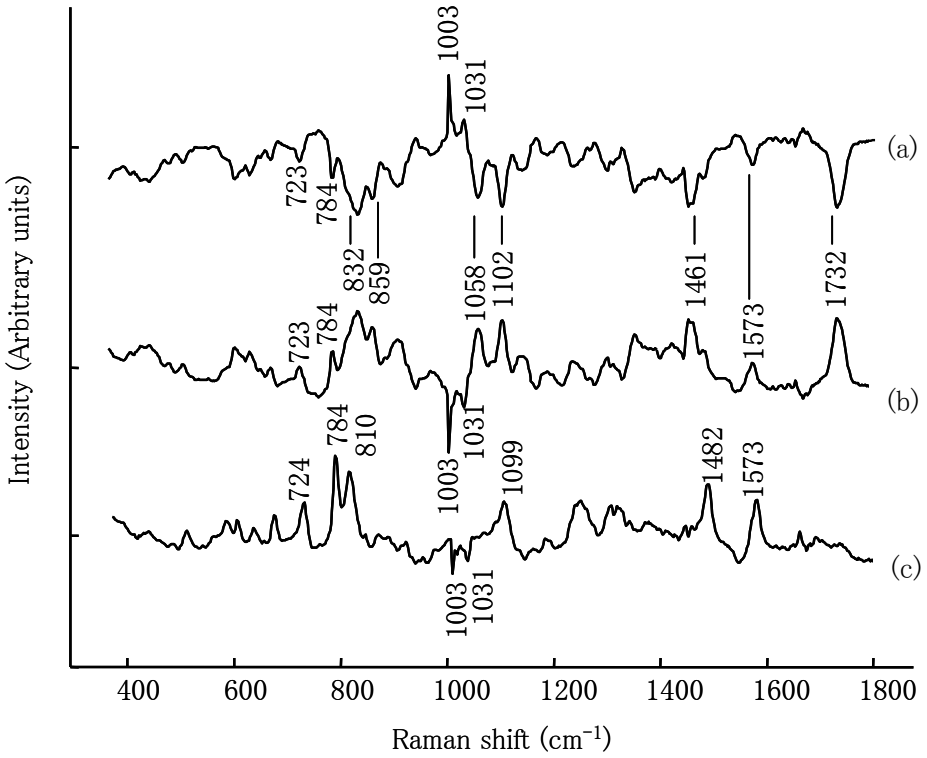


Figure 7.5 2D synchronous slice spectra at (a) 1003, (b) 832 and (c) 784 cm^{-1}

7.4 Conclusions

In this article we presented several approaches for the interpretation of bacterial Raman spectra, which were applied on a large data set containing Raman spectra from five stages of growth of *C. metallidurans*. A study of the Raman bands of the bacterial spectra as such

only provides information of molecules that are present in large amounts and/or that are intense Raman scatterers, such as Phe and Tyr. Difference spectra reveal more varying Raman bands. Indeed, some shoulders that were increasing or decreasing according to the growth stage, were not observed in the average spectra, but caused distinct Raman bands in the difference spectra. Since all Raman bands are the sum of bands from different biomolecules, more complex mathematical approaches were applied in order to distinguish the contributions of different biomolecules in the bacterial Raman spectra. Dot products and EMSC coefficients were used to reveal the evolution of biomolecule concentrations during growth. These methods both take all Raman bands of the reference spectra of the biomolecules into account. During the EMSC procedure, all reference spectra of the biomolecules in our database were fitted simultaneously to the bacterial Raman spectra. Thus, contributions of biomolecules with similar Raman spectra could be distinguished from each other, which is an important advantage of the EMSC coefficients over the dot product values. With 2D correlation spectroscopy the most important changes in the dataset could easily be identified by the power spectrum and related to each other by the synchronous slice spectra of three representative wavelengths.

The reference spectra were recorded from pure, mostly solid products and therefore can have slightly different Raman bands compared to the same molecules in the physiological environment in the cell, which can induce anomalies in the calculation of the EMSC coefficients. A database of reference spectra of biomolecules in physiological environment would be useful, but may be hard to obtain. Nevertheless, the results of several techniques that were applied confirmed each other and produced biochemically acceptable trends.

The above mentioned techniques all confirm that the Raman bands at 724 and 783 cm^{-1} not only contain a contribution of A respectively C and U, both also of CoA and acetyl-CoA respectively citric acid. The EMSC coefficient profiles reveal the ceasing of D-fructose-6-phosphate in the stationary phase, which leads to a decrease in acetyl-CoA and citric acid levels that are consumed by the citric acid cycle and to an accumulation of citric acid cycle intermediates.

This paper shows that models as incorporated in an EMSC procedure are very useful to estimate the contribution of reference spectra in bacterial Raman spectra. The question rises to what extent the number of reference spectra added to the model influences the estimation of their contribution to the bacterial Raman spectra. Would the accuracy of the model be improved by raising the number of reference spectra/products or by adding only some representatives of each group of biomolecules? This, together with a study of the similarity of the Raman signals of biomolecules in pure solid state (reference spectra) and physiological conditions could improve the accuracy of the results, although the output of such models will always be an estimation. Therefore, the challenge should be to make the estimation as accurate as possible. It is also recommended to include always some basic cell compounds that contribute highly in bacterial Raman spectrum (even when they are not of interest in the study) to improve the model's accuracy. This issue was discussed by Sowa et al.[18], who evaluated several models for the estimation of the composition of simulated multicomponent systems. They showed the importance of the presence of the basis reference spectra in order to perform accurate least squares estimations (the EMSC procedure includes a least squares estimation).

The developers of the EMSC procedure warn for possible colinearity problems[13], which can occur e.g. in case highly similar reference spectra are added to the model. When spectra of structurally very similar biomolecules are added, the model cannot distinguish between the contributions of these biomolecules to the bacterial spectra. In the next chapter, a possible solution for this problem is supplied and illustrated for the analysis of the fatty acid composition of bacterial cells.

References

1. Czarnik-Matusewicz B., Murayama K., Tsenkova R., Ozaki Y. (1999) Analysis of near-infrared spectra of complicated biological fluids by two-dimensional correlation spectroscopy: Protein and fat concentration-dependent spectral changes of milk. *Appl. Spectrosc.* 53: 1582-1594.
2. Decker M., Nielsen P.V., Martens H. (2005) Near-infrared spectra of Penicillium camemberti strains separated by extended multiplicative signal correction improved prediction of physical and chemical variations. *Appl. Spectrosc.* 59: 56-68.
3. De Gussem K., Verbeken A., Vandenabeele P., De Gelder J., Moens L. (2006) Raman spectroscopic monitoring of *Lactarius* latex. *Phytochemistry* 67: 2580-2589.
4. Escoriza M.F., Vanbriesen J.M., Steart S., Maier J. (2006) Studying bacterial metabolic states using Raman spectroscopy. *Appl. Spectrosc.* 60: 971-976.
5. Hutsebaut D., Vandenabeele P., Moens L. (2005) Evaluation of an accurate calibration and spectral standardization procedure for Raman spectroscopy. *Analyst* 130: 1204-1214.
6. Hutsebaut D., Vandroemme J., Heyrman J., Dawyndt P., Vandenabeele P., Moens L., De Vos P. (2006) Raman microspectroscopy as an identification tool within the phylogenetically homogeneous '*Bacillus subtilis*'-group. *Syst. Appl. Microbiol.* 29: 650-660.
7. Karasinski J., Andreescu S., Sadik O.A., Lavine B., Vora M.N. (2005) Multiarray sensors with pattern recognition for the detection, classification, and differentiation of bacteria at subspecies and strain levels. *Anal. Chem.* 77: 7941-7949.
8. Kirschner C., Maquelin K., Pina P., Thi N.A.N., Choo-Smith L.P., Sockalingum G.D., Sandt C., Ami D., Orsini F., Doglia S.M., Allouch P., Mainfait M., Puppels G.J., Naumann D. (2001) Classification and identification of enterococci: a comparative phenotypic, genotypic, and vibrational spectroscopic study. *J. Clin. Microbiol.* 39: 1763-1770.
9. Lopez-Diez E.C., Goodacre R. (2004) Characterization of microorganisms using UV resonance Raman spectroscopy and chemometrics. *Anal. Chem.* 76: 585-591.

10. Maquelin K., Choo-Smith L.P., Endtz H.P., Bruining H.A., Puppels G.J. (2002) Rapid identification of *Candida* species by confocal Raman micro spectroscopy. *J. Clin. Microbiol.* 40: 594-600.
11. Maquelin K., Choo-Smith L.P., van Vreeswijk T., Endtz H.P., Smith B., Bennett R., Bruining H.A., Puppels G.J. (2000) Raman spectroscopic method for identification of clinically relevant microorganisms growing on solid culture medium. *Anal. Chem.* 72: 12-19.
12. Maquelin K., Kirschner C., Choo-Smith L.P., van den Braak N., Endtz H.Ph., Naumann D., Puppels G.J. (2002) Identification of medically relevant microorganisms by vibrational spectroscopy. *J. Microbiol. Meth.* 51: 255-271.
13. Martens H., The EMSC toolbox for MATLAB http://www.models.kvl.dk/source/emsctoolbox
14. Martens H., Nielsen J.P., Engelsen S.B. (2003) Light scattering and light absorbance separated by extended multiplicative signal correction. Application to near-infrared transmission analysis of powder mixtures. *Anal. Chem.* 75: 394-404.
15. Naumann D., Keller S., Helm D., Schultz C., Schrader B. (1995) FT-IR spectroscopy and FT-Raman spectroscopy are powerful analytical tools for the noninvasive characterization of intact microbial cells. *J. Mol. Struct.* 347: 399-504.
16. Noda I., Ozaki Y. (2004) Two-Dimensional Correlation Spectroscopy - Applications in Vibrational and Optical Spectroscopy. John Wiley & Sons Ltd., Chichester.
17. Oust A., Moretro T., Naterstad K., Sockalingum G.D., Adt I., Manfait M., Kohler A. (2006) Fourier transform infrared and Raman spectroscopy for characterization of *Listeria monocytogenes* strains. *App. Environ. Microb.* 72: 228-232.
18. Sowa M.G., Smith M.S.D., Kendall C., Brock E.R., Ko A.C.-T., Choo-smith L., Stone N. (2006) Semi-parametric estimation in the compositional modeling of multicomponent systems for Raman spectroscopic data. *Appl. Spectrosc.* 60: 877-883.

Chapter 8: Methods for extracting biochemical information from bacterial Raman spectra: Focus on a group of structurally related biomolecules – fatty acids

Joke De Gelder, Kris De Gussem, Peter Vandenabeele, Marc Vancanneyt, Paul De Vos, Luc Moens

Analytica Chimica Acta, 603 (2007) 167-175

In this chapter, the EMSC procedure, which was introduced in Chapter 7, is used to estimate contributions of highly similar biomolecules in bacterial Raman spectra. As we focused on the fatty acid composition within a bacterial cell, highly similar fatty acid spectra were added to the EMSC model, which could suffer from colinearity problems[13,14]. This chapter illustrates the introduction of a PCA before EMSC is started as a solution to avoid colinearity problems.

8.1 Introduction

Over the past few years, Raman spectroscopy became attractive as analytical tool in microbiology. The development of identification procedures for microorganisms such as bacteria[7,8,10,16], yeasts[11] and fungal spores[5] have been successful. Several research groups have reported band assignments in bacterial Raman spectra for biomolecules such as phenylalanine, tyrosine, DNA and RNA bases, etc.[8,12,15]. However, many other bands in a bacterial Raman spectrum remain unassigned, due to the complex nature of the spectrum. Each band is most often the superimposed result of bands caused by several biomolecules present in the cell. The assignment of a band from bacterial Raman spectra to one specific biomolecule is therefore difficult or even impossible. De Gelder *et al.*[3] present several methods to extract more biochemical information from complex Raman spectra (Chapter 7). Using a reference database and mathematical approaches, information was obtained for several biomolecules present in the cell.

The question raised how detailed the biochemical information is that can be extracted from bacterial Raman spectra. It is our aim to explore the possibility of achieving relevant information on a group of very similar biomolecules and so this paper focuses on the contribution of fatty acids in bacterial Raman spectra. The fatty acid composition of biological materials has already been studied with Raman spectroscopy: Beattie *et al.*[2] analyzed adipose tissue, while Afsethi *et al.*[1] used self-made mixtures of fatty acids that serve as food model systems. In both cases, the Raman spectra consisted solely or mainly of bands that coincide with fatty acids. These spectra allowed determining the fatty acid

composition by partial least squares regression (PLS). Laucks *et al.*[9] reported that Raman spectra of marine and common mesophilic bacteria show large differences. One band was attributed to the presence of polyunsaturated fatty acids, but no single fatty acids were determined quantitatively. In this paper, we explore the possibilities to retrieve fatty acid information from bacterial Raman spectra. For this purpose, we selected *Enterococcus faecalis* LMG 7937 grown on two different media and *Bacillus coagulans* LMG 6326 incubated at two different temperatures. The Raman spectra of the two conditions for each strain were compared in order to answer the following questions: (i) can these culturing conditions be distinguished from one another by their Raman spectra, (ii) can these culturing conditions be distinguished from one another by the fatty acid information provided by the Raman spectra and (iii) how detailed is the fatty acid information that we can extract from bacterial Raman spectra? To confirm our findings, we performed in parallel gaschromatographic Fatty Acid Methyl Ester (FAME) analyses.

8.2 Experimental

8.2.1 Strains and culturing conditions

Enterococcus faecalis LMG 7937 was cultured at 28 °C on two different solid media namely TSA (tryptone soya agar) and MRS (De Man, Rogosa, Sharpe) Agar, while *Bacillus coagulans* LMG 6326 was cultured on TSA at two different temperatures namely 32 and 52 °C. All incubations occured for 24 h to obtain confluent growth. From each plate a sample for Raman analysis was taken with a 1 μL smear loop, while the rest was harvested from one or more plates to obtain sufficient biomass for gaschromatographic FAME analysis.

8.2.2 Raman spectroscopy

A Kaiser System Hololab 5000R modular Raman microspectrometer was used. A 785 nm diode laser (Toptica Photonics AG) was focused through an 100x objective of the microscope (Leica), to obtain ca. 45 mW of laser power on the sample. The scattered light was transferred through a confocal 15 µm aperture pinhole and a collection fiber to the spectrograph. Detection was performed by a back illuminated deep depletion Pelletier cooled (-70°C) CCD detector (Andor). The Raman signal was collected in the spectral interval of 150 cm^{-1} until 3500 cm^{-1}. Only the region from 365 till 1780 cm^{-1} was used because: i) the sensitivity is lower at higher wavenumbers and ii) for higher wavenumbers there are no standards that allow a good quality calibration of the Raman shift axis. The acquisition time was 60 s for all spectra with a spectral resolution of ca. 4 cm^{-1}.

The reference spectra of fatty acids and other biomolecules (DNA and RNA bases, amino acids and some primary metabolites) were presented by De Gelder *et al.*[4] (Chapter 5), where all reference products are listed together with their specifications and suppliers. For the measurement of bacterial Raman spectra, the samples were transferred to a CaF_2 plate and dried for 10 minutes on silica gel.

Spectral calibration was performed in MATLAB as described by Hutsebaut *et al.*[6]. The spectra were preprocessed by extended multiplicative signal correction (EMSC)[14], for which the toolbox by Martens was used[13]. The reference spectra were treated with datacase 103 (physical correction) for each product separately, which includes correction for additive and multiplicative effects[4]. The bacterial Raman spectra were treated (per strain) with datacase 106. This datacase includes a physical correction (elimination of additive and multiplicative effects) and models the spectra as a function of 'good spectra', which are in this case the reference spectra of fatty acids and other biomolecules. The EMSC coefficients for each reference spectrum that were calculated in the EMSC model are values for the contribution of a reference spectrum in a bacterial spectrum compared to a mal spectrum (mean spectrum of the dataset). These coefficients were used as relative values for the concentration of the different biomolecules. These values can be compared per reference product, but not between different reference products, because they do not

take the intrinsic Raman activity of the reference products into account. The spectra were collected in a small temperature interval of 19 to 21°C.

8.2.3 Gas chromatographic fatty acid methyl esters analysis (FAME)

Methyl esters of fatty acids were prepared and analyzed with the MIDI (Newark, Delaware) gaschromatographic system according to the manufacturer's prescription. The obtained fatty acid profiles were compared pair-wise per strain and in relation to the different incubation conditions.

8.3 Results and Discussion

8.3.1 FAME analysis

Multiple FAME analyses were performed in parallel to the Raman spectroscopic measurements. Average values of the FAME data for each culturing condition are shown in Figure 8.1. These percentages reflect for each fatty acid the relative concentration compared to the total fatty acid content of the cell. For *B. coagulans* LMG 6326 (Fig. 8.1a) the major differences after incubation at 32 and 52 °C, are the percentages of c15Aiso* and c17Aiso. PCA (data not shown) shows that FAME profiles from 2 groups are clearly separated from each other by PC1 according to the incubation temperature. The loading plot shows that c15iso does not contribute significantly to this separation, as its absolute value of the loading for PC1 is relatively low. Furthermore, it is remarkable that lower fatty acids such as c14iso, c14, c15Aiso and c16iso have a high score for the data collected at 32 °C. This trend is expected as the cells try to keep the fluidity of their membranes at a constant level. Therefore, the production of fatty acids with longer carbon chains (which have higher boiling points) is increased at higher incubation temperatures. The results for *E. faecalis* LMG 7937 (Fig. 8.1b) show that the largest differences between

* Abbreviations of fatty acids are listed in Table 8.1.

the data obtained after culturing on TSA and MRS are the percentages of oleic acid and vaccenic acid, which differ only in the position of the double bond in the carbon chain. Again, the PCA plot (data not shown) shows 2 distinct groups according to the media that were used. The loading plot reveals that oleic acid, vaccenic acid, c14 and two sum peaks† contribute the most to PC1, which is responsible for the separation of these groups.

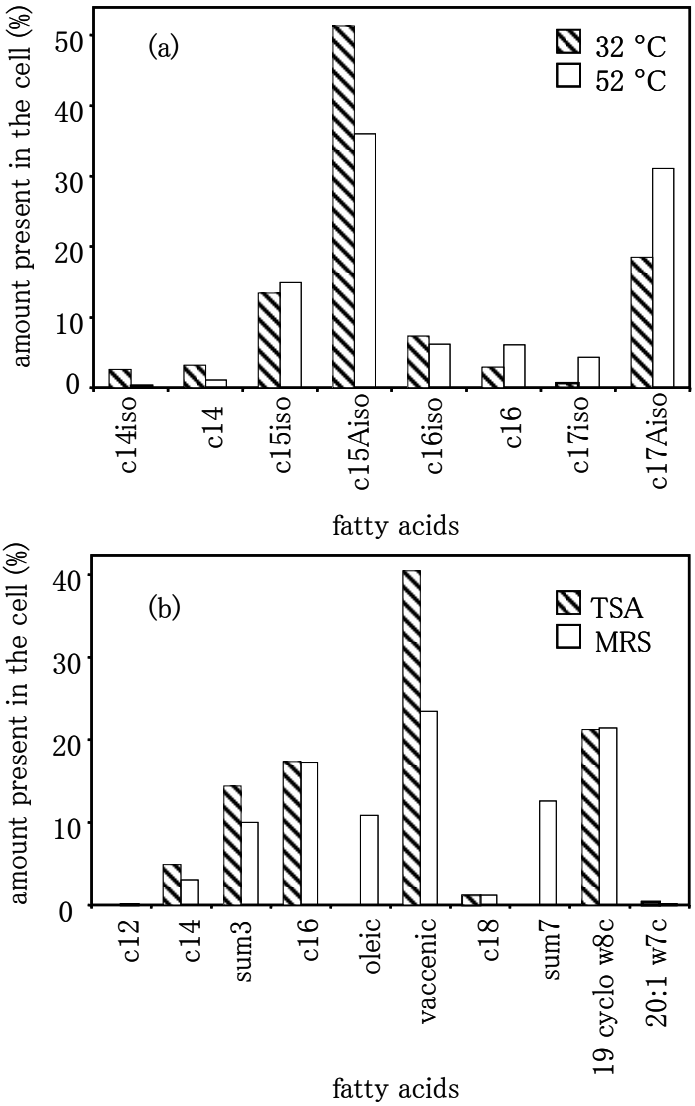


Figure 8.1 FAME analysis results for (a) *B. coagulans* LMG 6326 and (b) *E. faecalis* LMG 7937.

† Sum peaks contain contributions of several fatty acids that could not be separated by gas chromatography.

8.3.2 Discussion of the Raman spectra of fatty acids

The aim of this study is to examine whether useful information from a group of highly similar molecules, in this case fatty acids, can be extracted from a complex bacterial Raman spectrum. A logic first step is the study of the spectra of pure fatty acids that are commercially available and present in the cell as deduced from the FAME analyses of *B. coagulans* LMG 6326 and *E. faecalis* LMG 7937 (Table 8.1). Reference spectra of these products were recorded at multiple positions to include variations in the Raman spectra and are presented in De Gelder *et al.*[4], along with bands assignments (Chapter 5). The Raman spectra of saturated and unsaturated fatty acids are clearly different: i) the unsaturated fatty acids are in liquid state and therefore have much broader bands and ii) an intense band at 1655 cm^{-1} is present in the spectra of the unsaturated fatty acids and can be assigned to the ν(C=C) stretching vibration. The two unsaturated fatty acids, oleic acid and vaccenic acid, which differ from each other by the position of the double bond, show only small differences in their Raman bands.

Table 8.1 Fatty acids from which reference Raman spectra were recorded, along with their abbreviations used in the text and their specifications.

Product name	Abbreviation used in text	Specifications*
lauric acid	c12	98%
myristic acid	c14	Sigma grade, 99-100%
palmitic acid	c16	Free acid, Sigma grade
stearic acid	c18	Free acid, grade I, ~99%
12-methyltetradecanoic acid	c15Aiso	
13-methylmyristic acid	c15iso	
14-methylpentadecanoic acid	c16iso	
14-methylhexadecanoic acid	c17Aiso	
15-methylpalmitic acid	c17iso	
oleic acid		Reagent grade, 99%
cis-vaccenic acid		

* all products were supplied by Sigma-Aldrich, Belgium

PCA was performed on the Raman spectra of the saturated fatty acids to evaluate whether these spectra show sufficient differences to allow distinction between these fatty acids. The score plots in Figure 8.2 show that linear fats, iso branched fats and Aiso branched fats can be clearly distinguished from one another. Furthermore, the similar fatty acids c14 and c16 can be distinguished by PC1, while c15Aiso and c17Aiso are distinguished by PC3. Only the separation of the different iso fatty acids is not clear. These score plots illustrate the high similarity of the Raman spectra of these fatty acids and indicate that the interpretation of complex Raman spectra concerning the contributions of individual fatty acids will be difficult.

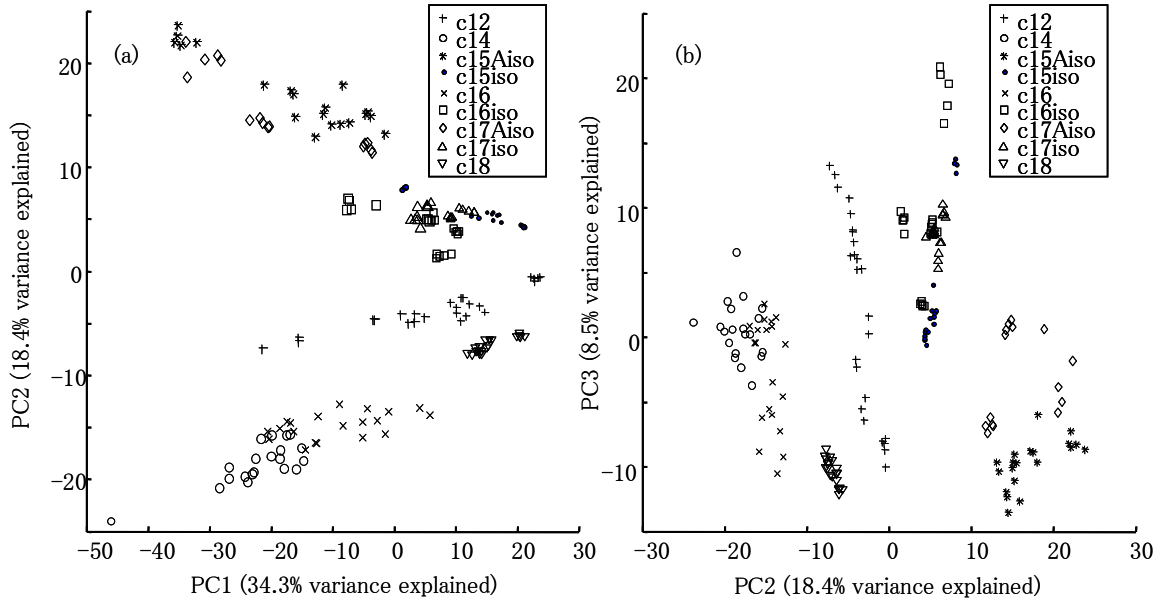


Figure 8.2 Score plots of the PCA of the reference spectra from saturated fatty acids: (a) PC2 versus PC1 and (b) PC3 versus PC2.

It should be stressed that the reference spectra are not necessarily identical to the contributions of the fatty acids in the bacterial Raman spectra, due to different physicochemical states. The concordance of our conclusions with these of FAME analysis in the discussion below shows that this issue is not of a major influence for the accuracy of the results.

8.3.3 Discussion of Raman spectra of B. coagulans and E. faecalis

After the discussion of the Raman spectra of the pure fatty acids, we studied whether information about (similar) fatty acids can be extracted from complex bacterial Raman spectra. The two strains were each cultured in two different conditions: *E. faecalis* LMG 7937 on TSA and MRS, and *B. coagulans* LMG 6326 on TSA at 32 and 52 °C. The Raman spectra were EMSC corrected (datacase 106[13]) with input of reference spectra from fatty acids and other biomolecules[4] (Chapter 5). PCA was performed on these preprocessed bacterial Raman spectra. Per strain, clearly distinct groups in the score plots of the first two principal components are formed according to the two different incubation conditions (Fig. 8.3).

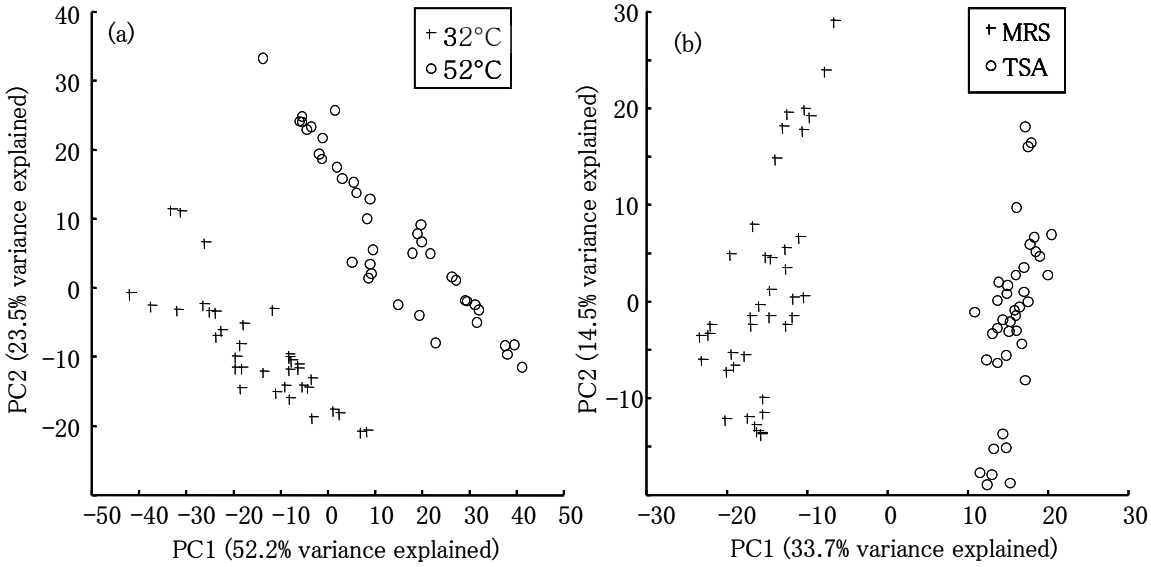


Figure 8.3 Score plots of the PCA of the bacterial Raman spectra of (a) *B. coagulans* LMG 6326 and (b) *E. faecalis* LMG 7937.

For each strain, the mean spectra were calculated per culturing condition (Fig. 8.4–8.5). Most bands are visible in all 4 spectra and the majority can be assigned to DNA/RNA or proteins. Indeed, the bands at ca. 622, 1004 and 1606 cm^{-1} are assigned to phenylalanine, while tyrosine causes bands at 644, 825, 853 and 1617 cm^{-1}. Bands of DNA and RNA can be found at 668 (G), 724(A), 782 (C,U) and 1573 cm^{-1} (G,A)[12], although those at 724 and 782 cm^{-1} can also contain contributions of primary metabolites acetyl coenzyme A and

citric acid, respectively[3,4]. Visible contributions of fats and fatty acids are observed at 1296 and 1449 cm^{-1} due to $\delta(CH_2)$ and δ(C–H) deformations, respectively. Saturated fatty acids contribute to bacterial Raman spectra by a triplet at ca. 1062, 1092–1103 and 1126–1135 cm^{-1} caused by ν(C–C) stretching, while unsaturated fatty acids contribute to these spectra at ca. 1655 cm^{-1} by ν(C=C) stretching.

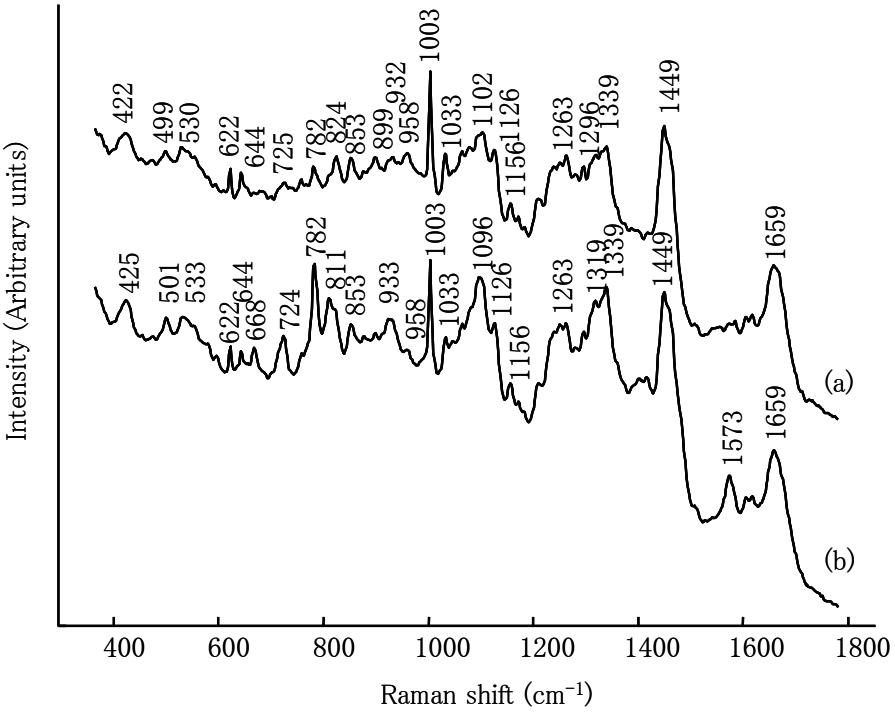


Figure 8.4 Raman spectra of *B. coagulans* LMG 6326 incubated at (a) 52 °C and (b) 32 °C

Although fatty acids seem to have a significant contribution in bacterial Raman spectra, some difficulties make it very hard to obtain accurate information about the fatty acid composition of the cell. First, most of the fatty acid bands visible in bacterial Raman spectra are caused in more or lesser extent by all saturated or all unsaturated fatty acids. As a consequence, at specific positions in the bacterial Raman spectra there is a superposition of signals of several fatty acids that cannot be separated. Second, at these Raman band positions contributions of other biomolecules can as well be observed. For example, bands at ca. 1092–1103 cm^{-1} do not only contain contributions from C–C skeletal vibrations, but also from ν(COC) 1,4 stretches from glycosidic links and $\nu(>PO_2^-)$ symmetric stretches. The same remark can be made for the band at ca. 1126 cm^{-1} where ν(C–C) and ν(C–N) stretches contribute, and for the band at ca. 1659 cm^{-1} where ν(C=C)

stretches of several biomolecules cause a Raman signal. When comparing both spectra of *B. coagulans* (Fig. 8.4), cultured under different conditions, the most remarkable differences are the larger intensities of the bands at 688, 724, 782 and 1573 cm^{-1} in the spectrum of incubation temperature 32 °C compared to 52 °C. These bands all contain contributions of the purine and pyrimidine bases present in DNA and RNA.

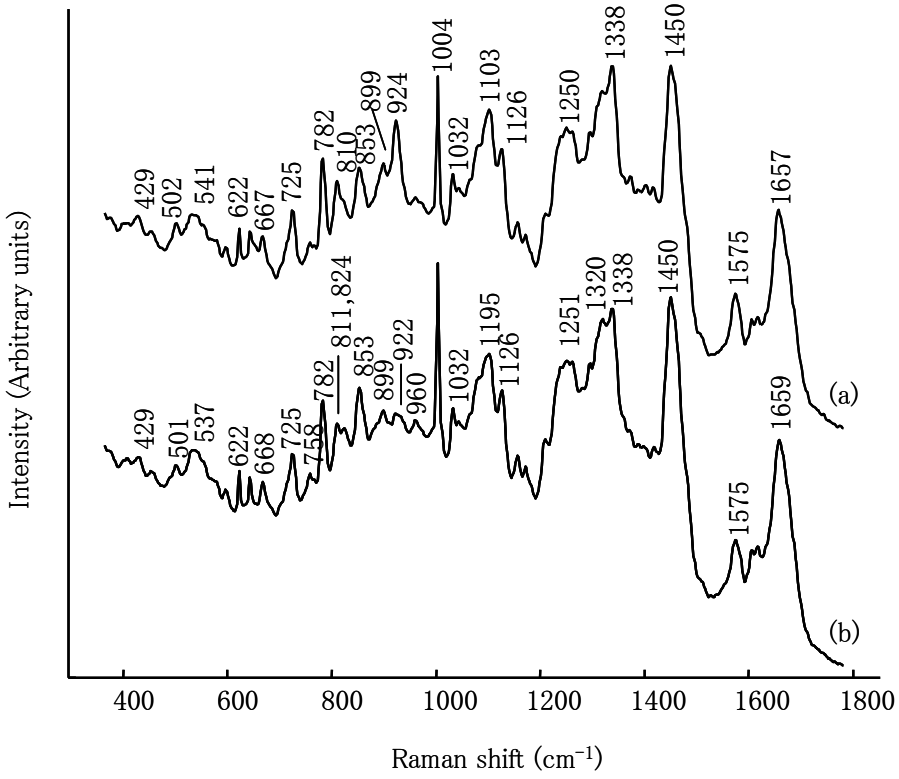


Figure 8.5 Raman spectra of *E. faecalis* LMG 7937 cultivated on (a) MRS and (b) TSA

The largest difference between both spectra of *E. faecalis* (Fig. 8.5) concerns the band at 924 cm^{-1}, which is of higher intensity in the spectrum of MRS compared to that of TSA. From these comparisons, we conclude that the most important differences in the Raman spectra are not related to the contribution of fatty acids. In order to visualize more differences, difference spectra were calculated between the two culturing conditions for each strain (Fig. 8.6). In the difference spectrum of *B. coagulans* (Fig. 8.6a), only the band at 1482 cm^{-1} can be related to fatty acids. Most Raman spectra of saturated fatty acids contain a band in that region as a shoulder to the intense band at 1449 cm^{-1}. Although this

band can thus be related to saturated fatty acids, it is not possible to draw conclusions about specific fatty acids, as this band is the result of contributions from various saturated fatty acids. In the difference spectrum of *E. faecalis* LMG 7937 (Fig. 8.6b) a negative band at in the region of 1655-1675 cm^{-1} is observed and seems to consist of several contributions, such as a shoulder at 1658 cm^{-1}. A band at 1655 cm^{-1} is only very weakly present in spectra of saturated fatty acids (ν(C=O) symmetrical stretch vibration), but is much more intense for unsaturated fatty acids (ν(C=C) stretch vibration). This difference spectrum thus suggests the presence of more unsaturated fatty acids when *E. faecalis* is cultured on TSA instead of on MRS. This observation is confirmed by the FAME data as the sum of the percentages of oleic and vaccenic acid is higher for the data collected with TSA instead of MRS grown cells. The most prominent difference between the reference spectra of oleic acid and vaccenic acid, is the presence of a band at 1118 cm^{-1}, which is also observed as a positive band in the difference spectrum of *E. faecalis*. This observation suggests the presence of more oleic acid when culturing on MRS, which is confirmed by the FAME data.

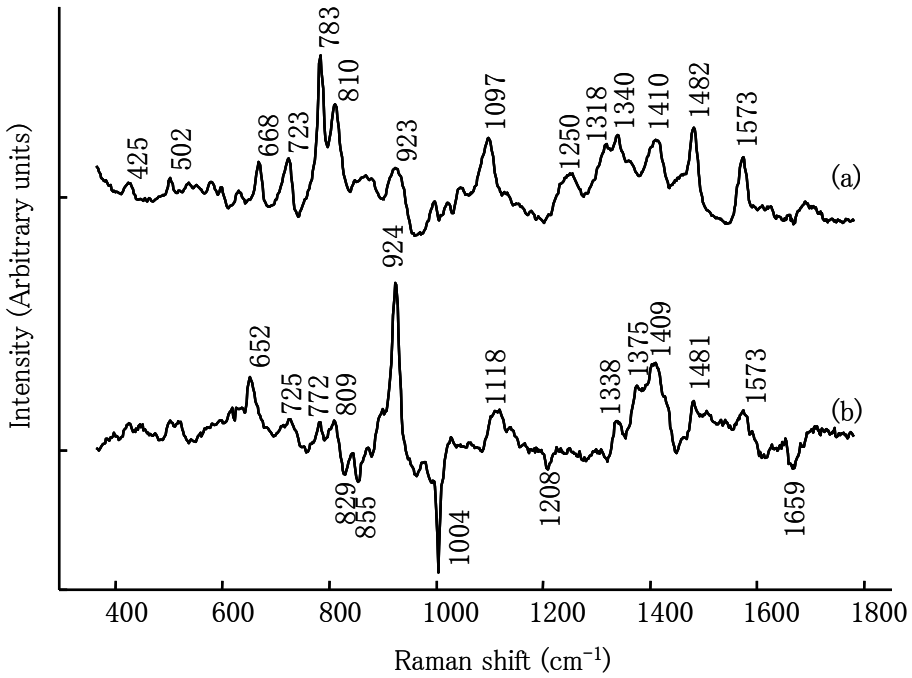


Figure 8.6 Difference spectra of both culturing conditions of (a) *B. coagulans* LMG 6326: 32°C-52°C and (b) *E. faecalis* LMG 7937: MRS-TSA.

In general, there is a significant contribution of fatty acids in bacterial Raman spectra. However, difference spectra show that most of the differences between the studied conditions are not related to fatty acids. Some minor differences can be attributed to fatty acids, although they are hard to assign to specific molecules, because these bands contain contributions of different fatty acids. Indeed, increasing contributions of certain fatty acids can be compensated by decreases of others and hence no difference is observed in the spectra. Band assignments are summarized in Table 8.2.

Table 8.2 Band assignments in the Raman spectra of *B. coaguluans* LMG 6326 and *E. faecalis* LMG 7937, based on Maquelin *et al.*[12] and De Gelder *et al.*[4]

Raman shift (cm^{-1})		
B. coagulans **LMG 6326**	***E. faecalis*** **LMG 7937**	*Band assignments*
425	429	
501	501	
533	537–540	S–S str, COC glycosidic ring def
622	622	Phe
644	644	Tyr
668	668	G
724	725	A, CoA, acetyl–CoA
782	782	C, U (ring str), citric acid
810	810	
824	825	“exposed” Tyr
853	853	“buried” Tyr
899	899	COC str
932	924	
958	960	
1003	1004	Phe
1033	1032	
1096–1102	1195–1103	C–C skeletal*, COC 1,4 str from glycosidic link, $>PO_2^-$ sym str
1126	1126	C–N and C–C str*
1156	1156	
1210	1210	
1263	1250	Amide III
1296	1296	CH_2 def*
1319	1320	
1339	1338	
1449	1450	$C–H_2$ def*
1573	1575	G, A (ring str)
1605	1606	Phe
1617	1617	Tyr
1659	1657–1659	Amide I, C=C str*

* Raman shifts where a contribution of fatty acids is possible.

8.3.4 Chemometric approach for the extraction of fatty acid information from bacterial Raman spectra

To obtain more information about the contribution of specific fatty acids to these Raman spectra, the coefficients of the fatty acids resulting from the EMSC correction were studied (datacase 106[13]), as described in Chapter 7[3]. These values (data not shown) seemed to be difficult to compare with the FAME values. The EMSC coefficients (for all spectra per strain) were subjected to PCA, which produced very complex loading plots that were hard to interpret. From these EMSC coefficients, no valuable conclusions could be drawn that were consistent with FAME analysis, likely because of colinearity problems in the EMSC procedure as described by Martens *et al.*[13,14]. To avoid colinearity, an alternative procedure was applied that is summarized in a flow chart in Figure 8.7.

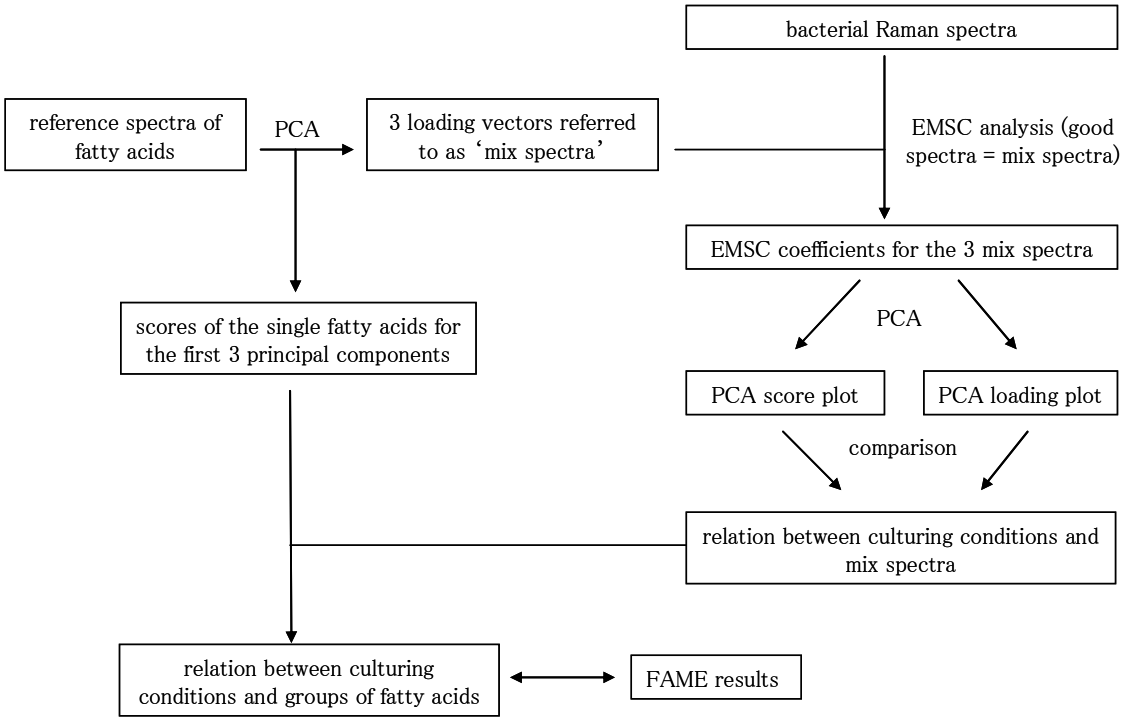


Figure 8.7 Flow chart of the data processing that was applied to obtain fatty acid information from bacterial Raman spectra.

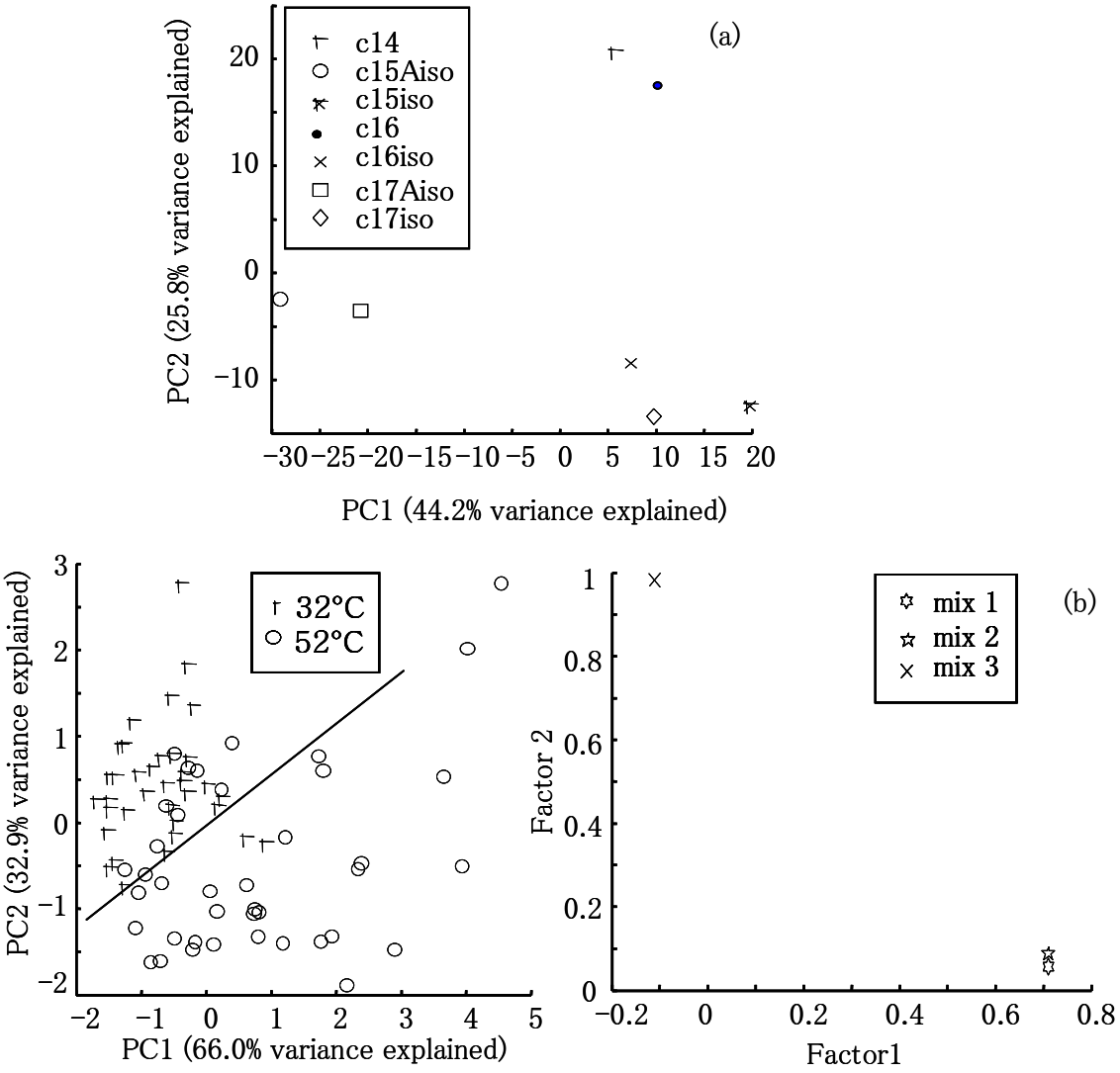


Figure 8.8 (a) Score plot from the PCA performed on the reference spectra of the fatty acids of interest for *B. coagulans* LMG 6326 and (b) score and loading plot from the PCA performed on the EMSC coefficients of the fatty acid mix spectra for *B. coagulans* LMG 6326.

We first performed PCA analysis on the reference Raman spectra of the fatty acids, where the loading vectors of a full rank PCA present an equivalent dataset where the vectors are orthogonal. For *B. coagulans* the score plot (PC1, PC2) in Figure 8.8a shows that PC1 makes a distinction between the Aiso fatty acids and the other fatty acids, while PC2 distincts linear from branched fatty acids. Other principal components do not allow making a useful differentiation between the fatty acids. The first three loading vectors of this PCA of the fatty acid reference spectra were selected to serve as new reference spectra. These spectra are linear combinations of the original fatty acid reference spectra and are orthogonal (further referred to as 'mix spectra'). Together with the reference spectra of

other biomolecules, these three loading vectors now serve as the 'good spectra' input for the EMSC procedure (datacase 106) applied on the bacterial Raman spectra. PCA analysis was performed on the resulting EMSC coefficients for the three mix spectra. Figure 8.8b shows the score plot for the first and the second principal component and the corresponding loading plot. The score plot shows a distinction (line drawn) between the data obtained after incubation at 32 and 52 °C. Although the distinction is not absolute, a trend is observed that most spectra obtained after incubation at 32°C have a lower score for PC1 and a higher score for PC2 compared to the spectra obtained after incubation at 52°C. Therefore, these culturing conditions can be distinguished from each other by fatty acid information extracted from Raman spectra.

Table 8.3 Scores of the fatty acids for the first three loading vectors (mix spectra), resulting from the PCA's of the fatty acid reference Raman spectra of interest for each strain.

(a) *B. coagulans* LMG 6326			
	mix1	mix2	mix3
c14	5.34	20.87	−3.17
c15Aiso	−29.58	−2.07	−11.95
c15iso	19.40	−12.21	−8.04
c16	9.87	17.71	2.33
c16iso	7.01	−8.21	11.05
c17Aiso	−21.29	−3.15	12.17
c17iso	9.25	−12.95	−2.38
(b) *E. faecalis* LMG 7937			
	mix1	mix2	mix3
c12	15.30	15.41	9.45
c14	13.85	−13.58	4.85
c16	15.14	−9.36	−1.33
c18	14.89	7.12	−12.96
oleic acid	−29.28	0.24	1.10
vaccenic acid	−29.89	0.17	−1.12

Comparison of the score and loading plots shows that mix1 and mix2 (loading vectors 1 and 2 of the PCA performed on the fatty acid reference spectra) are related to the spectra recorded after incubation at 52 °C. From the scores of the PCA applied on the fatty acid reference spectra (Table 8.3a), it can be seen that mix1 has a negative contribution for the

Aiso fatty acids and a positive contribution for the other fatty acids. Therefore, we can conclude that incubation at 52 °C results in a lower Aiso content compared to incubation at 32 °C. This observation is confirmed by FAME analysis, as the sum of the percentages of c15Aiso and c17Aiso is higher for incubation at 32 °C (69.7%) than at 52 °C (67.0%). Analogous, mix2 has a positive contribution from the linear acids and a negative contribution of the branched acids, so there appear to be more linear acids in the cell at 52 °C. Again, this observation is confirmed by FAME analysis as the sum of the percentages of c14 and c16 is higher for incubation at 52 °C (7.2%) than at 32 °C (6.1%). For mix3 no conclusions were drawn, as PC3 in the PCA analysis on the fatty acid reference spectra did not make a logic distinction between the fatty acids.

A similar procedure was performed for the spectra of *E. faecalis* LMG 7937. First, PCA was performed on the reference spectra of the fatty acids of interest for this strain. Figure 8.9a shows the score plots of the first three principal components. PC1 makes a clear distinction between saturated and unsaturated fatty acids while PC2 shows no logic information about the fatty acids. For the saturated acids, it is clear that a higher score for PC3 is related to a lower number of carbon atoms in the fatty acids chain. Again, the first three loading vectors were used as new reference spectra (mix spectra), which are linear combinations of the fatty acid reference spectra and are orthogonal. These three mix spectra were used, together with reference spectra of other biomolecules, as 'good spectra' input for the EMSC procedure (datacase 106), that was applied on the bacterial Raman spectra. The coefficients of the three mix spectra, resulting from this EMSC procedure, were treated with PCA (Fig. 8.9b).

The score plot shows that the two different media form clearly distinct groups, based on the fatty acids information and that only PC1 is responsible for this separation. The loading plot shows that mix1 and mix3 have the highest contributions in PC1. From comparison of the score plot and the loading plot, it is clear that mix1 and mix3 are related to cultivation on MRS. The scores of the first PCA (Table 8.3b) show that mix1 has a positive contribution from the saturated fatty acids and a negative contribution from the unsaturated fatty acids. Therefore, it is expected that more saturated fatty acids are present in the cell when culturing on MRS instead of on TSA. Analogous, mix3 has a

positive contribution of oleic acid, c12 and c14, so the sum of the concentration of these acids is expected to be higher when culturing on MRS instead of on TSA. These two findings match the FAME data as the sum of the unsaturated acids (oleic and vaccenic acid) is 34.5% and 40.5% for MRS and TSA grown cells respectively, while the sum of c12, c14 and oleic acid is 13.9% and 4.9% for MRS and TSA grown cells respectively. No conclusions were drawn for mix2, because mix2 has a lower (absolute) contribution in PC1 and a higher one in PC2 that is not responsible for the distinction between data of the two media. Moreover, mix2 did not contain any logic information according to the PCA of the fatty acid reference spectra.

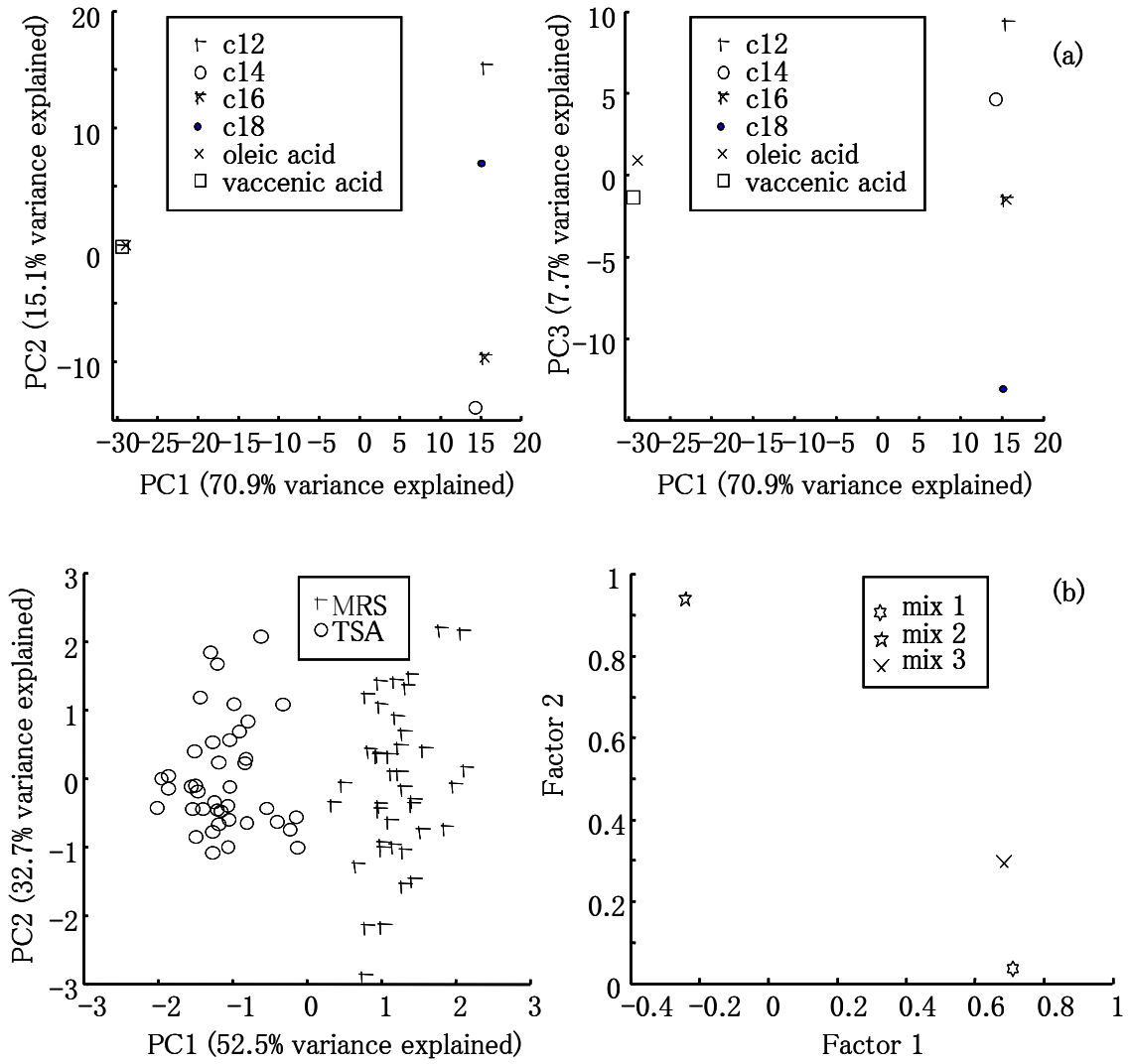


Figure 8.9 (a) Score plots from the PCA performed on the reference spectra of the fatty acids of interest for *E. faecalis* LMG 7937 and (b) score and loading plot from the PCA performed on the EMSC coefficients of the fatty acid mix spectra for *E. faecalis* LMG 7937.

8.4 Conclusions

In this study we explored the possibilities of Raman spectroscopy to provide information about the fatty acid composition of bacterial cells. Reference Raman spectra were recorded from several commercially available fatty acids. Most of these spectra could be distinguished from one another, although generally their Raman pattern is very similar which stresses the complexity of this study. The bacterial Raman spectra and their difference spectra are dominated by Raman bands from other biomolecules such as DNA and RNA bases, amino acids, etc. Furthermore, minor differences related to fatty acids are difficult to interpret because these bands contain contributions of several fatty acids. For a more detailed study of the data, we used the coefficients resulting from the EMSC procedures as values for the relative concentration of fatty acids in the cell. Due to the complexity of bacterial Raman spectra, it is very difficult to extract information about single fatty acids from these spectra. Indeed, PCA performed on the EMSC coefficients of the single fatty acids did not supply useful information, likely due to a colinearity problem in the EMSC procedure. Therefore, PCA was performed on the fatty acid reference spectra. Only three loading vectors (mix spectra) were used as 'good spectra' input for the EMSC procedure. The resulting coefficients of the mix spectra were subjected to PCA. In the score plots, the different culturing conditions form distinct groups. Thus, these culturing conditions could be easily distinguished from each other based on the fatty acid information in the bacterial Raman spectra. Comparison of the score and loading plots provided information about the relative amount of groups of fatty acids present in the cell. This information was consistent with the results from gaschromatographic FAME analysis. We conclude that Raman spectroscopy can provide information about fatty acid in bacterial cells, but the information is not as refined as that resulting from FAME analysis. We consider the described chemometric method useful to obtain fatty acids information when a whole cell profile has to be obtained, for which Raman spectroscopy is very suitable. Application of this chemometric method can be performed on similar problems that aim to achieve information from bacterial Raman spectra about any kind of biomolecule group with structurally very similar members.

This study shows that the use of an EMSC model for the estimation of the contribution of biomolecules to bacterial Raman spectra can be extended to highly similar biomolecules, provided PCA precedes the EMSC procedure. Unfortunately, in this way, it is not possible to obtain information about each of these compounds separately, but only about groups of them.

The previous chapters showed how reference spectra (Chapter 5) can be used to study bacterial compounds (Chapters 6, 7 and 8). Further research should focus on the vague borderline when to use the EMSC approach with or without precedent PCA analysis. The question remains whether also (semi-)quantitative information can be obtained from the biomolecules in bacterial cells. This will be dealt with in Chapter 9.

References

1. Afseth N.K., Segtnan V.H., Marquardt B.J., Wold J.P. (2005) Raman and near-infrared spectroscopy for quantification of fat composition in a complex food model system. *Appl. Spectrosc.* 59 : 1324-1332.
2. Beattie J.R., Bell S.E.J., Borgaard C., Fearon A., Moss B.W. (2006) Prediction of adipose tissue composition using Raman spectroscopy: Average properties and individual fatty acids. *Lipids* 41: 287-294.
3. De Gelder J., De Gussem K., Vandenabeele P., De Vos P., Moens L. (2007) Methods for extracting biochemical information from bacterial Raman spectra: An explorative study on *Cupriavidus metallidurans*. *Anal. Chim. Acta* 585: 234-240.
4. De Gelder J., De Gussem K., Vandenabeele P., Moens L. (2007) Reference database of Raman spectra of biological molecules. *J. Raman Spectrosc.* 38: 1133-1147.
5. De Gussem K., Vandenabeele P., Verbeken A., Moens L. (2007) Chemotaxonomical identification of spores of macrofungi: possibilities of Raman spectroscopy. *Anal. Bioanal. Chem.* 387: 2823-2832.
6. Hutsebaut D., Vandenabeele P., Moens L. (2005) Evaluation of an accurate calibration and spectral standardization procedure for Raman spectroscopy. *Analyst* 130: 1204-1214.
7. Hutsebaut D., Vandroemme J., Heyrman J., Dawyndt P., Vandenabeele P., Moens L., De Vos P. (2006) Raman microspectroscopy as an identification tool within the phylogenetically homogeneous '*Bacillus subtilis*'-group. *Syst. Appl. Microbiol.* 29: 650-660.
8. Kirschner C., Maquelin K., Pina P., Thi N.A.N., Choo-Smith L.P., Sockalingum G.D., Sandt C., Ami D., Orsini F., Doglia S.M., Allouch P., Mainfait M., Puppels G.J., Naumann D. (2001) Classification and identification of enterococci: a comparative phenotypic, genotypic, and vibrational spectroscopic study. *J. Clin. Microbiol.* 39: 1763-1770.

9. Laucks M.L., Sengupta A., Junge K., Davis E.J., Swanson B.D. (2005) Comparison of psychro-active arctic marine bacteria and common mesophillic bacteria using surface-enhanced Raman spectroscopy. *Appl. Spectrosc.* 59: 1222-1228.
10. Lopez-Diez E.C., Goodacre R. (2004) Characterization of microorganisms using UV resonance Raman spectroscopy and chemometrics. *Anal Chem* 76: 585-591.
11. Maquelin K., Choo-Smith L.P., Endtz H.P., Bruining H.A., Puppels G.J. (2002) Rapid identification of *Candida* species by confocal Raman micro spectroscopy. *J. Clin. Microbiol.* 40: 594-600.
12. Maquelin K., Krischner C., Choo-Smith L.P., van den Braak N., Endtz H.Ph., Naumann D., Puppels G.J. (2002) Identification of medically relevant microorganisms by vibrational spectroscopy. *J. Microbiol. Meth.* 51: 255-271.
13. Martens H., The EMSC toolbox for MATLAB
http://www.models.kvl.dk/source/emsctoolbox
14. Martens H., Nielsen J.P., Engelsen S.B. (2003) Light scattering and light absorbance separated by extended multiplicative signal correction. Application to near-infrared transmission analysis of powder mixtures. *Anal. Chem.* 75: 394-404.
15. Naumann D., Keller S., Helm D., Schultz C., Schrader B. (1995) FT-IR spectroscopy and FT-Raman spectroscopy are powerful analytical tools for the noninvasive characterization of intact microbial cells. *J. Mol. Struct.* 347: 399-504.
16. Oust A., Moretro T., Naterstad K., Sockalingum G.D., Adt I., Manfait M., Kohler A. (2006) Fourier transform infrared and Raman spectroscopy for characterization of *Listeria monocytogenes* strains. *App. Environ. Microb.* 72: 228-232.

Chapter 9: Monitoring poly-3-hydroxybutyrate in *Cupriavidus necator* DSM 428 with Raman spectroscopy

Joke De Gelder, Diana Willemse-Erix, Maarten J. Scholtes, Jorge I. Sanchez, Kees Maquelin, Peter Vandenabeele, Patrick De Boever, Gerwin J. Puppels, Luc Moens, Paul De Vos

Analytical Chemistry, 80 (2008) 2155-2160

This chapter explores the possibility of monitoring and of quantifying a specific bacterial compound. For this study, we selected poly-3-hydroxybutyrate (PHB), a compound that contributes highly to the bacterial Raman spectrum and shows an isolated band at 1734 cm^{-1}. In addition, bacterial PHB formation is of industrial importance for the production of bioplastics.

9.1 Introduction

Poly-3-hydroxybutyrate (PHB) is a biodegradable polymer that can be produced by certain bacteria in large amounts when unbalanced growth conditions such as a high carbon/nitrogen ratio are applied[1,26]. PHB is stored in granules and metabolized when the growth conditions become favorable or when lacking essential nutrients become available. Industries were interested in the production of PHB in microorganisms for the production of bioplastics[1,2]. Over the years, several species and culturing conditions have been evaluated in order to obtain an efficient PHB production. *Cupriavidus necator* (formerly *Ralstonia eutropha*) was the most preferred organism and fed-batch cultivation the most preferred method of production[23]. Oliveira *et al.*[21] described the characterization of PHB produced by *Cupriavidus necator* in solid state fermentation. Characterization of its chemical structure, thermal properties and crystalline morphology was performed by NMR and FTIR spectroscopy, differential scanning calorimetry, X-ray diffraction and polarizing optical microscopy. The results showed that solid state fermentation is an interesting alternative for submerged fermentation for the production of PHB with adequate properties. Obviously, the production rate needs to be monitored. The preferred methods for PHB quantification are GC-MS[20] and HPLC[9]. These labor-intensive techniques require a destruction of the samples and derivatization of PHB. A PHB ester is measured after GC by MS. In the case of HPLC, PHB is converted to crotonic acid monomers from which the absorbance at 210 or 214 nm is measured. As an alternative technique, flow cytometry[3,8] has been proposed. This requires a staining of a small amount of cells in suspension using the fluorescent lipid dye Nile Red. This dye penetrates cells in suspension and

preferentially accumulates in a lipid-rich environment such as PHB granules. Cell counts and fluorescence intensities can be recorded during the flow cytometry analysis using an excitation and emission wavelenth of 488 nm and 585 nm, respectively. Fourier transform infrared spectroscopy (FT-IR)[18] has also been described as an alternative for HPLC. On-line determination is possible using a sequential analysis flow system[7].

Raman spectroscopy has already been reported as a powerful tool for identification of species[10,13] or strains[6,15,22]. However, this technique also allows the extraction of information about cell components. Several bands in bacterial Raman spectra are already assigned to (groups of) biomolecules[16,19]. Because of band overlap, there is often a need for mathematical fitting of reference spectra to these bacterial Raman spectra, in order to obtain information about specific cell compounds[4] (Chapter 7). Raman spectroscopy is potentially suitable for the determination of PHB for process control, as it requires only small sample volumes and minimal sample preparation, the analysis is fast and on-line detection systems could be developed.

In this paper, the possibilities of Raman spectroscopy to follow PHB production are described. The Raman spectra of pure PHB and HB were collected. Raman spectra collected from *Cupriavidus necator* DSM 428 (H16) were studied for their contribution of PHB and HB signals. The possibility to monitor PHB production during growth was evaluated. In a separate setup, the results from Raman spectroscopy were compared to PHB concentrations determined by parallel HPLC analysis.

9.2 Experimental

9.2.1 Strains, culturing conditions and sampling

Cupriavidus necator DSM 428 (H16) producing PHB[25] and its mutant strain *Cupriavidus necator* DSM 541[24] unable to produce PHB, were cultured in 50 mL of a mineral medium (Table 9.1) inoculated from an overnight trypticase soy broth (TSB) culture to obtain a start $OD_{660\ nm}$ of 0.01.

Table 9.1 Composition of the mineral medium

Mineral medium: per L MilliQ-water*:	
Sodium acetate	2.4 g
$MgSO_4.7H_2O$	600 mg
NH_4Cl	160 mg
Sodium EDTA	100 mg
K_2HPO_4	92 mg
KH_2PO_4	45 mg
$CaCl_2.2H_2O$	70 mg
Spore solution	2 mL
Spore solution: per L MilliQ-water:	
$FeCl_3.6H_2O$	150 mg
H_3BO_3	150 mg
$CoCl_2.6H_2O$	150 mg
$MnCl_2.4H_2O$	120 mg
$ZnSO_4.7H_2O$	120 mg
$Na_2MoO_4.2H_2O$	60 mg
$CuSO_4.5H_2O$	30 mg

*MilliPore, Vergeze, France

To monitor PHB production and consumption, two sets of three cultures of DSM 428 and one culture of DSM 541 were shaken at 28 °C. The first set was sampled after 26, 28, 30, 32, 34, 36, 50, 52, 54, 56, 58 and 60 h of incubation, while the second set was sampled after 14, 16, 18, 20, 22, 24, 38, 40, 42, 44, 46 and 48 h of incubation. In this way, Raman measurements were performed every two hours between 14 and 60 h of incubation, in triplicate for DSM 428 and once for the PHB negative mutant DSM 541. Samples of 0.5 mL were washed 2 times with 0.5 mL of water (9000 × g, 1 min) and were transferred to a fused silica slide. The samples were dried for 15 min at 37 °C and analyzed with instrument 1 (see below).

For the comparison of Raman spectroscopic and HPLC analysis of PHB, samples were taken in parallel from separate cultures after 18, 22, 26, 30, 34, 36, 38, 40, 42, 46, 50 and 54 h of incubation. From each culture, two samples of 10 to 20 mL were taken for HPLC analysis and three samples of 2 to 4 mL were used for Raman spectroscopy (for cultures with lower OD values, larger samples are necessary). The samples were washed 2 times with 1 mL water (9000 × g, 2 min), transferred to a CaF_2 slide, dried for 10 min on silica and analyzed with instrument 2 (see below).

The range of the studied incubation times differs for the experiments on Raman instrument 1 and 2, because the use of different incubators causes differences in growth rate.

9.2.2 Products

Poly(3-hydroxybutyric acid) (natural origin) and (R)-(-)-3-hydroxybutyric acid (sodium salt, 98%) were purchased from Sigma-Aldrich (Belgium).

9.2.3 Raman spectroscopy

Instrument 1

Raman spectra were collected using a Model 2500 High Performance Raman Module (HPRM) (River Diagnostics BV, Rotterdam, The Netherlands), coupled to a custom-build inverted microscope stage, with an automated XY-stage (River Diagnostics) and operated using RiverICon software (River Diagnostics), version 1.63. The microscope contained a custom-designed microscope objective (River Diagnostics) with a numerical aperture of 0.7, 20× magnification, a working distance of 100mm, and optimized for Raman experiments in the 750-1000 nm wavelength region. The objective focused laser light emitted by the Model 2500 HPRM through the fused silica slide into the samples on top of the slide. The measurement volume of the instrument is approximately 5 μm in the lateral direction and 15 μm along the optical axis. The objective also collected Raman scattered light from the samples. Samples were excited using laser light from a 785 nm diode laser (Sacher Lasertechnik, Marburg, Germany), delivering approximately 150 mW to the sample. The spectrometer was calibrated according to the manufacturer's guidelines.
Automated data collection and signal pre-treatment was performed using the RiverICon software, requiring 100 s to sample 100 independent locations per sample. A correction for the signal contribution of the fused silica substrate was performed.

Instrument 2

For practical reasons, we used a Raman instrument that was located close to the HPLC instrument for the comparative study. Raman spectra were recorded with a Kaiser System Hololab 5000R modular Raman microspectrometer. Laser light of a 785 nm diode laser (Toptica Photonics AG, Grafelfing, Germany) was focused through a 100x objective of the microscope (Leica, Wetzlar, Germany) to obtain a power of 50-60 mW at the sample. The scattered light is transferred to the spectrograph by a confocal aperture collection fiber (15 µm N.A.) where it was detected by a back illuminated deep depletion Peltier cooled (-70 °C) CCD detector (Andor, Belfast, Northern Ireland). On each smear of the CaF_2 plate 4 Raman spectra of 60 s were collected around a central focus point. The calibration was performed as described by Hutsebaut *et al.*[5]. MATLAB version 6.5 (The Mathworks, Natick, Massachusetts) was used for data preprocessing.

9.2.4 Data analysis

Extended multiplicative signal correction (EMSC) was applied[17] and spikes were removed from the spectra. For the spectra collected on instrument 1 and 2 the average of the 100 and 4 spectra per sample, respectively, was calculated. Band intensities were obtained at the pixel with the highest signal for the band at 1734 cm^{-1}. The spectra collected on instrument 2, for the comparison with HPLC analysis, were subjected to a background subtraction (modified polynomial fit[11]) before calculating band intensities.

9.2.5 HPLC analysis

PHB content of the biomass was determined by acid hydrolysis of PHB to crotonic acid. Culture samples were centrifuged for 10 minutes at 7000 x g. The pellets were resuspended in 1 mL of distilled water and transferred to pre-weighted eppendorf tubes. Samples were centrifuged for 5 minutes at 13000 x g, dried overnight at 100 °C and weighted again for the determination of total biomass. Dried pellets were digested with 1 mL of 96% H_2SO_4 at

100 °C for 1 h to form crotonic acid. The reaction mixture was cooled to room temperature, samples were 50-fold diluted with distilled water and crotonic acid was determined by HPLC using a Dionex ASI-100 autosampler injector (Dionex Corporation, Sunnyvale, CA, USA) equipped with an Aminex HPX-87H ion-exchange organic acids column (300 x 7.8 mm) (BioRad). The solvent used was 0.014 N H_2SO_4 at a flow rate of 0.7 mL/min. The elution peaks were monitored at 210 nm with a Dionex UV detector. PHB content was calculated from a calibration curve for standards of commercial PHB (Goodfellow Cambridge Ltd., Huntingdon, England) treated in the same way as the samples.

9.3 Results and discussion

9.3.1 Raman spectra of PHB and its monomer HB

A Raman spectrum of commercially available PHB was recorded (Fig. 9.1a). The most prominent bands in this spectrum are located at 433, 839 and 1725 cm^{-1} and can be assigned to δ(CC) skeletal deformations, ν(CC) skeletal stretches and a ν(C=O) stretching vibrations[12]. The regions from 1040 to 1140 cm^{-1} and from 1250 to 1460 cm^{-1} show bands caused by respectively ν(CC) skeletal stretches and δ(CH), δ(CH_2), δ(CH_3) deformations[12]. When recording spectra of bacterial samples, all Raman active biomolecules will contribute to the resulting spectrum and their signals may overlap. Therefore, it is important to study reference spectra of compounds related to PHB, for example its monomer. Therefore, a spectrum of the sodium salt of β-3-hydroxybutyric acid (HB) was recorded (Fig. 9.1b). In this spectrum, the bands of δ(CC) skeletal deformations and ν(CC) skeletal stretches seem to be shifted to 483 and 853 cm^{-1} respectively. At 909 and 953 cm^{-1}, two bands of medium intensity are observed and can be assigned to ν(C-C-O) stretches of the hydroxyl group at the β-position in the monomer[12]. Because this monomer contains a COOH or COO^- function instead of an ester bond, its spectrum shows no band between 1720 and 1740 cm^{-1} [12]. Indeed, the ν(C=O) stretching vibration of carboxylic acids contributes only weakly in a Raman spectrum in the region between 1625 and 1687 cm^{-1}, while the

symmetric and asymmetric ν(C=O) stretching vibrations of carboxylic acid salts are observed at 1540–1695 cm^{-1} (weak) and 1360–1450 cm^{-1} (strong) respectively[12].

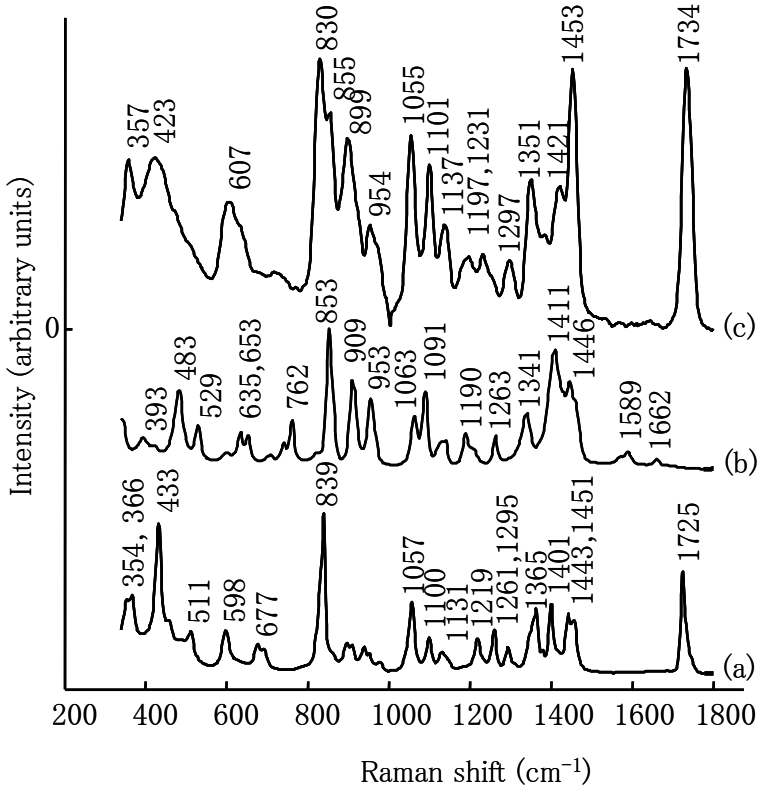


Figure 9.1 (a) Reference Raman spectra of poly-3-hydroxybutyrate (PHB) and (b) the sodium salt of β-3-hydroxybutyric acid (HB) and (c) difference spectrum between *C. necator* DSM 428 and the PHB negative mutant DSM 541 after 26 h of incubation (instrument 1).

9.3.2 Monitoring of PHB production and consumption of C. necator DSM 428 and its PHB negative mutant DSM 541 during growth

Raman spectra of *C. necator* DSM 428 and its PHB negative mutant DSM 541 were recorded during growth (14 to 60 h, step 2 h). Examples of these spectra for both strains collected after 26 h of incubation are shown in Figure 9.2 (band assignments in these spectra are summarized in Table 9.2). Compared to the spectrum of the PHB negative mutant DSM 541 (Fig. 9.2a), the spectrum of the strain DSM 428 (Fig. 9.2b) shows significant higher intensities at 1452 and 1734 cm^{-1}, and in the regions from 340 to 430, from 800 to 960 and from 1040 to 1130 cm^{-1}. The differences between the PHB producing

strain and the mutant strain were more pronounced when a difference spectrum was calculated between their spectra collected after 26 h of incubation time (Fig. 9.1c). Comparison of this difference spectrum to the reference spectrum of pure PHB (Fig. 9.1a), shows that the most prominent contributions of PHB to a bacterial Raman spectrum are located at ca. 357, 423, 607, 830, 1055, 1101, 1351, 1453 and 1734 cm^{-1}. Raman band positions of PHB in the reference spectrum and in the bacterial spectra may differ several wavenumbers, because of the different physicochemical state of PHB and a different matrix in bacteria. The reference product was measured as a dry powder, but in the cell PHB is present a suspended form inside granules[14].

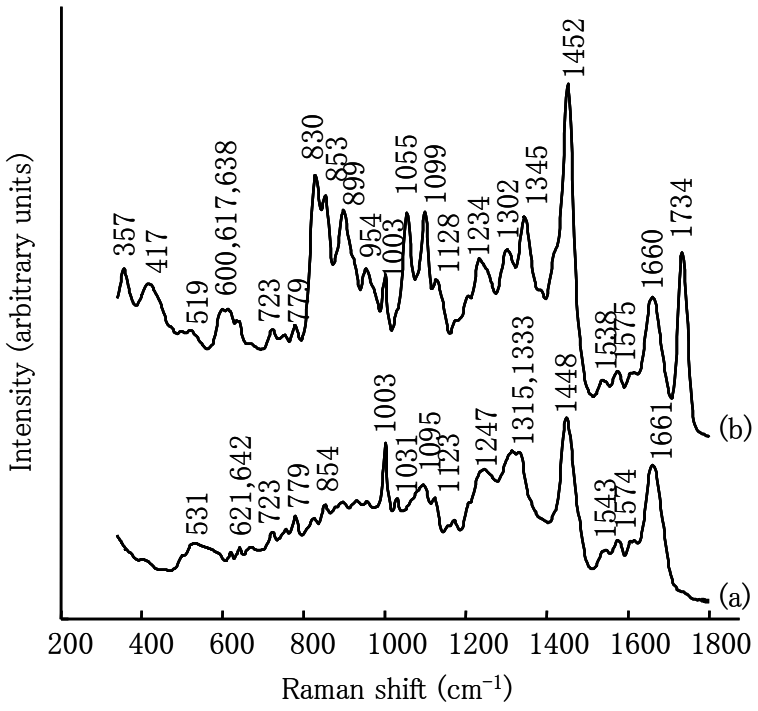


Figure 9.2 Raman spectra of (a) *C. necator* DSM 541 (PHB negative mutant) and (b) *C. necator* DSM 428 after 26 h of incubation (instrument 1).

In Figure 9.3, spectra of one culture series sampled at 14, 24, 38 and 48h are plotted. This figure shows that the above mentioned bands that were assigned to PHB vary as a function of incubation time. Because the band at 1734 cm^{-1} is isolated in the spectra of DSM 428 and not present in the spectra of the PHB negative mutant DSM 541 (Fig. 9.2a), its intensity seems to be a potential marker for the PHB content of bacteria. A possible contribution from the monomer of PHB in the bacterial Raman spectra was considered.

Indeed, the HB bands at 853, 909 and 953 cm^{-1} (Fig. 9.1b) could also contribute in the difference spectrum (Fig. 9.1c). Moreover, these bands seem to vary as a function of incubation time in the bacterial spectra in Figure 3, analogously to the bands of PHB. This observation confirms the contribution of HB in the bacterial spectra. The Raman spectrum of the monomer HB (Fig. 9.1b) does not show bands above 1700 cm^{-1}, which suggests that the band at 1734 cm^{-1} in bacterial Raman spectra contains a dominant contribution of the polymer PHB.

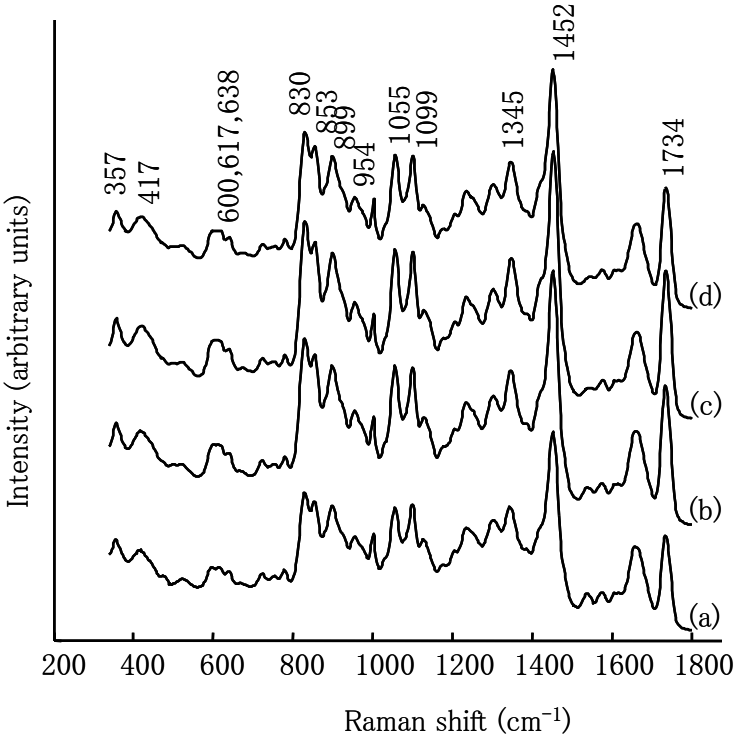


Figure 9.3 Mean spectra per incubation time of some spectra of *C. necator* DSM 428 in culture series 1: (a) 14 h, (b) 24 h, (c) 38 h and (d) 48 h (instrument 1).

Therefore, we used the intensity of the band at 1734 cm^{-1} as a marker for the PHB concentration in the cells. Band intensities at 1734 cm^{-1} were calculated for all Raman measurements of *C. necator* DSM 428 and DSM 541 and plotted as a function of the incubation time (Fig. 9.4). The error bars represent the standard deviation on the 100 independent measurements that were collected for each sample. Because cells in a culture are not all in the same state, these bars mainly represent within-culture variation instead of instrumental variation. Especially in the case of PHB, which occurs in granules, there is heterogeneity within a culture. As can be seen from the error bars, this

heterogeneity is larger in the exponential phase than in the stationary phase. For the PHB negative mutant DSM 541, the intensity is limited to the background level and no evolution during incubation time is observed. For DSM 428 a clear evolution of production and consumption of PHB can be observed by the Raman intensity at 1734 cm^{-1}. For the applied culturing conditions, the maximum PHB yield was reached after 28 to 30 h of incubation. An offset is observed between the two culture series due to different microbial growth, as these series were initiated from different inocula.

Table 9.2 Band assignments for the Raman spectra of *C. necator* DSM 428 and the PHB negative mutant DSM 541.

Raman bands (cm^{-1}) in the spectra of		Band assignments in bacterial Raman spectra*
C. necator DSM 428	*C. necator* DSM 541 (mutant)	
357 (m)		PHB
417 (m,br,400-450)		PHB
519 (w)	531 (m)	ν(S-S)
600 (m)		PHB
617 (m)	621 (w)	Phe
638 (m)	642 (w)	Tyr
	669 (w)	G
723 (w)	723 (m)	A
755 (w)	757 (w)	
779 (w)	779 (m)	C, U
830 (vs)	825 (m)	Tyr, PHB
853 (s)	853 (m)	Tyr, HB
899 (s)	896 (w)	ν(COC), HB
	933 (w)	
954 (m)	957 (w)	HB
1003 (m)	1003 (vs)	Phe
	1031 (m)	
1055 (s)		PHB, HB
1101 (s)	1095 (s)	PHB, HB, ν(>PO_2^-) sym
1128 (m)	1123 (m)	ν(C-N), ν(C-C)
1175 (w,sh)	1171 (w)	
1205 (m,sh)		
1234 (m)	1247 (s)	Amide III
1302 (m)		
	1315 (s)	
1345 (s)	1332 (s)	
1452 (vs)	1448 (vs)	δ(CH_2), e.g. PHB, HB
1538 (w)	1543 (m)	
1575 (w)	1574 (m)	G, A
1617 (w,sh)	1616 (w,sh)	Tyr
1660 (s)	1661 (vs)	Amide I
1734 (vs)		PHB

* 'PHB', 'HB' indicate contributions of these molecules in the Raman spectra of DSM 428
w = weak, m = medium, s = strong, vs = very strong, sh = shoulder, br = broad

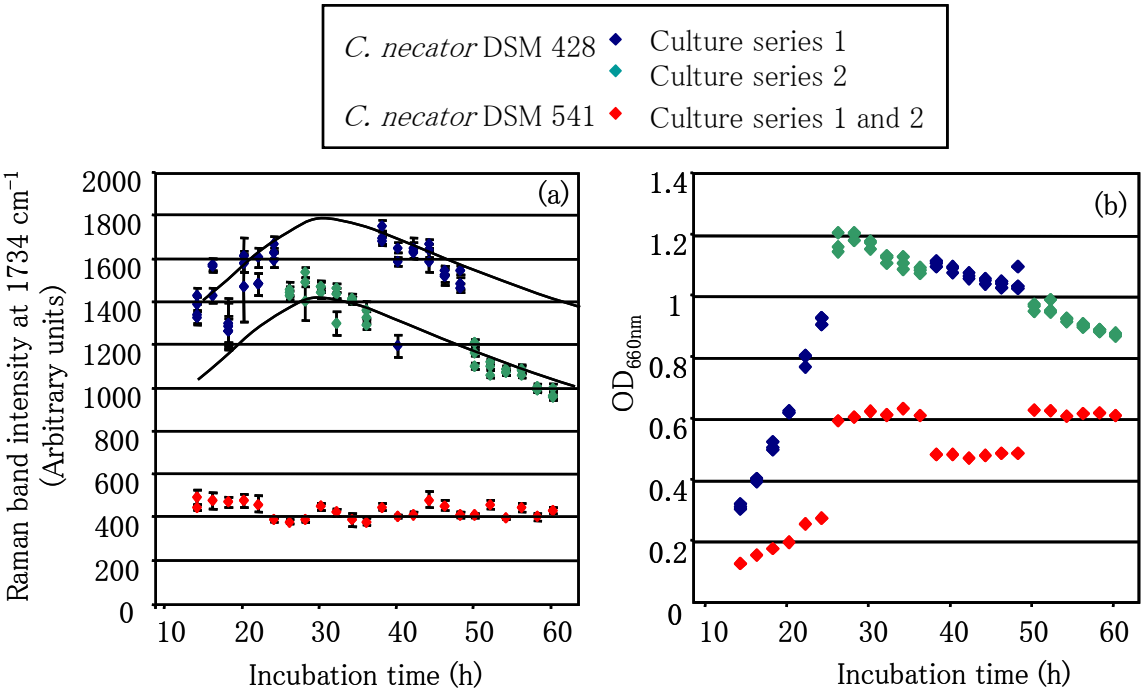


Figure 9.4 (a) Band intensities at 1734 cm^{-1} (PHB) in the Raman spectra of *C. necator* DSM 428 and the PHB negative mutant DSM 541 recorded during growth (instrument 1). The error bars represent the standard deviation of the 100 independent measurements collected per sample and reflect mainly within-culture variation. For sake of clarity, manually drawn curves have been added to indicate the evolution of the two culture series of DSM 428. (b) OD_{660nm} as a function of incubation time, measured in parallel to the Raman experiment.

9.3.3 Determination of the PHB content of C. necator DSM 428: comparison of Raman band intensities and absolute concentrations obtained from HPLC analysis

In order to relate the intensity of the Raman feature at 1734 cm^{-1} with the PHB content of the cultured cells, a comparative study with HPLC was performed. Twelve analyses of independently grown cultures between 18 and 54 hours of incubation were performed for *C. necator* DSM 428 to cover a large range of the assumed PHB content. Therefore, the twelve cultures were prepared by adding an overnight TSB grown inoculum to 50 mL mineral medium so OD_{660nm} 0.01 was obtained. After the appropriate incubation time, these cultures were sampled in parallel for Raman spectroscopy and HPLC analysis. From each culture (50 mL), three Raman samples (2-4 mL) and two HPLC samples (10-20 mL) were taken. A background subtraction based on a modified polynomial fit[11] was performed before

calculating the Raman band intensities at 1734 cm^{-1} (Fig. 9.5). Raman bands intensities and absolute concentrations (mg/L culture) are given in Table 9.3.

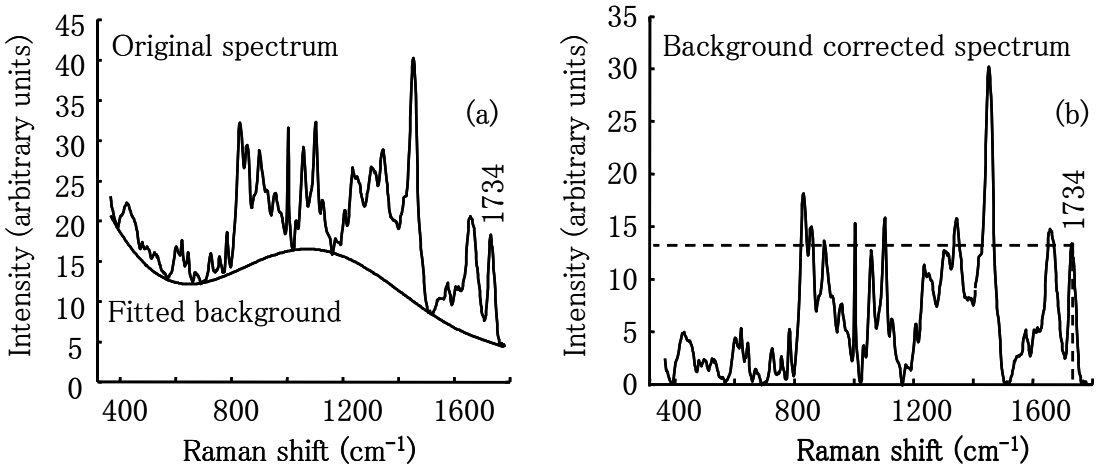


Figure 9.5 Illustration of (a) the modified polynomial fit, (b) background subtraction and calculation of band intensity at 1734 cm^{-1}, applied on a Raman spectrum of *C. necator* DSM 428 incubated for 36 h (instrument 2).

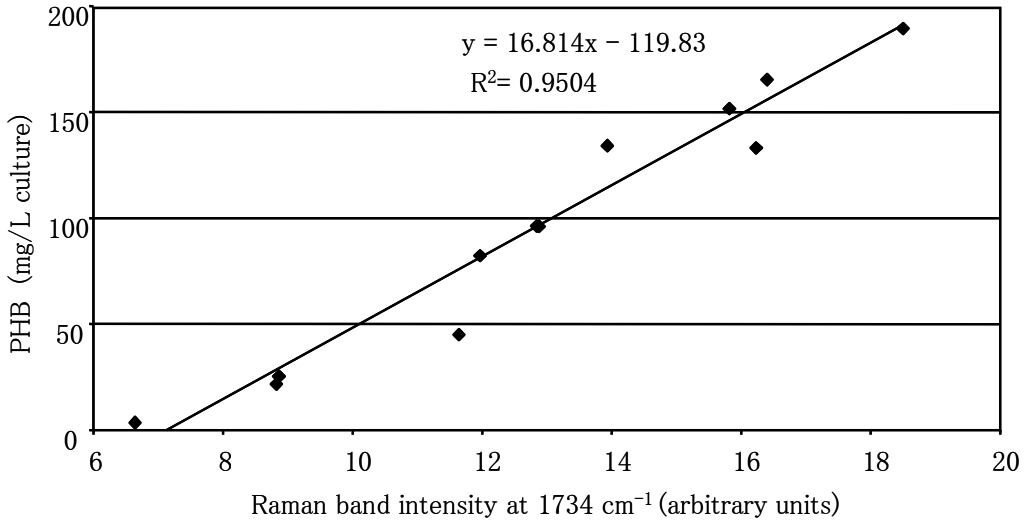


Figure 9.6 Determination of the PHB content of *C. necator* DSM 428: comparison of the relative values deduced from Raman spectra and the absolute concentrations (mg/L culture) determined by HPLC analysis (instrument 2).

PHB concentrations determined by HPLC can vary substantially, probably due to variations in sample preparation (recovery for the conversion to crotonic acid). The variations in Raman band intensities are smaller and likely to be caused by a combination of heterogeneity in PHB formation between the individual cells of a culture and the small

sampling volume of the laser (ca. 22 μm^3). Per incubation time (per culture), the average values were calculated for Raman spectroscopy as well as for HPLC analysis. In Figure 9.6 these values are plotted, in order to visualize the relation between the values obtained for both techniques. The Raman band intensity at 1734 cm^{-1} was related linearly (R^2 = 0.95) with the concentration of PHB determined by HPLC and expressed as mg PHB per L culture. This observation shows the potential of Raman spectroscopy for the quantification of PHB in bacterial cells and the ability to qualitatively follow the PHB production as a function of culture time.

Table 9.3 Raman band intensities at 1734 cm^{-1} and PHB content (mg/L culture) determined by parallel HPLC analysis.

Incubation time (h)	PHB content determined by				
	Raman band intensities at 1734 cm^{-1} (arbitrary units)			HPLC (mg/L culture)	
18	6.55	6.40	6.88	4.80	4.20
22	9.55	8.49	8.46	20.99	31.52
26	9.36	7.83	9.20	18.42	26.59
30	11.29	11.24	12.31	42.46	49.30
34	11.65	11.05	13.11	103.10	63.20
36	13.37	11.86	13.33	114.10	79.80
38	13.34	11.60	13.49	115.10	79.20
40	13.01	14.58	14.12	154.40	115.50
42	15.74	16.20	16.66	145.80	121.97
46	18.80	18.99	17.61	184.73	195.92
50	16.29	16.42	16.41	155.06	177.36
54	15.60	15.87	15.89	146.56	158.47

9.4 Conclusions

The presence of PHB in *C. necator* DSM 428 and most likely in other bacterial species can be deduced from its contribution in bacterial spectra to Raman bands at ca. 357, 423, 830, 1055, 1101, 1351, 1453 and 1734 cm^{-1}. The band at 1734 cm^{-1} seems to be suitable for monitoring PHB, as no contribution from other biomolecules is observed at 1734 cm^{-1} in the Raman spectra of the mutant strain DSM 541 which is unable to produce PHB. Moreover, the ν(C=O) stretching vibration of the monomer, hydroxybutyric acid, is located

below 1700 cm^{-1} and has weak intensity. Therefore, production and consumption of PHB in *C. necator* DSM 428 could be observed as a function of the incubation time by the Raman band intensity at 1734 cm^{-1}. For the PHB negative mutant, these intensities remained at the background level during the observed growth, while for the PHB producing strain DSM 428, the Raman band intensities at 1734 cm^{-1} are linearly related to the absolute concentrations of PHB (mg/L culture) determined by HPLC (R^2 = 0.95).

We conclude that Raman spectroscopy could be used for the fast monitoring of PHB production and consumption. Raman band intensities at 1734 cm^{-1} can be used as relative values for the PHB content in bacteria and show potential for its quantitative determination.

This chapter shows the potential of Raman spectroscopy for quantitative determination of a bacterial cell compound. Further research could try to determine the absolute concentration of a compound directly by using, for instance, a PLS model. However, this approach requires the possibility to prepare reference samples with different known concentrations of that compound, so that their Raman spectra could be used for calibration of the model.

The previous chapters showed the applicability of reference spectra and several processing methods for studying bacterial compounds with Raman spectroscopy. Chapter 10 reports on a case study where some of the described methods were applied to study the influence of microgravity on Ralstonia metallidurans LMG 1195 with Raman spectroscopy.

References

1. Anderson A.J., Dawes, E. A. (1990) Occurrence, metabolism, metabolic role, and industrial uses of bacterial polyhydroxyalkanoates. *Microbiol. Rev.* 54: 450-472.
2. Byrom D. (1987) Polymer synthesis by microorganisms - technology and economics. *Trends Biotechnol.* 5: 246-250.
3. Degelau A., Scheper T., Bailey J.E., Guske C. (1995) Fluorometric measurement of poly-beta hydroxybutyrate in *alcaligenes-eutrophus* by flow-cytometry and spectrofluoromety. *Appl. Microbiol. Biotechnol.* 42: 653-657.
4. De Gelder J., De Gussem K., Vandenabeele P., De Vos P., Moens L. (2007) Methods for extracting biochemical information from bacterial Raman spectra: An explorative study on *Cupriavidus metallidurans*. *Anal. Chim. Acta* 585: 234-240.
5. Hutsebaut D., Vandenabeele P., Moens L. (2005) Evaluation of an accurate calibration and spectral standardization procedure for Raman spectroscopy. *Analyst* 130: 1204-1214.
6. Hutsebaut D., Vandroemme J., Heyrman J., Dawyndt P., Vandenabeele P., Moens L., De Vos P. (2006) Raman microspectroscopy as an identification tool within the phylogenetically homogeneous *Bacillus subtilis*-group. *Syst. Appl. Microbiol.* 29: 650-660.
7. Jarute G., Kainz A., Schroll G., Beana J. R., Lendl B. (2004) On-line determination of the intracellular poly(beta-hydroxybutyric acid) content in transformed *Escherichia coli* and glucose during PHB production using stopped-flow attenuated total reflection FT-IR Spectrometry. *Anal. Chem.* 76: 6353-6358.
8. Kacmar J., Carlson R., Balogh S.J., Srienc F. (2006) Staining and quantification of poly-3-hydroxybutyrate in *Saccharomyces cerevisiae* and *Cupriavidus necator* cell populations using automated flow cytometry. *Cytom. Part A* 69A: 27-35.
9. Karr D.B., Waters J.K., Emerich D.W. (1983) Analysis of poly-beta-hydroxybutyrate in *Rhizobium-japonicum* bacteroids by ion-exclusion high-pressure liquid-chromatography and UV detection. *Appl. Environ. Microbiol.* 46: 1339-1344.

10. Kirschner C., Maquelin K., Pina P., Thi N.A.N., Choo-Smith L.P., Sockalingum G.D., Sandt C., Ami D., Orsini F., Doglia S.M., Allouch P., Mainfait M., Puppels G.J., Naumann D. (2001) Classification and identification of enterococci: a comparative phenotypic, genotypic, and vibrational spectroscopic study. *J. Clin. Microbiol.* 39: 1763-1770.
11. Lieber C.A., Mahadevan-Jansen A. (2003) Automated method for subtraction of fluorescence from biological Raman spectra. *Appl. Spectrosc.* 57: 1363-1367.
12. Lin-Vien D., Colthup N.B., Fateley W.G., Grasselli J.G. (1991) The handbook of Infrared and Raman characteristic frequencies of organic molecules. Academic press, San Diego.
13. Lopez-Diez E.C., Goodacre R. (2004) Characterization of microorganisms using UV resonance Raman spectroscopy and chemometrics. *Anal. Chem.* 76: 585-591.
14. Madigan M.T., Martinko J.M., Parker J. (2003) Brock Biology of Microorganisms. 10th ed., Pearson Education Inc., Upper Saddle River.
15. Maquelin K., Dijkshoorn L., van der Reijden T.J.K., Puppels G.J. (2006) Rapid epidemiological analysis of *Acinetobacter* strains by Raman spectroscopy. *J. Microbiol. Meth.* 64: 126-131.
16. Maquelin K., Krischner C., Choo-Smith L.P., van den Braak N., Endtz H.Ph., Naumann D., Puppels G.J. (2002) Identification of medically relevant microorganisms by vibrational spectroscopy. *J. Microbiol. Meth.* 51: 255-271.
17. Martens H., Stark E. (1991) Extended multiplicative signal correction and spectral interference subtraction - new preprocessing methods for near-infrared spectroscopy. *J. Pharmaceut. Biomed.* 9: 625-635.
18. Misra A.K., Thakur M.S., Srinivas P., Karanth N.G. (2000) Screening of poly-beta-hydroxybutyrate-producing microorganisms using Fourier transform infrared spectroscopy. *Biotechnol. Lett.* 22: 1217-1219.
19. Naumann D., Keller S., Helm D., Schultz C., Schrader B. (1995) FT-IR spectroscopy and FT-Raman spectroscopy are powerful analytical tools for the noninvasive characterization of intact microbial cells. *J. Mol. Struct.* 347: 399-504.

20. Odham G., Tunlid A., Westerdahl G., Mardén P. (1986) Combined determination of poly-beta-hydroxyalkanoic and cellular fatty-acids in starved marine-bacteria and sewage-sludge by gas-chromatography with flame ionization or mass-spectrometry detection. *Appl. Environ. Microbiol.* 52: 905-910.
21. Oliveira F.C., Dias M. L., Castilho L.R., Freire D.M.G. (2007) Characterization of poly(3-hydroxybutyrate) produced by *Cupriavidus necator* in solid-state fermentation. *Bioresource Technol.* 98: 633-638.
22. Oust A., Moretro T., Naterstad K., Sockalingum G.D., Adt I., Manfait M., Kohler A. (2006) Fourier transform infrared and Raman spectroscopy for characterization of *Listeria monocytogenes* strains. *App. Environ. Microb.* 72: 228-232.
23. Patnaik P.R. (2006) Dispersion optimization to enhance PHB production in fed-batch cultures of *Ralstonia eutropha*. *Bioresource Technol.* 97: 1994-2001.
24. Schlegel H.G., Lafferty R., Krauss I. (1970) Isolation of mutants not accumulating poly-beta-hydoxybutyric acid. *Arch. Microbiol.* 71: 283-294.
25. Schlegel H.G., Vonbartheld R., Gottschalk G. (1961) Formation and utilization of poly-beta-hydroxybutyric acid by knallgas bacteria. *Nature* 191: 463-465.
26. Shi H.D., Shiraishi M., Shimizu K. (1997) Metabolic flux analysis for biosynthesis of poly(beta-hydroxybutyric acid) in *Alcaligenes eutrophus* from various carbon sources. *J. Ferment. Bioeng.* 84: 579-587.

Chapter 10: Raman spectroscopic analysis of *Cupriavidus metallidurans* LMG 1195 cultured in low-shear microgravity conditions (case study)

Joke De Gelder, Peter Vandenabeele, Patrick De Boever, Max Mergeay, Luc Moens, Paul De Vos

Microgravity Science and Technology, Submitted

This chapter presents a case study of Ralstonia metallidurans LMG 1195 cultured in microgravity conditions and a control setup. The cellular differences induced by microgravity are subtle and so it is challenging to identify them. Some of the processing techniques presented in Chapter 6 and 7 were used to extract relevant information about these differences from bacterial Raman spectra.

10.1 Introduction

Future long-term manned space trips demand the possibility to recycle waste en produce food in a space environment. Therefore, microbial based life support systems are being developed[9,11,12]. These systems can only be used in a reliable way if one understands the behaviour of microorganisms in spaceflight conditions (e.g. microgravity and cosmic radiation). Currently, studies are being performed in the orbital station International Space Station (ISS) to investigate this issue[2,21]. Wilson *et al.*[39] have shown that the pathogenicity of a *S. thyphimurium* strain increased when it was cultured in space flight conditions (on the Space Shuttle). This observation is of paramount importance for estimating the possible impact of bacteria on the astronaut's health. The knowledge about the behaviour (pathogenicity in particular) of bacteria in a space flight environment is of importance in the frame of long-term manned space missions because the presence of bacteria can not be avoided during the missions[15,32]. Space flight experiments are limited, extremely expensive and have many experimental constraints (limited access to the ISS, small experimental setups, little experimental control, etc.). The application of laboratory models to simulate certain aspects of spaceflight has proven to be useful to overcome a number of these practical constraints. Furthermore, these models can be used to elaborate hypotheses about the effects of spaceflight conditions on living cells. In this respect, the Rotating Wall Vessel (RWV) technology was developed by the National Aeronautics and Space Administration (NASA), and is used in the field of space research to simulate the low-shear environment inherent to microgravity[17]. The RWV is a cylindrical bioreactor which can be completely filled with medium and which is rotated on an axis parallel with the ground.

Subsequently, a solid body mass rotation of the culture medium is obtained creating a low fluid-shear environment (calculated to be less than 0.001 Pa[10,28]) allowing bacterial cells to remain in suspension in a restricted orbit[23]. Several research groups apply the RVW technology for cultivation of various bacteria[8,22,38,39]. Wilson *et al.*[39] reports that differences detected between cultures from the Space Shuttle and an identical ground control setup were confirmed with the ground based microgravity culture model in the RWV. This strengthens the assumption that the RWV can mimic (only) low-shear microgravity effects. In addition to space-flights, microgravity studies can be useful in ground-based applications. Studying changes in pathogenesis, metabolism and physiology induced by microgravity can lead to a better understanding of their mechanisms[30]. Studies, for space as well as ground-based applications, have already reported differences in gene expression, physiology and pathogenesis and are summarized in Nickerson *et al.*[31] and Leys *et al.*[21]. However, some processes such as DNA repair do not seem to be altered by microgravity[18,34].

Raman spectroscopy is a vibrational spectroscopic technique. Raman band positions are related to certain vibrations and therefore the spectra provide information about the composition of a sample. A Raman spectrum recorded from bacterial cells is thus a superposition of signals from most cell constituents. Raman spectroscopy has been explored for its possibilities in microbiology and has proven to be useful for the identification of microorganisms on the species[19] and even strain level[14,24,33]. In addition, information on the composition of bacterial cells can be obtained[3,5,25,29]. In this work we used Raman spectroscopy to compare cultures of *Cupriavidus metallidurans* LMG 1195 that were grown under simulated microgravity (SMG) conditions and a control setup, both in the RWV. *Cupriavidus metallidurans* was chosen because it was used in previous studies as a model organism for studying the influence of microgravity[27]. Our aim was to show that Raman spectroscopy is a useful technique 1) for the discrimination between bacterial cells grown in SMG and control conditions and 2) for the extraction of information about the constitution and/or metabolomics of bacterial cells.

10.2 Materials and methods

10.2.1 Strain and culturing conditions

The inoculation of the Rotating Wall Vessel (RWV) bioreactors (Cellon, Bereldange, Luxembourg) occurred as follows: an overnight shaken culture of *Cupriavidus metallidurans* LMG 1195 was diluted in fresh mineral medium to obtain OD(590nm) of 0.01. A detailed composition of the medium is given by De Gelder *et al.*[3] (Chapter 7). The RWV bioreactors were filled completely with circa 50 mL of the diluted cell culture. In order to avoid undesired shear stress, air bubbles were carefully removed through the sampling ports using syringes (without needle). The vessels were mounted on separate RWV devices in an incubator. One device was put in a horizontal and one in a vertical position representing the control and the SMG setup, respectively (Fig. 10.1). Gas exchange in the RWV during growth was ensured by the gas-permeable silicone membrane present at the back of each RWV culture vessel. Evaporation of the culture medium through the gas-permeable silicone membrane was minimized by enlarged humidity in the incubator (approximately 60%). Bacterial growth in RWV conditions was allowed for 24 or 48 h at 28 °C. The vessels were continuously rotated at 20 rpm in forward direction. The experiments were repeated up to 6 times on different days. OD(590nm) values were recorded for each culture.

10.2.2 Raman spectroscopy

The Raman spectrometer and its calibration are described by Hutsebaut *et al.*[13]. For Raman analysis, 2 samples of 2 mL culture were taken from each vessel and were centrifuged at 15000 × g for 2 minutes. The pellets were washed three times with 1 mL physiological water. The final centrifugation lasted 10 minutes in order to obtain a dense pellet. The pellets were transferred to a CaF_2 plate and dried for 15 minutes on silica. On each smear, a focus point was chosen on the CaF_2 plate and four spectra were recorded in the surrounding area, each with a collection time of 60 s.

10.2.3 Data processing

Data processing of the Raman spectra were performed using Matlab version 6.5 (The Mathworks, Natick, Massachusetts) and SPSS 12 (SPSS inc., Chicago, Illinois). Principal components analysis (PCA) was performed on the spectra, which is a data reduction method that summarizes most variability (read 'information') of a dataset in a few new variables (principal components, PC's). A plot of the first two PC's (called score plot) shows the relatedness between the samples. In this way, groups of samples with different composition can be distinguished from one another. The calculation of difference spectra and the use of an extended multiplicative signal correction (EMSC) procedure to extract chemical information from the Raman spectra, are described by De Gelder *et al.*[3].

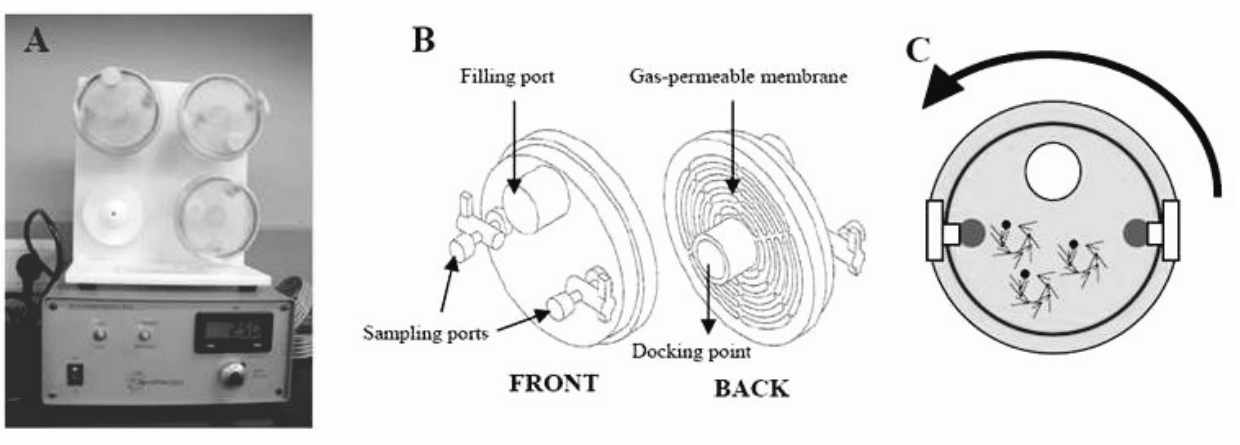


Figure 10.1 The rotating wall vessel (RWV) technology: (a) vertical setup to generate low-shear microgravity, (b) the RWV bioreactors and (c) rotation of the RWV in this way the cells are maintained in suspension in a restricted fluid orbit.

10.3 Results

Raman spectra were recorded from *Cupriavidus metallidurans* LMG 1195 cultured for 24 h in the RWV, one vessel mounted in SMG setup and one in control setup. From each culture, the OD(590nm) value was determined and a sample was taken for Raman measurement. A paired T-test (between smg en controle samples originating from the same inoculum) showed that the OD(590nm) values of the smg and controle cultures do not significantly differ from each other on the 95% confidential level. All spectra were subjected

to autoscaling (mean centering for each Raman shift) and subsequent principal component analysis (PCA). The score plots did not show a clear distinction between the spectra from samples cultured under SMG and the control condition (data not shown). This suggests that Raman spectroscopy as expressed in PCA can not detect differences in cell composition induced by microgravity above the day-to-day variations caused by the use of different inocula.

Mean spectra were calculated for all SMG and control samples. These mean spectra (Fig. 10.2) show subtle differences (indicated by arrows). To clearly visualize these differences, the difference was calculated between the mean SMG spectrum and the mean control spectrum was calculated. The resulting spectrum, further referred to as difference spectrum (Figure 10.3a, microgravity minus control spectrum), shows some clear bands that can be related to the storage polymer poly-3-hydroxybutyrate (PHB). Several bands in the difference spectrum (434, 600, 839, 1058, 1100, 1360-1460 and 1721 cm^{-1}) match the most intense bands of the reference Raman spectrum of pure PHB (Fig. 10.3b). In addition, some bands of the reference Raman spectrum of PHB's monomer β-3-hydroxybutyrate (HB) (Fig. 10.3c) such as the bands at 853 and 916 cm^{-1} can be observed in the difference spectrum (Fig. 10.3a). Band positions in the bacterial spectra can differ slightly from those in the reference spectra due to a different physicochemical state in the cell compared to the reference product. However, it has been shown that the bands in the bacterial spectra that were assigned to PHB, indeed reflect the PHB concentration in the cell[6] (Chapter 9). These results are thus indicative for a higher concentration of PHB in *C. metallidurans* LMG 1195 when grown for 24 h in SMG compared to control conditions. This observation was checked in the individual difference spectra for each cultivation and measurement day (SMG spectrum minus control spectrum, pair-wise per measurement day and thus per inoculum). All these spectra (not shown) contained bands (to a variable extent) of PHB and HB and it can therefore be concluded that there is indeed an indication for more PHB production in SMG after 24 h of incubation. The reason for not observing two distinct groups according to microgravity and control setup when performing PCA, is probably that the main spectral differences are caused by the use of different inocula and thus uncontrolled fluctuations in growth.

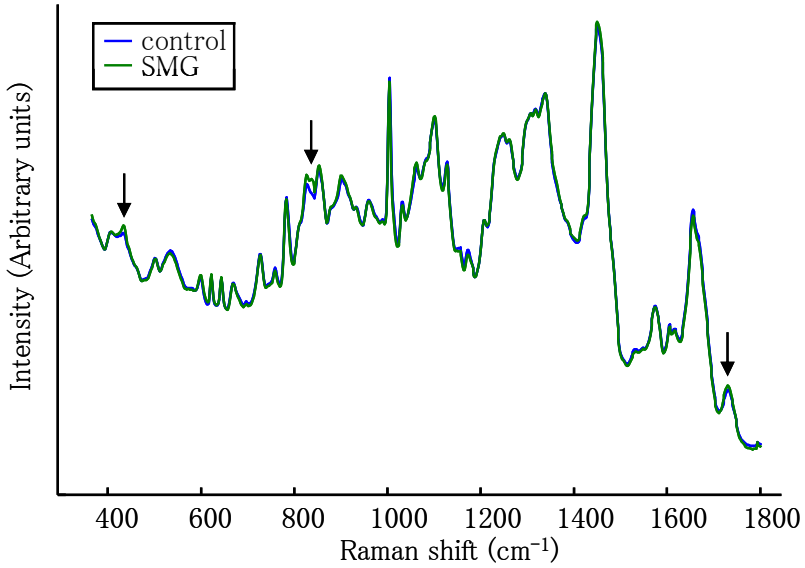


Figure 10.2 Mean Raman spectra of *Cupriavidus metallidurans* LMG 1195 cultivated for 24 h in a RWV simulating low-shear microgravity (SMG) and a control setup.

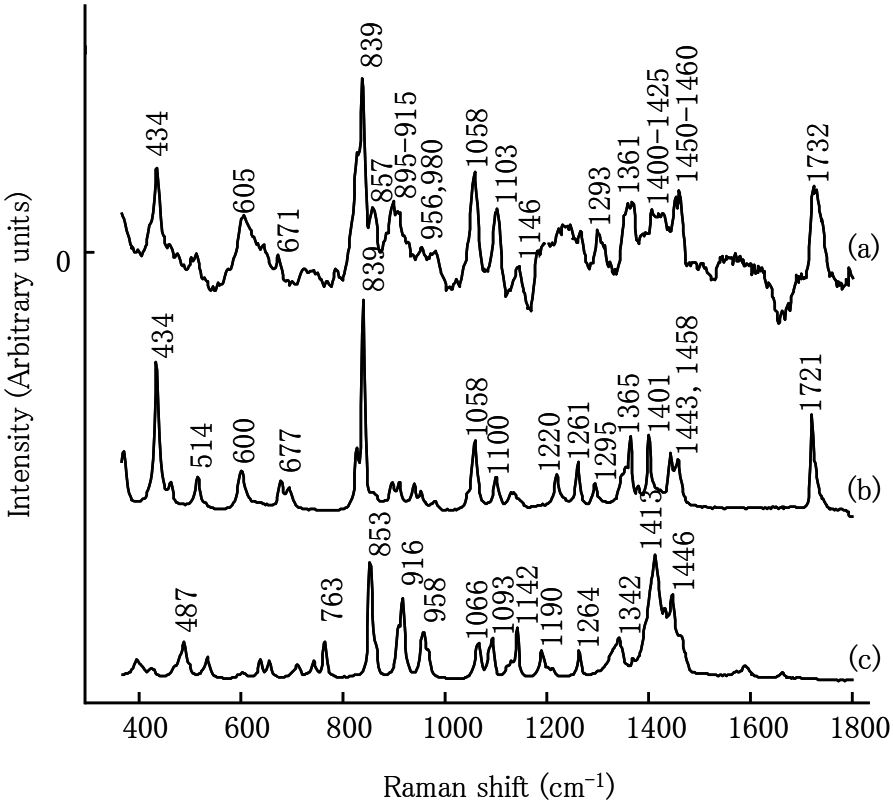


Figure 10.3 (a) Difference spectrum between the two mean Raman spectra (Fig. 10.2) of *Cupriavidus metallidurans* LMG 1195 cultivated for 24 h in a RWV simulating low-shear microgravity (SMG) and a control setup (SMG minus control) and reference Raman spectra of (b) poly-3-hydroxybutyrate (PHB) and of (c) its monomer β-3-hydroxybutyrate (HB).

Next, a mathematical procedure, known as extended multiplicative signal correction (EMSC)[26], was applied to extract information about spectral differences smaller than those caused by PHB and HB[3] (Chapter 7). This procedure allows estimating the contribution of several biomolecules to a bacterial Raman spectrum by fitting reference spectra of these biomolecules from a database present in our lab[4] (Chapter 5). The EMSC technique identified several biomolecules that contributed differentially to the SMG and the control spectra. Next to differences concerning PHB and HB, higher levels of acetyl-CoA, citric acid and malic acid were found in the SMG compared to the control, whereas the contribution of succinic acid was more pronounced in the spectra of control experiment.

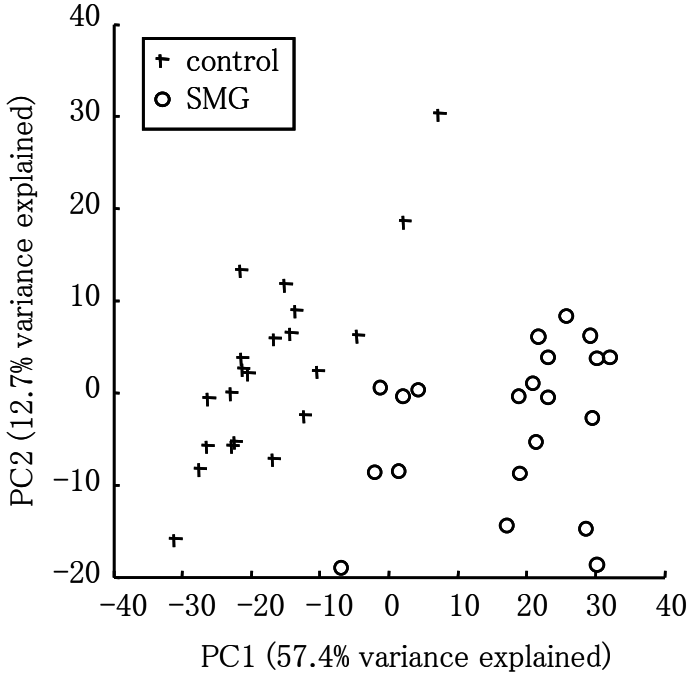


Figure 10.4 Score plot of the first two principal components of the PCA of the spectra from *Cupriavidus metallidurans* LMG 1195 cultivated for 48 h in a RWV simulating low-shear microgravity (SMG) and a control setup.

Similar analyses were performed on Raman spectra that were recorded after a cultivation time of 48 h cultures (spectra not shown). The mean spectra for the SMG conditions and the control setup showed much larger differences than the 24h-old cultures. After autoscaling of the spectra and subsequent PCA, the score plot of the first two principal components clearly shows two distinct groups of samples according to the culturing condition (smg or control) (Fig 10.4). Raman spectroscopy can thus distinguish between

bacterial cells cultured for 48 h in smg conditions from those of the control setup. In accordance to the results for the 24h-old cultures, EMSC revealed higher levels of acetyl-CoA, citric acid and malic acid, and a lower level of succinic acid in SMG conditions compared to the control setup. In contrast to 24h-old cultures, PHB and HB appeared to be present in lower amounts in the SMG samples compared to the control samples. However, OD(590nm) values differed slightly, but significantly at the 95% confidential level, between smg and control cultures and thus cells sampled from both setups might not be in the same metabolic state.

10.4 Discussion

We have studied the behavior of *Cupriavidus metallidurans* LMG 1195 in RWV conditions. The aim of our study was to use Raman spectroscopy to discriminate biopatterns obtained in simulated microgravity (SMG) and the respective control condition. A second aim was to extract information about the biomolecules that could be responsible for the different biopatterns.

Raman spectra were collected for 24h- and 48h-old *C. metallidurans* cultures using independent cultures that were set up on different days. The PCA of the 24h-spectra did not reveal a significant difference between SMG and the control. This lack of differentiation is regarded as being caused by subtle differences in the inocula that were used. The day-to-day variations in the Raman spectra caused by the use of different inocula probably dominated the differences induced by simulated microgravity. This is confirmed by the observation of the same Raman bands in the difference spectra (smg minus control) that were calculated per measurement day (thus per inoculum). These difference spectra all resemble the difference spectrum in Figure 10.3a (mean SMG spectrum minus mean control spectrum). Clear bands could be identified for a relative increase in the amount of PHB and its monomer HB in SMG conditions (Fig. 10.3b,c). EMSC indicated a relative increase in the levels of acetyl-CoA, citric acid and malic acid in SMG and a decrease of succinic acid. These changes in the Krebs cycle compounds indicate that the carbon metabolism may be affected in low-shear simulated microgravity as mimicked in the RWV. A possible

explanation why the results for succinic acid are opposite to the other citric acid cycle compounds might be the fermentation of succinic acid to propionate as in *Propiogenium modestum*[16]. In this metabolic pathway, the decarboxylation of succinate allows sodium export, which causes ATP synthesis when re-imported via the Na^+-pump when sufficient Na is present. This kind of mechanism has so far not been reported in Gram-negatives although decarboxylases are well known in fermentative metabolism of Gram-negatives as well. Changes in bacterial metabolism caused by SMG were reported previously: Lam *et al.* 1998[20] reported a higher production of the antibiotic monorden by *Humicola fuscoatra* in the Space Shuttle compared to ground control samples, while Demain *et al.* 2000[7] observed changes in levels of several secondary metabolites induced by SMG. The change in PHB and HB also indicate an effect at the level of carbon metabolism. PHB is a storage polymer, which is formed when the carbon-nitrogen ratio is high[1,35]. The surplus of carbon is stored as PHB that can be used when the cell is in need of an additional carbon source or when the growth conditions become favourable. The observed increase in PHB after 24 h is in accordance with Thiruvenkatam and Scholz[37] who reported a higher PHB production (especially after 18 and 24 h of incubation) and no lag phase for *Azotobacter vinelandii UWD* (ATCC 53799) upon culturing in microgravity conditions using the RWV compared to a control setup consisting of shaked flasks with similar oxygen flow as in the RWV. The authors assumed that reduced gravity is a stress factor that stimulates the bacteria to produce storage polymers.

The spectra obtained after an incubation of 48 h, showed the same evolution for the Krebs cycle compounds. In contrast to the 24h-data, the levels of PHB and HB appeared to be lower in the SMG cultures compared to the control cultures. The reason for this change in function of the time is difficult to explain at present. Bacterial growth is a dynamic process controlled by many factors at the transcript level and metabolic level. It is plausible that the growth conditions between 24 h at 48 h of culture changed and triggered the cells in SMG to start consuming the PHB that was accumulated in the first 24 h. In addition, the OD(590nm) values after 48 h of culturing were slightly higher for the control setup than under microgravity conditions. Apparently growth is going faster in the control setup,

which might be caused by sinking of the bacterial mass towards the membrane (visually observed) which can cause a difference in oxygen exchange compared to the microgravity setup. Thus, the observation that there is more PHB and HB in the control setup than in the microgravity setup after 48 h, might not be solely caused by the microgravity effect itself. Interplay of the SMG environment, different oxygen demand and different growth characteristics may be responsible for the difference in Raman patterns. In this respect, Tavares *et al.*[36] found a more efficient PHB production of *Ralstonia eutropha* (*Cupriavidus necator*) when cultivated in an airlift bioreactor compared to a stirred tank bioreactor. These results are hard to compare to the ones presented here, because the growth conditions are completely different. The conclusions made by Tavares *et al.*[36] do not include microgravity effects, but state that the production of the same amount of PHB can be obtained with a lower oxygen demand in the airlift bioreactor compared to the stirred tank. Because other effects than microgravity might have interfered, no conclusions were drawn concerning the effect of low-shear microgravity on *Cupriavidus metallidurans* after 48 h of culturing.

In conclusion, Raman spectroscopy holds potential as a non-destructive method for studying whole-cell chemical composition of bacteria. It can be classified as a high content data methodology, which may generate biological information that can be used to uncover the bacterial response to a particular growth condition or treatment. Our results indicate that Raman spectroscopy could be used for discriminating biopatterns of *C. metallidurans* grown in simulated microgravity using the RWV and a control setup.

This case study shows that Raman spectroscopy can provide useful information about small differences in metabolism when culturing under different conditions.

References

1. Anderson A.J., Dawes E.A. (1990) Occurrence, metabolism, metabolic role, and industrial uses of bacterial polyhydroxyalkanoates. *Microbiol. Rev.* 54: 450-472.
2. De Boever P., Ilyin V.L., Forget-Hanus D., Van der Auwera G., Mahillon J., Mergeay M. (2008) Conjugation-mediated plasmid exchange between bacteria grown under space flight conditions. *Microgravity Sci. Technol.* 19: 138-144.
3. De Gelder J., De Gussem K., Vandenabeele P., De Vos P., Moens L. (2007) Methods for extracting biochemical information from bacterial Raman spectra: An explorative study on *Cupriavidus metallidurans* LMG 1195. *Anal. Chim. Acta* 585: 234-240.
4. De Gelder J., De Gussem K., Vandenabeele P., Moens L. (2007). Reference database of Raman spectra of biological molecules. *J. Raman spectrosc.* 38: 1138-1147.
5. De Gelder J., Scheldeman P., Leus K., Heyndrickx M., Vandenabeele P., Moens L., De Vos P. (2007) Raman spectroscopic study of bacterial endospores. *Anal. Bioanal. Chem.* 389: 2143-2151.
6. De Gelder J., Willemse-Erix D., Scholtes M.J., Sanchez J.I., Maquelin K., Vandenabeele P., De Boever P., Puppels G.J., Moens L., De Vos P. (2008) Monitoring poly(3-hydroxybutyrate) production in *Cupriavidus necator* DSM 428 (H16) with Raman spectroscopy. *Anal. Chem.* 80: 2155-2160.
7. Demain A.L., Fang A. (2001) Secondary metabolism in simulated microgravity. *The Chemical Record* 1: 333-346.
8. England L.S., Gorzelak M., Trevors J.T. (2003) Growth and Membrane Polarization in *Pseudomonas aeruginosa* UG2 Grown in Randomized Microgravity in a High Aspect Ratio Vessel. *Biochim. Biophys. Acta - Gen. Subj.* 1624: 76-81.
9. Fulget N., Poughon L., Richalet J., Lasseur Ch. (1999) MELISSA: Global control strategy of the artificial ecosystem by using first principles models of the compartments. *Adv. Space Res.* 24: 397-405.

10. Goa H., Ayyaswamy P. S., Ducheyne P. (1997). Dynamics of a microcarrier particle in the simulated microgravity environment of a rotating-wall vessel. *Microgravity Sci. Technol.* X/3 : 154-165.
11. Godia F., Albiol J., Montesinos J.L., Pérez J., Creus N., Cabello F., Mengual X., Montras A. Lasseur, Ch. (2002) MELISSA: a loop of interconnected bioreactors to develop life support in space. *J. Biotechnol.* 99: 319-330.
12. Hendrickx L., De Wever H., Hermans V., Mastroleo F., Morin N., Wilmotte A., Janssen P., Mergeay, M. (2006) Microbial ecology of the closed artificial ecosystem MELiSSA (Micro-Ecological Life Support System Alternative): Reinventing and compartmentalizing the Earth's food and oxygen regeneration system for long-haul space exploration missions. *Res. Microbiol.* 157: 77-86.
13. Hutsebaut D., Vandenabeele P., Moens L. (2005) Evaluation of an accurate calibration and spectral standardization procedure for Raman spectroscopy. *Analyst* 130: 1204-1214.
14. Hutsebaut D., Vandroemme J., Heyrman J., Dawyndt P., Vandenabeele P., Moens L., De Vos P. (2006) Raman microspectroscopy as an identification tool within the phylogenetically homogeneous '*Bacillus subtilis*'-group. *Syst. Appl. Microbiol.* 29: 650-660.
15. Ilyin V.K. (2005) Microbiological status of cosmonauts during orbital spaceflights on Salyut and Mir orbital stations. *Acta Astronaut.* 56: 839-850.
16. Janssen P.H., Liesack W., Kluge C., Seeliger S., Schink B., Harfoot C.G. (1996) Sodium-dependent succinate decarboxylation by a new anaerobic bacterium belonging to the genus *Peptostreptococcus*. *Antonie Van Leeuwenhoek* 70: 11-20.
17. Klaus D.M. (2001) Clinostats and bioreactors. *Gravit. Space Biol. Bull.* 14: 55-64.
18. Kiefer J., Pross H.D. (1999) Space radiation effects and microgravity. *Mutat. Res.-Fundam. Mol. Mech. Mutagen.* 430: 299-305.
19. Kirschner C., Maquelin K., Pina P., Thi N. A. N., Choo-Smith L. P., Sockalingum G. D., Sandt C., Ami D., Orsini F., Doglia S. M., Allouch P., Mainfait M., Puppels G. J., Naumann D. (2001) Classification and identification of enterococci: a

comparative phenotypic, genotypic, and vibrational spectroscopic study. *J. Clin. Microbiol.* 39: 1763-1770.

20. Lam K.S., Mamber S.W., Pack E.J., Forenza S., Fernandes P.B., Klaus D.M. (1998) The effects of space flight on the production of monorden by *Humicola fuscatra* WC5157 in solid-state fermentation. *Appl. Microbiol. Technol.* 49: 579-583.
21. Leys N.M.E.J., Hendrickx L., De Boever P., Baatout S., Mergeay M. (2004) Space flight effects on bacterial physiology. *J. Biol. Regul. Homeost. Agents* 18: 193-199.
22. Lynch S.V., Mukundakrishnan K., Benoit M.R., Ayyaswamy P.S., Matin A. (2006) *Escherichia coli* biofilms formed under low-shear modeled microgravity in a ground-based system. *Appl. Environ. Microbiol.* 72: 7701-7710.
23. Lui T. Q., Li X. Q., Sun X. Y., Ma X. H., Cui Z. F. (2004) Analysis on forces and movement of cultivated particles in a rotating wall vessel bioreactor. *Biochem. Eng. J.* 18: 97-104.
24. Maquelin K., Dijkshoorn L., van der Reijden T. J. K., Puppels G. J. (2006) Rapid epidemiological analysis of *Acinetobacter* strains by Raman spectroscopy. *J. Microbiol. Meth.* 64: 126-131.
25. Maquelin K., Krischner C., Choo-Smith L.P., van den Braak N., Endtz H.P., Naumann D., Puppels G.J. (2002) Identification of medically relevant microorganisms by vibrational spectroscopy. *J. Microbiol. Meth.* 51: 255-271.
26. Martens H., Stark E. (1991) Extended multiplicative signal correction and spectral interference subtraction: new preprocessing methods for near infrared spectroscopy. *J. Pharmaceut. Biomed. Anal.* 9: 625-635.
27. Mergeay M. (2006) Editorial. *Res. Microbiol.* 157: 1-4.
28. Nauman E.A., Ott C.M., Sander E., Tucker D.L., Pierson D., Wilson J.W., Nickerson C.A. (2007) Novel quantitative biosystem for modeling physiological fluid shear stress on cells. *Appl. Environm. Microbiol.* 73: 699-705.
29. Naumann D., Keller S., Helm D., Schultz C., Schrader B. (1995) FT-IR spectroscopy and FT-Raman spectroscopy are powerful analytical tools for the non-invasive characterization of intact microbial cells. *J. Mol. Struct.* 347: 399-406.

30. Nickerson C.A., Ott C.M., Wilson J.W., Ramamurthy R., LeBlanc C.L., Bentrup K.H.Z., Hammond T., Pierson D.L. (2003) Low-shear modeled microgravity: a global environmental regulatory signal affecting bacterial gene expression, physiology, and pathogenesis. *J. Microbiol. Meth.* 54: 1-11.
31. Nickerson C. A., Ott C. M., Wilson J. W., Ramamurthy R., Pierson D. J. (2004). Microbial responses to microgravity and other low-shear environments. *Microbiol. Mol. Biol. R.* 68: 345-361.
32. Novikova N., De Boever P., Poddubko S., Deshevaya E., Polikarpov N., Rakova N., Coninx I., Mergeay M. (2006) Survey of environmental biocontamination on board the International Space Station. *Res Microbiol.* 157: 5-12.
33. Oust A., Moretro T., Naterstad K., Sockalingum G. D., Adt I., Manfait M., Kohler A. (2006) Fourier transform infrared and Raman spectroscopy for characterization of *Listeria monocytogenes* strains. *Appl. Environ. Microb.* 72: 228-232.
34. Pross H.D., Casares A., Kiefer J. (2000) Induction and repair of DNA double-strand breaks under irradiation and microgravity. *Radiat. Res.* 153: 521-525.
35. Shi H., Shiraishi M., Shimizu K. (1997) Metabolic flux analysis for biosynthesis of poly(beta-hydroxybutyric acid) in *Alcaligenes eutrophus* from various carbon sources *J. Ferment. Bioeng.* 84: 579-587.
36. Tavares L.Z., da Silva E.S., Pradella J.G.D. (2004) Production of poly(3-hydroxybutyrate) in an airlift bioreactor by *Ralstonia eutropha*. *Biochem. Eng.* J. 18: 21-31.
37. Thiruvenkatam R., Scholz C. (2000) Synthesis of poly(β-hydroxybutyrate) in simulated microgravity: an investigation of aeration profiles in shake flask and bioreactor. *J. Polym. Environ.* 8: 155-159.
38. Tucker D.L., Ott C.M., Huff S., Fofanov V., Willson R.C., Fox G.E. (2007) Characterization of *Escherichia coli* MG1655 grown in a low shear modeled microgravity environment. *BMC Microbiology* 7: Art. No. 15.
39. Wilson J.W., Ott C.M., Bentrup K.H., Ramamurthy R., Quick L., Porwollik S., Cheng P., McClelland M., Tsaprailis G., Radabaugh T., Hunt A., Fernandez D., Richter E., Shah M., Kilcoyne M., Joshi L., Neiman-Gonzalez M., Hing S., Parra

M., Dumars P., Norwood K., Bober R., Devich J., Ruggles A., Goulart C., Rupert M., Stodieck L., Stafford P., Catella L., Schurr M.J., Buchanan K., Morici L., McCracken J., Allen P., Baker-Coleman C., Hammond T., Vogel J., Nelson R., Pierson D. L., Stefanyshyn-Piper H.M., Nickerson C.A. (2007) Space flight alters bacterial gene expression and virulence and reveals a role for global regulator Hfq. *Proc. Natl. Acad. Sci. U. S. A.* 104: 16299-16304.

Chapter 11: Conclusions and future prospects

The Raman effect relies on molecular vibrations and therefore Raman spectra provide information about a samples' composition. More specific, the presence of certain molecules in samples can be deduced from their Raman active vibrations. Raman spectra of bacterial cells thus comprise signals from all Raman active vibrations of the cell's compounds. Therefore, on the one hand, Raman spectra of bacterial cells are rich in information, but on the other hand they are complex.

When using Raman spectroscopy as a tool for studying bacterial cell compounds, variations in cell composition within and between cultures should be considered. In this work, for each experiment, several independent cultures were prepared, sampled and measured on different days in order to obtain a reliable Raman fingerprint of a strain cultivated in certain conditions. In qualitative Raman spectroscopy, very often differences between two or more strains and/or cultivation conditions are compared. Before proceeding to the interpretation of the bacterial spectra, they should first be subjected to PCA in order to assess whether the within- and between-culture variation is smaller than the differences caused by the variety of the studied strains and/or cultivation conditions.

The basic information needed for the interpretation of bacterial Raman spectra are the Raman signals caused by specific cell compounds. Therefore, a database of reference Raman spectra was constructed from biomolecules (or their building blocks) which were commercially available. The information comprised in these reference spectra can be used in several ways to extract information from bacterial Raman spectra (summarized in Table 11.1). The proper method(s) to do so depend(s) on the position (isolated band or band overlap) and intensity of the Raman signal from the studied biomolecule(s).

When studying compounds that contribute highly to a bacterial Raman spectrum, information can be retrieved by focussing on one or a few characteristic Raman band(s). The presence of these compounds in the cells can be simply deduced by comparison of their reference spectra and the bacterial spectra. Difference spectra and PCA reveal differences in cell concentration of compounds between different strains or cultivation

conditions. However, many bands in a Raman spectrum contain contributions of several biomolecules. To study these compounds, more enhanced approaches are necessary. A dot product between a reference spectrum and a bacterial spectrum reflects the contribution of the reference spectrum from the studied biomolecule to the bacterial spectrum. These dot products can be compared per biomolecule for different bacterial spectra to determine which samples show a higher or lower content of that biomolecule. However, these values are not appropriate for comparison between different biomolecules, as they do not take into account the intrinsic Raman activity of the biomolecules. The advantage of dot products lies in the consideration of the whole reference spectrum of the studied biomolecule, instead of focussing on a single or a few characteristic bands. These dot products reflect the contribution of a reference spectrum that can not be distinguished from contributions of signals from other biomolecules. Dot products can thus overestimate the relative concentration of a biomolecule in a sample. To address this problem, the EMSC model, which was previously used for pre-processing, was extended by adding reference spectra. The coefficients resulting from this model reveal similar information as the dot products, but should be more reliable for distinguishing contributions of several biomolecules. Indeed, in the EMSC procedure, the reference spectra are fitted together towards the bacterial spectrum and so overfitting is minimized. To overcome colinearity problems when adding very similar reference spectra (of structurally similar biomolecules) to the EMSC model, PCA can be performed on these reference spectra and some of the resulting loading vectors can be used as input in EMSC. The resulting EMSC coefficients can be related to the original biomolecules by studying the scores of the PCA. In this way, information is retrieved from groups of biomolecules instead of from each separate biomolecule.

The above mentioned methods allow to compare the relative concentration of a particular biomolecule in different bacterial samples. In this way, samples originating from different strains, metabolic states or culturing conditions can be studied for their differences. However, it would be interesting to be able to determine absolute concentrations of a given compounds in a cell. The potential to do so was shown by the linear relation between the intensity of an intense characteristic band of a PHB and the PHB concentration (mg/L

culture) as determined by HPLC. This approach was possible because we focussed on a PHB band that was completely isolated in the bacterial spectrum and appropriate pre-processing was applied. When band overlap occurs, other processing methods should be applied to allow quantitative determination of compounds. For example the use of PLS models could provide quantitative data when calibration with standards is possible. Problems with the implementation of this approach could be solubility in water of some compounds and the cost to prepare solutions of sufficient concentration of compounds that are expensive in purified form. These standards were not necessary for the application of EMSC as this procedure provides relative values instead of absolute concentrations.

Table 11.1 Overview of the applied methods for studying bacterial cell compounds with Raman spectroscopy

approach	qualitative	quantitative	applicability
based on separate Raman bands	Difference spectra, PCA (Ch6)	Band intensity (Ch9)	When a biomolecule shows an intense and isolated band in bacterial Raman spectra.
based on a whole reference spectrum	Dot product (Ch7)		When a biomolecule shows a minor and/or not isolated contribution in bacterial Raman spectra.
based on multiple reference spectra	EMSC (Ch7,8)		

This work has shown that Raman spectroscopy is a potential powerful tool for microbial analysis. On the one hand, this technique allows in vivo detection, monitoring or (semi-)quantification of a specific cell compound (illustrated in this work for CaDPA and PHB). On the other hand, bacterial Raman spectra provide whole cell profiles and are thus suitable for fast detection of differences in general cell constitution between various samples. In this way, the influence of specific culturing conditions on the cell's composition or metabolism can be tracked, provided the use of a standardized protocol that excludes other variations. The requirement of only small sample volumes and the minimal sample preparation are attractive characteristics in the microbial laboratory. The non-destructive nature of Raman spectroscopy allows the construction of in situ monitoring systems. Recent developments of instrumentation allowing single-cell measurements even broaden the field of application. Some important practical constraints of this technique are: (i) the

studied compounds have to be Raman-active, (ii) fluorescence may mask the spectrum, (iii) slimy bacteria need extra washing steps in the sample preparation, in order to obtain high-quality spectra. As Raman spectra include information about the amount of specific (building blocks of) biomolecules present in the cell, no sequence information of any kind can be extracted from them. Therefore, Raman spectroscopy should be used as a supplementary technique in the microbial lab or as a first whole cell screening tool instead of a substitute for existing microbial techniques of analysis.

At the end of Chapter 5 and 7 two important remarks on the presented approaches were discussed, which each lead to ideas for future research.

Firstly, reference spectra are estimations of the biomolecules' signal in a bacterial Raman spectrum. These estimations can be improved in such a way that the output of the data processing, such as dot product calculations or EMSC, becomes more accurate. In theory, this could be achieved by collecting Raman spectra from these biomolecules in an environment resembling the bacterial cell. However, it is impossible to imitate a cell's composition completely and by trying so the collected Raman spectrum would contain interfering effects from the other constituents. Perhaps, in future, quantum chemistry could help to predict Raman signals of biomolecules in a cell environment, as this field is developing fast.

Secondly, one can imagine that the amount of reference spectra added to the EMSC model can affect the obtained results. It is obvious that the set of reference spectra should include all biomolecules with the most intense contributions in a bacterial Raman spectrum. However, one can wonder whether the set can contain too many of such reference spectra or whether the presence of reference spectra that do not contribute to the bacterial spectrum affects the accuracy of the model. Future research may focus on the determination of fitting errors on artificial data and fitting errors on experimental data by comparison to data obtained from other techniques of analysis. However, it is hard to find appropriate techniques for such a comparison as these should provide quantitative data of (building blocks of) biomolecules that are detected by Raman spectroscopy.

In conclusion, Raman spectroscopy is an interesting non-destructive tool for studying bacterial cell compounds. According to the contribution of the studied biomolecule in the bacterial spectrum, different processing methods can/should be applied. Although till present, mainly relative values for the concentrations of biomolecules can be obtained and compared for different strains or cultivation conditions, the technique has shown its potential for quantitative analysis of bacterial cell compounds. Further research should focus on quantitative determination and the accuracy of the used models. In the microbial laboratory, it is an interesting technique supplementary to others and it allows fast whole cell profiling.

Summary / Samenvatting

Summary

Raman spectroscopy is an attractive tool for microbial analysis because of the very small sample volumes needed for analysis, the minor sample preparation and the speed of analysis. Moreover, Raman spectra provide information of all Raman active molecules in the bacterial cell. This technique has already proven to be successful for bacterial identification at the species and strain level. For this purpose, Raman spectroscopy is used as a fingerprint technique. However, these spectra also contain valid information about biochemical composition of the cells. Because bacterial Raman spectra are the sum of signals from all Raman active cell compounds, they are complex. Although in literature some bands have been assigned to (specific or groups of) biomolecules, only very few studies are published where Raman spectroscopy is used to study specific bacterial cell compounds. Therefore, the aim of this work was to develop methods for extracting information from these complex bacterial spectra.

The foundation of this work was laid by constructing a database that contains reference Raman spectra of biomolecules that are known to constitute bacterial cells (Chapter 5). These reference spectra were further used throughout this work to extract information about biomolecules from bacterial Raman spectra.

In Chapter 6, sporulating bacteria and endospores obtained from several strains or after different cultivation conditions were analysed. Because the studied compounds show intense bands in the bacterial spectra, information can be obtained by: (i) comparison of reference spectra and bacterial spectra, (ii) comparison of reference spectra and difference spectra obtained from spectra originating from different strains or cultivation conditions, (iii) comparison of the score and loading plots from PCA. In this way, sporulation of *Bacillus licheniformis* LMG 7634 could be followed as a function of incubation time by the characteristic band at 1018 cm^{-1} of calcium dipicolinate (CaDPA). Spores obtained from various strains cultured under different conditions, were found to differ in the concentration of various compounds such as CaDPA, cysteine, phenylalanine and tyrosine.

Although the approaches presented in Chapter 6 are very useful to study bacterial cell compounds, they are less accurate when the signals of different biomolecules overlap in the bacterial spectra.

To tackle the problem of compounds that show band overlap in the bacterial spectra, more enhanced approaches are necessary to extract information on individual components. Two approaches are illustrated in Chapter 7: (i) the calculation of dot products between reference spectra and bacterial spectra and (ii) the application of an EMSC (extended multiplicative signal correction) procedure including reference spectra of biomolecules in the model. These approaches have an advantage over the ones presented in Chapter 6: they take the whole reference spectrum into account to evaluate the presence of a compound in a sample. Calculation of dot products is easy, but has the drawback that more overfitting can occur compared to the EMSC model, because the last one fits all considered reference spectra simultaneously to the bacterial Raman spectra. A third approach, illustrated in Chapter 7, is two dimensional correlation spectroscopy, which reveals correlating bands in a perturbed dataset. These three techniques were applied in this chapter on Raman spectra of *Cupriavidus metallidurans* LMG 1195 in different metabolic states.

In literature, the developers of the EMSC procedure for pre-processing warn for possible colinearity problems which can occur for example when similar reference spectra (thus spectra from structurally similar biomolecules) are added to the model. To avoid this problem, PCA can be applied to these reference spectra and some of the resulting loading vectors (orthogonal) can be added as reference spectra to the EMSC model, as illustrated in Chapter 8. After applying the EMSC model as described in Chapter 7, information about the contribution of the original reference spectra could be obtained by studying the PCA scores. In this way, information extraction is possible, although the retrieved information is less detailed, as the drawn conclusions concern groups of biomolecules instead of single biomolecules. This approach was applied to retrieve information about the fatty acid composition of bacteria. The results were in accordance to parallel FAME analyses.

In Chapter 9, the potential of Raman spectroscopy for quantitative determination of bacterial cell compounds was explored. For this purpose poly-3-hydroxybutyrate (PHB), a storage compound, was chosen which shows an intense isolated band at 1734 cm^{-1} in bacterial Raman spectra. When applying appropriate pre-processing, the intensity of this band showed a linear relation with the PHB concentration (mg/L culture) in *Cupriavidus necator* DSM 428 retrieved from HLPC analysis.

Chapter 10 demonstrates a case study of the influence of microgravity on *Cupriavidus metallidurans* LMG 1195. Methods that were discussed in Chapter 6 and 7 (difference spectra, EMSC procedure) were applied to search for differences between *Cupriavidus metallidurans* LMG 1195 cultured in a RWV (rotating wall vessel) simulating low-shear microgravity and a similar control setup. Differences in PHB production and in general carbon metabolism could be identified.

In conclusion, Raman spectroscopy is a powerful tool for studying bacterial cell compounds. Different processing methods should be applied according to the Raman signal of the studied biomolecules in bacterial spectra. These methods provide information on the relative concentration of compounds in bacterial cells. As these approaches do not take into account the intrinsic Raman activity of the compounds, they should only be compared per compound, not between different compounds. The potential of Raman spectroscopy for quantitative analysis of bacterial cell compounds has been shown by the study of PHB production in *C. necator*. Raman spectroscopy could be an interesting tool in the microbial lab for fast whole cell profiling or for studying specific cell compounds in various cultivation conditions. Further research should focus on: (i) the amelioration of reference spectra by taking cell conditions into account, (ii) how the EMSC model's accuracy is influenced by using different sets of reference spectra and (iii) calibration of PLS models for quantitative determination of compounds.

Samenvatting

Ramanspectroscopie is aantrekkelijk voor microbiële analyse omdat slechts kleine staal volumes vereist zijn, de staalvoorbereiding minimaal is en de analyse snel is. Bovendien bevatten Ramanspectra informatie over alle Raman-actieve componenten in de bacteriële cel. Deze techniek werd in het verleden succesvol ingezet voor de identificatie van bacteriën tot op het species- en stamniveau. Hierbij worden Ramanspectra enkel als een fingerprintpatroon gebruikt. Nochtans bevatten deze spectra ook waardevolle informatie over de biochemische samenstelling van de cel. Bacteriële Ramanspectra zijn complex, aangezien ze de som zijn van de signalen van alle Raman-actieve bestanddelen in de cel. Alhoewel in de literatuur reeds enkele banden werden toegekend aan (specifieke of groepen van) biomoleculen, wordt er weinig gerapporteerd over het gebruik van Ramanspectroscopie voor de studie van specifieke bacteriële componenten. Het doel van dit werk is dan ook om methoden te ontwikkelen die informatie uit complexe bacteriële spectra extraheren.

De basis van dit werk is de aanleg van een databank van referentiespectra van biomoleculen waarvan de aanwezigheid in bacteriële cellen gekend is (Hoofdstuk 5). Deze referentiespectra werden doorheen dit werk gebruikt voor de extractie van informatie over biomoleculen uit bacteriële Ramanspectra.

In Hoofdstuk 6 werden sporulerende bacteriën en endosporen, geïsolleerd uit verschillende stammen en gecultiveerd onder verschillende condities, geanalyseerd. Omdat de bestudeerde celcomponenten intense banden vertonen in bacteriële spectra, kan informatie bekomen worden via: (i) vergelijking van referentiespectra en bacteriële spectra, (ii) vergelijking van referentiespectra en verschilspectra tussen spectra gecollecteerd van verschillende stammen of verschillende cultivatie omstandigheden, (iii) vergelijking van de score en loading grafieken bekomen uit PCA. Op deze wijze werd de sporulatie van *Bacillus licheniformis* LMG 7634 gevolgd in functie van de incubatietijd aan de hand van de karakteristieke band bij 1018 cm^{-1} veroorzaakt door calciumdipicolinaat (CaDPA). Sporen,

geïsoleerd uit verschillende stammen en gecultiveerd onder verscheidene condities, bleken te verschillen in de concentratie aan componenten zoals CaDPA, cysteine, fenylalanine en tyrosine. Alhoewel de benaderingen voorgesteld in Hoofdstuk 6 erg geschikt zijn voor de studie van bacteriële celcomponenten, zijn ze minder accuraat wanneer de Ramansignalen van verschillende biomoleculen overlappen in de bacteriële spectra.

Om het probleem van overlappende banden in bacteriële spectra aan te pakken, zijn meer gevorderde technieken nodig zodat informatie verkregen kan worden over individuele celcomponenten. Twee mogelijke benaderingen worden geïllustreerd in Hoofdstuk 7: (i) de berekening van inproducten tussen referentiespectra en bacteriële spectra en (ii) het toepassen van een EMSC (extended multiplicative signal correction) procedure die referentiespectra van biomoleculen opneemt in het model. Deze benaderingen hebben een belangrijk voordeel tegenover diegene in Hoofdstuk 6: ze houden rekening met het volledige referentiespectrum om de aanwezigheid van een component in een staal te evalueren. Het berekenen van inproducten is eenvoudig, maar heeft tegenover EMSC het nadeel dat er meer gevaar is voor overfitting. De EMSC procedure fit alle beschouwde referentiespectra simultaan aan de bacteriële spectra. Een derde benadering die in Hoofdstuk 7 geïllustreerd wordt, is twee-dimensionele correlatie spectroscopie die gecorreleerde banden aangeeft in een geperturbeerde dataset. Deze drie technieken werden in dit hoofdstuk toegepast op Ramanspectra van *Cupriavidus metallidurans* LMG 1195 gecollecteerd in verschillende metabolische toestanden.

In de literatuur waarschuwen de ontwikkelaars van de EMSC procedure voor mogelijke colineariteitsproblemen die bijvoorbeeld kunnen optreden wanneer gelijkaardige referentiespectra (spectra van structureel gelijkaardige biomoleculen) in het model worden opgenomen. Om dit probleem te vermijden, wordt in Hoofdstuk 8 de toepassing van PCA vòòr EMSC geïllustreerd. Hierbij worden enkele loadingvectoren (orthogonaal) uit de PCA gebruikt als referentiespectra in het EMSC model. Nadien kan informatie bekomen worden over de bijdrage van de originele referentiespectra via de scores van de PCA. De bekomen informatie is minder gedetailleerd, aangezien ze groepen van biomoleculen in plaats van

specifieke biomoleculen betreft. Deze benadering werd toegepast om informatie te bekomen over de vetzuursamenstelling van bacteriën. De resultaten waren in overeenkomst met deze van parallelle FAME analyses.

In Hoofdstuk 9 werden de mogelijkheden van Ramanspectroscopie voor de kwantitatieve bepaling van bacteriële celcomponenten verkend. Hiervoor werd poly-3-hydroxybutyraat (PHB) gekozen, een component voor koolstof-opslag. Deze vertoont een instense en geïsoleerde band bij 1734 cm^{-1} in bacteriële Ramanspectra. Mits het toepassen van geschikte pre-processing, werd een linear verband gevonden tussen de intensiteit van deze band en de PHB concentratie (mg/L cultuur) in *Cupriavidus necator* DSM 428 bekomen door HPLC analyse.

Hoofdstuk 10 beschrijft een case study van de invloed van microgravitéit op *Cupriavidus metallidurans* LMG 1195. Methodes voorgesteld in de Hoofdstukken 6 en 7 (verschilspectra en EMSC) werden toepepast om verschillen te detecteren tussen *Cupriavidus metallidurans* LMG 1195, gecultiveerd in een RWV (rotating wall vessel) die lage-wrijvingsmicrograviteit simuleert en in een gelijkaardige controle opstelling. Verschillen in PHB-productie en meer algemeen in koolstofmetabolisme werden geïdentificeerd.

Als besluit kunnen we stellen dat Ramanspectroscopie een krachtig hulpmiddel is voor de analyse van bacteriële celcomponenten. Verschillende verwerkingsmethoden kunnen toegepast worden naargelang het Ramansignaal van de bestudeerde componenten in bacteriële spectra. Via deze methoden wordt informatie over de relatieve concentratie van componenten in bacteriele cellen bekomen. Omdat deze benaderingen geen rekening houden met de instrinsieke Raman-activiteit van de componenten, mogen de bekomen waarden enkel met elkaar vergeleken worden per component en niet tussen verschillende componenten. De mogelijkheden van Ramanspectroscopie voor kwantitatieve analyse van bacteriële celcomponenten werd geïllustreerd door de studie van PHB-productie in *Cupriavidus necator* DSM 428. Ramanspectroscopie kan nuttig zijn in microbiële

laboratoria voor een snelle screening van bacteriële cellen (*whole cell profiling*) en voor de studie van specifieke celcomponenten in verscheidene cultivatie-condities. Verder onderzoek kan zich toespitsen op: (i) het verbeteren van referentiespectra door celcondities in rekening te brengen, (ii) de invloed van het gebruik van verschillende sets referentiespectra op de accuratesse van EMSC modellen en (iii) de calibratie van PLS modellen voor kwantitatieve analyse van componenten.

Dankwoord

Het dankwoord is allicht het meest gelezen onderdeel van een doctoraat. Soms lijkt het zo'n verplicht nummertje, maar ik voel dit niet zo aan. Alles hieronder is van harte.

Uiteraard wil ik beginnen met het bedanken van onze vakgroepvoorzitter Prof. Strijckmans en mijn promotoren-team, de directe ondersteuning van mijn werk. Prof. Moens, bedankt om dit werk te promoten. Ondanks het feit dat het niet mogelijk was dat wij elkaar vaak spraken, heb ik een goed gevoel bij onze samenwerking. Je bleef altijd geïnteresseerd in mijn werk. Als er een artikel werd ingediend of aanvaard kreeg ik steevast een mail met felicitaties en dank voor het gedane werk. Dit was echt een hart onder de riem voor mij!

Peter, jij kreeg de taak om mij binnen de vakgroep analytische chemie in de praktijk te ondersteunen. Alhoewel we beiden werkten met Ramanspectroscopie lagen onze toepassingsgebieden ver uit elkaar. Geen gemakkelijke taak, maar ik denk dat we er het beste van gemaakt hebben. Hoe meer het doctoraat vorderde, hoe meer ik èn ook jij thuis werden in de wereld van dataverwerking. Ik wil je bedanken voor je wetenschappelijke ondersteuning, ook al was dit niet altijd makkelijk. Ik ben je ook dankbaar dat je minstens 10 keer in mijn bureau bent komen argumenteren waarom een artikel over onze databank nut had en dat velen hierin zouden geïnteresseerd zijn. Zonder zou jouw aandringen was dit nooit iets geworden.

Paul, jij stond in voor mijn ondersteuning op microbiologisch vlak. In het begin was ik ook erg afhankelijk van jou om mijn eerste experimenten uit te denken. Bedankt dat je me toen op de rails gezet hebt. Ik weet nog hoe onze eerste gesprekken eraan toe gingen. Nadat ik mijn prille ideeën had uitgelegd, zei je: "Hoho, dat is allemaal heel schoon op papier, maar die beestjes doen niet altijd wat jij wilt hoor". Mijn analytische geest kreeg direct schock-therapie, precies wat ik nodig had om er iets van te maken.

Naast mijn promotoren-team zijn er natuurlijk nog vele mensen waarmee ik samengewerkt heb, vooreerst binnen het labo Analytische Chemie.

Didier, met jou is alles begonnen. Ik wou doctoreren en zocht een onderwerp. “Wil je niets met bacteriën doen?” Hmmm ja, het leek me onmiddellijk interessant om twee velden te combineren. Ook al was je toen zelf aan het schrijven, ging je de eerste keer mee naar microbiologie lab om me te leren enten. Een tip die je vaak herhaalde was: “Je moet goed KIJKEN naar je data”. Dat leek erg banaal, maar met de jaren begreep ik meer en meer wat je bedoelde met KIJKEN. Je enthousiasme was erg aanstekelijk en ik vond het dan ook jammer dat je niet kon blijven.

Kris, bedankt voor de hulp met MATLAB. In het begin was het al rood en error wat de klok sloeg. “Kriiiiiiis?!” Ik heb veel bijgeleerd en ben ondertussen het MATLAB-monster de baas.

Niko, bedankt voor die talloze calibraties. Zonder jou had ik veel minder werk kunnen verzetten. Terwijl ik op de Ledeganck mijn stalen voorbereide, bracht jij het toestel in orde, ik kon op je rekenen.

Marleen, Sylvia, Isabel, Annelien, over wat heb ik allemaal niet tegen jullie gezaagd! En wat luchtte dat altijd op! Bedankt voor de steun.

Bij een doctoraat horen ook computerproblemen, papierwerk en andere probleempjes… en daarom wil ik ook Bart, Rudy, Jacques, Chantal, Lucien, Roger, Pieter en Marc bedanken.

Toen ik binnenkwam op het labo Microbiologie werd mij iemand toegwezen voor al mijn praktische vragen, en dat waren er heel wat. Liesbeth, erg bedankt voor de ondersteuning in een wereld die voor mij geheel nieuw was. Ik herinner me nog je gezicht toen ik vroeg waar de analytische balans was om mijn TSA medium af te wegen… Jeroen en Rieka, jullie waren fijne bureau genoten in de eerste periode van mijn doctoraat. Kim, ik had al op de baksteen van je huisje geschreven dat ik nooit zal vergeten wat je zei in de flow: “Joke, alles komt goe, gewoon omdat het moet”. En inderdaad, het is goedgekomen… bedankt voor alle goede raad en leuke babbels. Verder wil ik alle anderen van het “klein labo” bedanken: Caroline, An, Bram, Jeroen A, Emly, Ines; maar daarbuiten ook Margo, Paul S, Peter D, Marc V en Cindy voor hun hulp.

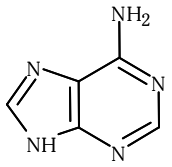
Voor enkele deelprojecten werd samengewerkt met andere instellingen. Patrick, Kees, Diana, Maarten, Gerwin, Nacho, Marc, Patsy, bedankt voor de samenwerkingen die stuk voor stuk deel uit maken van dit werk. Kees, ondanks de erg lastige dagen en het broodje kroket dat op mijn maag lag, ben ik blij dat ik heb kunnen kenismaken met jullie labo.

Eva, Sofie, tegen jullie kon ik mijn hart luchten... over het defecte toestel, de vervelende kweek-uren, weer het defecte toestel, etc. Bedankt voor jullie luisterend oor, mijn verhalen waren allicht supersaai. Verder wil ik ook mijn familie bedanken voor alle steun: mijn ouders, zus en grootouders. Leen, jij maakte het eerste deel van mijn doctoraat mee vanuit Idaho. Onze telefoontjes en chat-sessies betekenden veel voor mij want als ik het niet goed meer wist, zou mijn zus wel raad weten. Opa G., wat jammer dat je dit niet meer kan meemaken. Ik weet dat je dit werk met evenveel interesse zou doorbladerd hebben als mijn licentiaatsthesis en dat je met evenveel trots op mijn verdediging zou hebben gestaan als op mijn proclamaties.

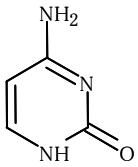
Olivier, ik kan hier onmogelijk vermelden wat je voor mij allemaal hebt gedaan. Het zou ook belachelijk zijn om dat te proberen. Het is dat onbeschrijfbare gevoel van iemand die achter je staat. Bedankt!

Appendix: Chemical structures

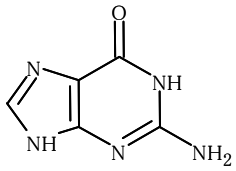
DNA/RNA bases



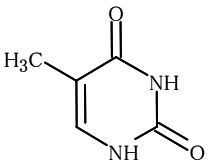
adenine



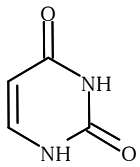
cytosine



guanine

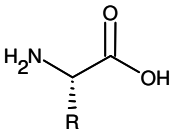


thymine



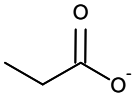
uracil

Amino acids

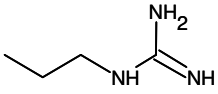


With R =

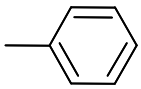
H	glycine
CH_3	L–alanine
iPr	L–valine
CH_2OH	L–serine
CH_2SH	L–cysteine



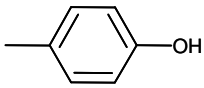
L–glutamate



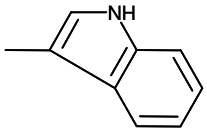
L–arginine



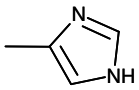
L–phenylalanine



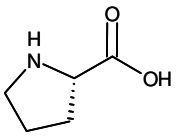
L–tyrosine



L–tryptophan

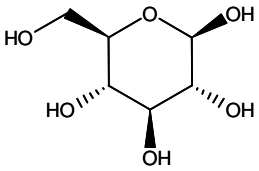


L–histidine

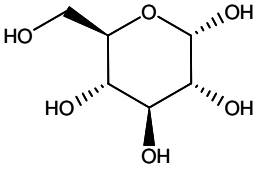


proline

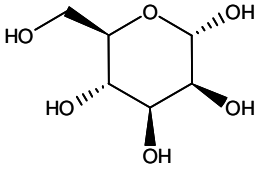
Saccharides



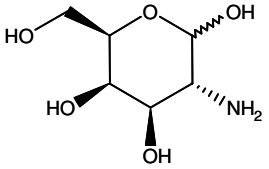
β-D-glucose



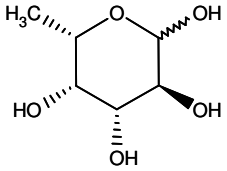
D(+)-dextrose



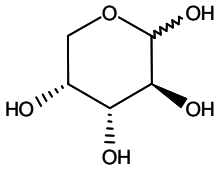
D(+)-mannose



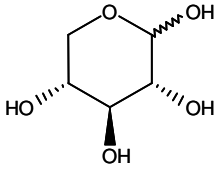
D(+)-galactosamine



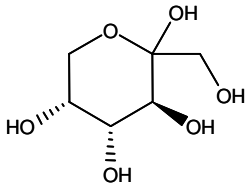
D(+)-fucose



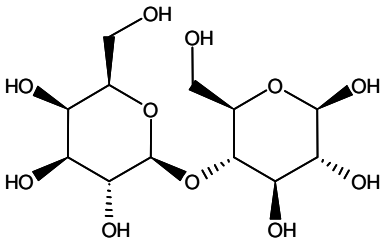
D(-)-arabinose



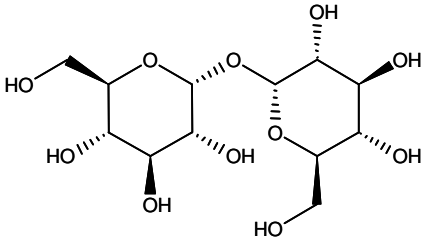
D(+)-xylose



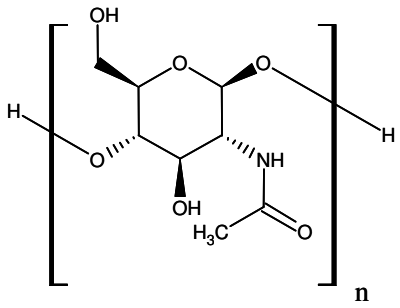
D(-)-fructose



lactose

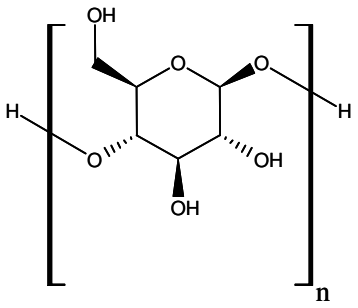


D(+)-trehalose



n = 1: N-acetyl-D-glucosamine

n >> 1: chitin



α-D-Glucose:

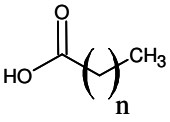
linear polymer: amylose

branched polymer: amylopectin

β-D-Glucose:

linear polymer: cellulose

Fatty acids and fats

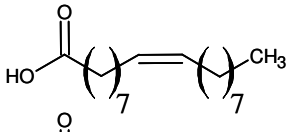


n = 12: lauric acid

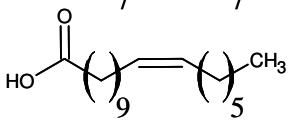
n = 14: myristic acid

n = 16: palmitic acid

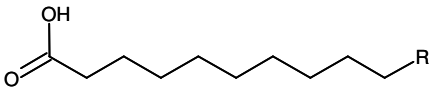
n = 18: stearic acid



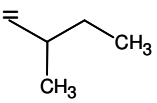
oleic acid



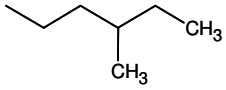
vaccenic acid



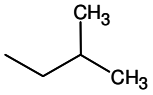
With R =



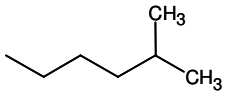
12–methyltetradecanoic acid (15Aiso)



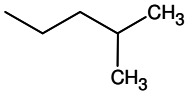
14–methylhexadecanoic acid (17Aiso)



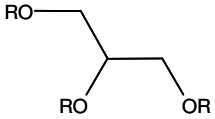
13–methylmyristic acid (15iso)



15–methylpalmitic acid (17iso)



14–methylpentadecanoic acid (16iso)



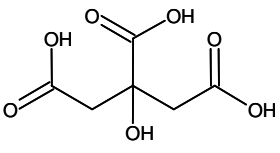
R = H: glycerol

-(C=O)-$(CH_2)_7$-CH=CH-$(CH_2)_7$-CH_3: triolein

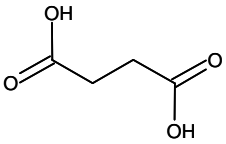
-(C=O)-$(CH_2)_7$-CH=CH-CH_2-CH=CH-$(CH_2)_4$-CH_3: trilinolein

-(C=O)-$(CH_2)_7$-CH=CH-CH_2-CH=CH-CH_2-CH=CH-CH_2-CH_3: trilinolenin

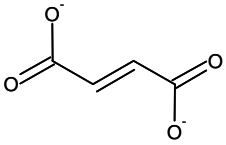
Metabolites



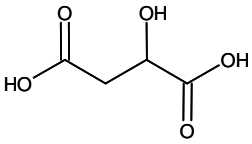
citric acid



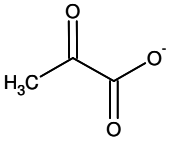
succinic acid



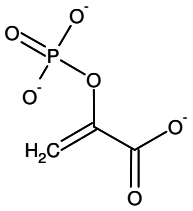
fumarate



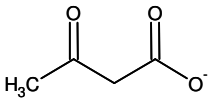
malic acid



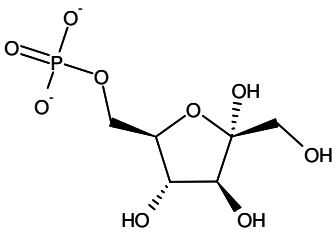
pyruvate



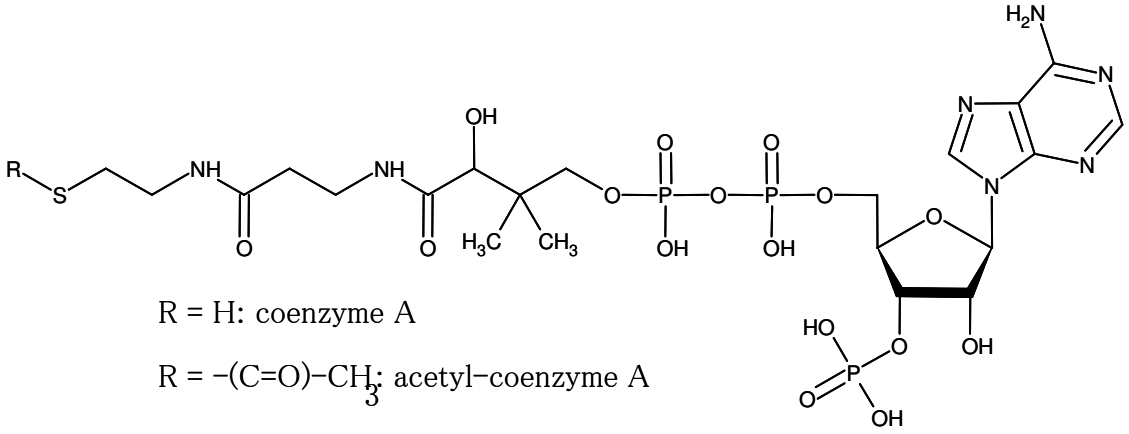
phosphoenolpyruvate



acetoacetate



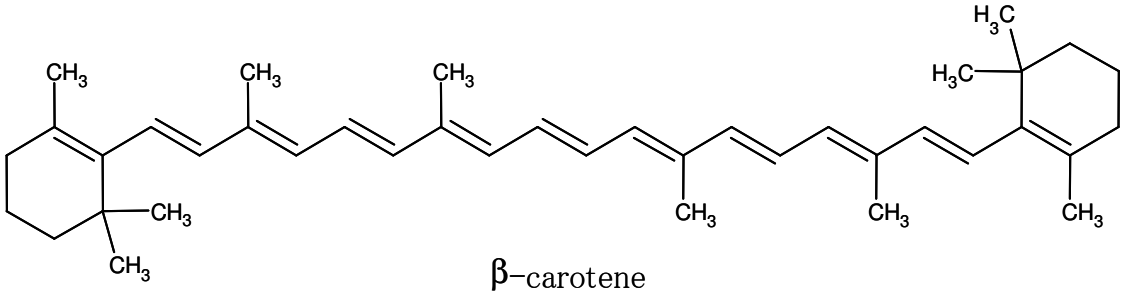
D–fructose–6–phosphate



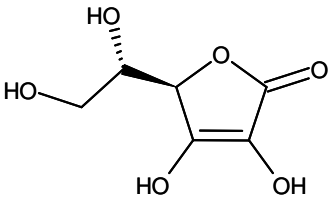
R = H: coenzyme A

R = −(C=O)−CH_3: acetyl−coenzyme A

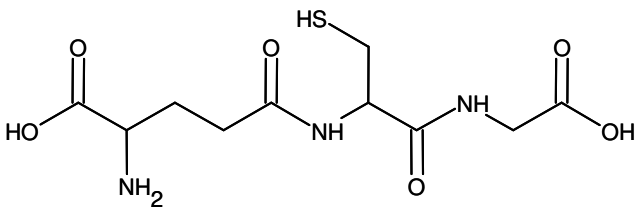
Others



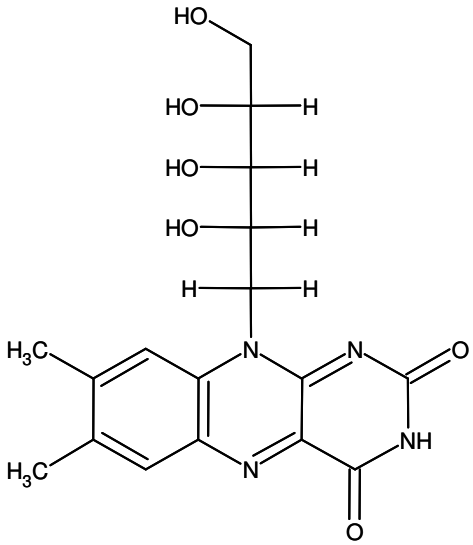
β−carotene



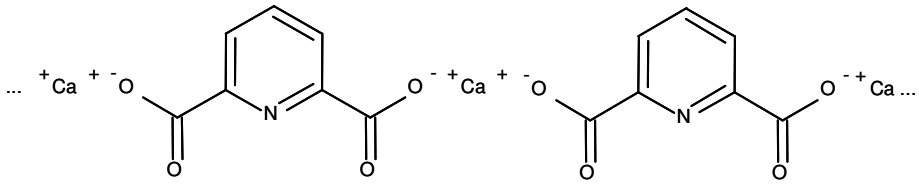
ascorbic acid



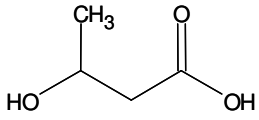
gluthatione



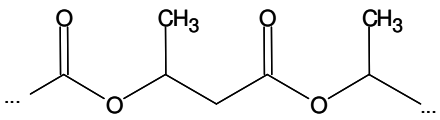
riboflavin



calcium dipicolinate



β-3-hydroxybutyrate



poly-3-hydroxybutyrate